BIRD SPECIATION IN THE GULF OF GUINEA

Martim Ferreira Pinto Pinheiro de Melo

A thesis submitted to the University of Edinburgh
in candidacy for the degree of Doctor of Philosophy

August 2006

"I have made this letter longer than usual, only because I have not had time to make it shorter."

Blaise Pascal (1623-1662)

ABSTRACT

The Gulf of Guinea island system, West Africa, constitutes a spectacular centre of bird endemism, with 33 species unique to the region. It comprises three oceanic islands (Annobón, São Tomé, Príncipe), one land-bridge island (Bioko) and one ecological island (Mount Cameroon), all part of the Cameroon line of volcanoes. This thesis used genetic, morphological and behavioural data from finches, white-eyes, thrushes and kingfishers to investigate: i) the importance of isolation for the speciation process; ii) the applicability of the current 'ecological model' of speciation, which was developed in parapatric and sympatric situations, to allopatric situations, and iii) the link between character divergence at the population level and the evolution of reproductive isolation.

Molecular phylogenies revealed that previous systematic assessments based on phenotypic characters were often incorrect. High levels of phenotypic differentiation of island taxa are not related to time since origin and can evolve within very short periods. For example, *Speirops* (Zosteropidae) and *Neospiza* (Fringillidae) are well within the genera *Zosterops* and *Serinus* respectively. 'Aberrant' characters did not constitute evidence for shared ancestry, e.g. the "genus" *Speirops* proved not to be monophyletic. The two Gulf of Guinea endemic kingfishers *Alcedo* spp. are island populations of the mainland Malachite kingfisher *A. cristata*, rather than being distinct island species. In contrast, molecular evidence in combination with phenotypic data strongly supported the elevation of the thrush population from Príncipe *Turdus olivaceofuscus xanthorhynchus* to species status. Molecular data also revealed a possible case of cryptic speciation within the Príncipe white-eye *Zosterops ficedulinus.*

The high number of endemic bird species in the Gulf of Guinea islands was the result of recent speciation events rather than the accumulation of relict species extinct on the mainland. Therefore, the Gulf of Guinea constitutes a very important centre of bird speciation in Africa. Because the oceanic islands are surrounded by one of the richest centres of biodiversity in the world they were colonised by several bird

groups which, by occupying different niches, reduced the possibilities of radiations within the archipelago. Therefore, most species originated by diverging in isolation from their source populations (allospeciation). Data from the Príncipe seedeater *Serinus rufobrunneus* showed that selection rather than drift was the main driver of divergence in allopatry, thereby supporting the applicability of the ecological model of speciation to allopatric situations. At the same time, the most divergent species were those that speciated after establishing sympatry with related populations, providing strong evidence for the importance of secondary contacts in promoting phenotypic diversification and speciation. Molecular evidence suggested that the São Tomé grosbeak *Neospiza concolor* may have speciated in full sympatry – which if confirmed would make it unique among birds.

In *S. rufobrunneus*, mate recognition traits were the first to diverge and may therefore be implicated in the first stages of the speciation process. This was further supported by playback experiments showing that populations no longer recognise the songs of foreign populations, suggesting that reproductive isolation may evolve as a by-product of independent divergence of mate recognition systems in allopatry.

Overall, this study supported the view of speciation as a selection-driven process more likely to be completed in sympatry after an initial period of isolation. This is likely to constitute the most general model of speciation.

RESUMO EM PORTUGUÊS / PORTUGUESE SUMMARY

O Golfo da Guiné, na África Ocidental, constitui um centro de diversidade biológica espectacular. Este sistema é constituído por três ilhas oceânicas (Annobón, São Tomé, Príncipe), uma ilha continental (Bioko) e uma ilha ecológica (Monte Camarões), todas parte da linha de vulcões dos Camarões. As ilhas oceânicas, separadas por mar com uma profundidade acima dos 1800 m nunca estiveram ligadas entre elas nem ao continente, enquanto que Bioko separado da costa africana por 32 km e profundidades inferiores a 60 m foi inúmeras vezes uma península de África. O grau de endemismo é particularmente extraordinário nas aves, com 33 espécies únicas a esta região. Este endemismo está concentrado nas três ilhas oceânicas que, com uma área total de cerca de 1000 km^2, têm 30 espécies endémicas. Em comparação o famosíssimo arquipélago das Galápagos com 13 ilhas com uma área total de 8000 km^2 tem 22 espécies de aves endémicas. Na realidade, em relação ao seu tamanho, o número de aves endémicas das duas maiores ilhas oceânicas do Golfo da Guiné (São Tomé e Príncipe) não tem paralelo no mundo (ver Fig. 1.1).

Esta tese procurou compreender os motivos desta riqueza. As aves do Golfo da Guiné servem assim como modelo para abordar um dos problemas mais fascinantes de biologia: a formação de espécies ou especiação. Para que uma espécie nova se forme é necessário em primeiro lugar que populações da mesma espécie se diferenciem (em morfologia, comportamento...) e, em segundo lugar, que consigam manter essa diferenciação. Tal só é possível se as populações divergentes não se reproduzirem entre si (o que eliminaria as diferenças). A evolução de isolamento reprodutor é por isso o ponto chave no processo de especiação. Durante a maior parte do século 20, considerou-se que o isolamento reprodutor só poderia evoluir em populações separadas por barreiras geográficas. Neste modelo clássico, sobretudo desenvolvido por Ernst Mayr, o isolamento geográfico permite mudanças aleatórias das frequências genéticas das populações isoladas (deriva genética). Com um fundo genético diferente não só as diferentes populações vão responder de forma distinta ao meio que as rodeia como a probabilidade de produzirem híbridos inférteis caso se voltem a encontrar aumenta. Nos últimos anos o interesse pelo problema de

especiação ganhou um novo fôlego graças à acumulação de dados que sugerem que a formação de espécies não requer isolamento geográfico. Um enorme volume de trabalhos deu origem a um novo modelo onde a formação de espécies é um processo dirigido por factores ecológicos (selecção natural e sexual) e onde a deriva genética não tem qualquer papel. Neste modelo, a formação de espécies é explicada pelos mesmos processos que permitem a adaptação das populações ao seu ambiente. No entanto, para demonstrar que o isolamento geográfico não é necessário, este modelo foi desenvolvido em situações em que as populações estão em contacto. É por isso importante testar se este modelo é generalizável às situações em que as populações estão isoladas (como em ilhas), ou se aqui factores aleatórios são importantes – como proposto por Mayr.

Este estudo combina dados genéticos, morfológicos e comportamentais de pica-peixes (Alcedinidae), tordos (Turdinae), olho-brancos (Zosteropidae) e canários (Fringillidae) para investigar: i) a importância do isolamento geográfico para a formação de espécies; ii) se o modelo de especiação actual aplica-se também para a formação de espécies a partir de populações isoladas; iii) o elo entre diferenciação de populações e a evolução de isolamento reprodutor.

A história das diferentes espécies analisadas neste estudo foi reconstruída através de análises genéticas. Técnicas moleculares permitem uma leitura directa do material genético, o ADN. Visto que este material contem a informação para a construção dos indivíduos, quanto mais próximos forem dois indivíduos mais semelhante é o seu ADN. Do mesmo modo, quanto mais próximas forem duas espécies mais semelhante é o seu ADN. Este principio é utilizado para a construção de filogenias que representam as relações entre espécies (por exemplo, Fig. 2.5). Além disso, com um conhecimento da velocidade a que o material genético muda (taxa de mutações) é possível inferir aproximadamente as datas em que as diferentes espécies se formaram.

As filogenias moleculares mostraram que as hipóteses existentes, baseadas na morfologia, sobre a relação das diferentes espécies estavam quase sempre erradas.

Assim, as espécies que apresentam um nível de diferenciação morfológica elevado não são espécies antigas como se pensava. Nas ilhas os caracteres morfológicos podem evoluir muito rapidamente. Por exemplo, as espécies do género *Speirops* (olho-grosso; Zosteropidae) e o *Neospiza* (enjolo; Fringillidae) fazem afinal parte dos géneros *Zosterops* e *Serinus* respectivamente, sendo espécies recentes. Também demonstrou-se que os caracteres aberrantes de muitas espécies insulares não são indicativos de parentesco, mas podem evoluir várias vezes em grupos diferentes. É o caso das quatro espécies que tinham sido incluídas no género *Speirops*: as espécies de Bioko e do Monte Camarões não são aparentadas com as espécies de São Tomé e Príncipe. Os dois pica-peixes (ou conóbia; Alcedinidae) que eram considerados espécies endémicas são afinal populações da espécie que ocorre no continente *Alcedo cristata*. Por outro lado, a população de tordo do Príncipe, que era considerada a mesma espécie que o tordo que ocorre em São Tomé, é claramente uma espécie distinta, tanto a nível genético como morfológico. Esta nova espécie (*Turdus xanthorhynchus*) só ocorre nas florestas do sul do Príncipe e em densidades muito reduzidas, sendo por isso necessário estudos para avaliar o grau de ameaça de extinção que possa correr. Os dados genéticos sugerem também que a população de olho-branco do Príncipe (*Zosterops ficedulinus ficedulinus*) constitui uma espécie nova, diferente do olho-branco de São Tomé (*Zosterops ficedulinus feae*) apesar de morfologicamente ambas serem muito semelhantes. Tal como o tordo do Príncipe, o olho-branco do Príncipe é uma espécie rara.

O elevado número de espécies de aves endémicas do Golfo da Guiné é o resultado da formação recente de espécies nas ilhas e não o resultado da acumulação de espécies que se extinguiram no continente (as ilhas funcionando neste caso como um refúgio durante as alterações climáticas ligadas aos ciclos das glaciações). Assim, o Golfo da Guiné constitui um centro de diversificação de aves de enorme importância em África. A localização geográfica das ilhas oceânicas, rodeadas a norte e a este por um dos centros de biodiversidade mais ricos do mundo, permitiu que fossem colonizadas por espécies de grupos diversos. Estas espécies, ocupando diferentes nichos ecológicos, reduziram a possibilidade de radiações dentro do arquipélago – o processo através do qual várias espécies se formam a partir de uma, com cada nova

espécie adaptada a recursos ecológicos diferentes (como o famoso caso dos tentilhões das Galápagos: onde a espécie que chegou às ilhas deu origem a 13 novas espécies). Assim, no Golfo da Guiné a maior parte das espécies originou-se por divergir em isolamento das populações colonizadoras. Dados do canário-castanho de São Tomé e Príncipe (ou padê ou chota-café, *Serinus rufobrunneus*) mostraram que a divergência das populações isoladas é explicada sobretudo pela selecção natural e não pela deriva genética. Este resultado sugere por isso que o modelo de especiação actual é generalizável aos vários contextos geográficos onde a especiação tem lugar. No entanto, este estudo mostra também que as espécies com características mais distintas são aquelas que completaram o processo de especiação depois de voltarem a encontrarem-se com a população da qual se separaram inicialmente. Estes contactos secundários têm assim um papel importante na diversificação morfológica e na formação de espécies. Finalmente, os dados genéticos sugerem, com um grau de confiança elevado, que a origem do enjolo (*Neospiza concolor*) ocorreu sem isolamento geográfico – se isto vier a ser confirmado será um caso único nas aves.

No canário-castanho de São Tomé e Príncipe os caracteres utilizados para a atracção e reconhecimento de parceiros (canto e cor) são os primeiros a divergir, podendo por isso participar no processo de especiação logo desde o inicio. Esta possibilidade foi confirmada por experiências no terreno que demonstraram que cada uma das populações desta espécie (São Tomé, Príncipe e Boné de Jóquei) já não reconhece o canto das outras populações. Em populações isoladas, o isolamento reprodutor poderá assim evoluir como consequência de mudanças independentes dos sistemas de atracção sexual.

Em sumário, este estudo sugere que a formação de espécies é favorecida pelo contacto entre populações que começaram o processo de divergência em isolamento geográfico. Após este encontro é a população imigrante (mais pequena) que vai divergir mais a nível morfológico. Factores ecológicos são a força principal durante todo o processo (em isolamento e em contacto). Este constituirá provavelmente o modelo mais geral de formação de espécies.

GENERAL ACKNOWLEDGEMENTS

This thesis was only possible with the generous help of an extraordinary number of people. This made it a very special project for me. Help has come in all sorts and I must have certainly forgotten to mention someone below. My apologies to those I have not named.

~~~

Firstly, I would like to thank Peter Jones, my principal supervisor, for his help and enthusiasm in all stage of the process. This PhD was driven by my enormous interest for the Gulf of Guinea. Peter shared this same interest. It was a great pleasure to build this PhD on the foundations so clearly laid out by Peter (Jones & Tye 2006). Peter's fabulous ornithological library was always a pleasure to explore and certainly saved me many trips to the Edward Grey Institute together with many inter-library loans… The fact that Peter revised my manuscripts almost before I had finished them is also greatly appreciated. As were the tapenades and other delicacies that he would bring regularly from southern shores.

Graham Stone, my co-supervisor, provided invaluable advice at key points of the thesis – making sure I knew what I was doing.

Josephine Pemberton welcomed me in her lab, and always took care that I had everything I needed. It was a great pleasure to work in such a technically advanced lab. I also appreciate Josephine's interest in the progression of my thesis.

Claire Doutrelant (CEFE-CNRS, Montpellier) taught me everything about bird song and recordings and received me each year in her lab and home. She was also the main force behind the analysis of bird coloration and supervised the 'honours' (Maîtrise) project of Geoffroy Colin on this subject (see 'collaborations'). Claire came also to help and provide guidance in the field. Travelling to France to stay with Claire, Rémy, Maya, and, in the last visit, Cyane was always a highlight of the year. Merci mes beaux amis et à tout de suite!

Claire and I thank the expert help of our friend Albertine Leitão (Max-Planck Institute) with song analyses, and of Staffan Andersson (Gottenbörg) with colour analyses. (I am sorry Staffan that you left the forest one day before the grosbeak was captured; we will have to go back).

Bengt Hansson (Edinburgh and Lund) and Thomas Smith (Center for Tropical Research, UCLA), played the very important role of non-official supervisors. They were always ready to discuss ideas and answer my queries, and revised proposals and manuscripts. While doing so they played also the vital role of injecting renewed enthusiasm into the project in those periods when one no longer sees the whole picture.
~~~

I also would like to warmly thank Tom Smith for my visit to his lab in UCLA in October 2005. Tom turned it into an incredible brainstorming session. And it was great to meet after 9 years of correspondence: it all started before I travelled to the Gulf of Guinea for the first time in 1996. I arrived on São Tomé, where I lived for a year, with a pack of reprints from Tom. Most were on the evolution of polymorphisms. Reading them while being surrounded by endemic birds goes a long way in explaining the origin of this thesis.

I sincerely thank all members of Tom's lab who took time to discuss many issues with me and present me their interesting work. John McCormack, pursuing similar interests, gave me his PhD proposal, which clarified a lot of things. It was also an inspiration for my general introduction. Tom Dietsch pointed out to me an important flaw of my study design (comparing divergence between habitats of very different ages); Ravinder Sehgal and Camille Bonneaud gave me useful advice on the study of blood parasites in birds that I will be pursuing after this thesis. John McCormack and Gabriela Castañeda were incredible hosts – I appreciated immensely the time spent together. Thank you also to Brenda Larisson for having me at her and Tom's place.

Many people helped in various aspects of lab work and genetic analyses. Apart from everyone mentioned above I would like to thank: Alex Hayward, Richard Chabblis, Rauri Bowie, Jérôme Fuchs, Mark Pagel, Andrew Rambaut, Ben Warren and, from the Workshop on Molecular Evolution at CCMAR (Algarve), David Swofford, Mark Holder and Gavin Naylor. Sonya Preuss generously lent me her computer for running phylogenetic searches for days on end.

Robert Prys-Jones (British Museum of Natural History) welcomed me at Tring, where I could handle birds collected by the 19th century naturalists, including the only museum specimen of São Tomé grosbeak in the world.

Richard Ranft (British Library National Sound Archive) provided me the recording material for the study of the Príncipe seedeater song.

Practical life at the Institute was made easy with the help of Carole Ferrier, Jayne Glendinning, and David Walker.

~~~

Thomas Smith and Marc Languy provided all the necessary advice and contacts needed to organise a successful field trip to Cameroon. Roger Fotso welcomed me and organised the research permit. Fieldwork in a new ornithological world was only possible with the knowledge and commitment of Francis Njie, to whom I am very grateful. Okah Monya provided very useful advice on where to carry the work. At the Limbe Botanic and Zoological Garden, Dr. Ndam welcomed me and gave me permission to capture birds in the garden, and Joe Mbelle was always available to help in any logistic issue even before my arrival in Cameroon. Thanks to everyone that helped in the field, namely Dan, Max, and Anthony. I would like to thank Dan and his family in particular for hosting me in their home.

~~~

~~~

The work in São Tomé and Príncipe was made under the permission of the Ministry of the Environment and Natural Resources. I particularly want to thank Eng. Arlindo Ceita de Carvalho not only for providing the permit, but above all for his kindness and his keen interest in the progress of my work. Fieldwork could only be carried out with the invaluable logistic support of ECOFAC-São Tomé. Thanks to Eng. Salvador Sousa Pontes for establishing such a close collaboration. I would like to thank the indefatigable Aurélio and Guilhermino for all the travels back and forth to the Bom Sucesso Botanic Garden and to most remote parts of the island – often at very inappropriate times. I thank them their good humour and friendship. At 1100 m, Bom Sucesso was the perfect place to stay. I would like to thank its keeper, Lagoas, for his help and company. I would also like to thank the good humour brought daily by Estevão, Francisco, Aurélio, Fátima and Lúcio.

If I managed to capture birds from species that were seen only for a handful of times in the last 100 years it was thanks to the professionalism and enthusiasm of the field assistants. On São Tomé, I thank Sr. Pedro, Luis Mário and Lúcio, the three best guides on the entire island. On Príncipe, I thank 'the one and only' Bikegila. His commitment had no parallel: he wouldn't be happy until I got the birds I wished for. Failure was not allowed, and often I saw Bike climbing 15m tall trees to set nets where he thought to be a good place. And he was always right. On top of this, Bike was a storyteller a par with Sr. Pedro (i.e. the best there is) and the long discussions we had were always illuminating. Octávio offered his meticulous help whenever he could, both on São Tomé and on Príncipe. The periods spent on the terrain that the collector José Correia described in 1928 as the "bad among the bad or worse among the worse" were the best times of this thesis. Thank you all for your friendship.

São Tomé and Príncipe is a home away from home not because it has beautiful forests and spectacular beaches but because of all the friends I have there. Angus Gascoigne, 'our man in São Tomé', has always given us all help we needed and even more. Angus, Nanda, Jaime and, in the last year, Kezia, have taken Rita and myself and all our fieldwork chaos under their roof for months on end. We are extremely grateful for their help. Staying with Angus had the added bonus of long discussions on Gulf of Guinea biodiversity. Lúcia, Paula e Sameiro from the NGO 'Leigos para o Desenvolvimento' were extremely generous in providing a base in São Tomé for the first 6-month field season. Thanks also to Quaresma and Noémia for their care. I am very grateful for Lúcia's friendship along the years and her help in al sort of situations. And of course, in São Tomé, I can never thank enough Octávio, Aldinha and the rest of my Alto Douro family. It was so good to rest under your care of laughter, dancing, swims, and being treated to the best food on the island. I did learn how to do a fast-food version of Kalulú – the one that takes only 12 hours to prepare. In Príncipe, I thank again Bikegila and Dáda and their family for their friendship and care, and Zôzô for her beautiful smile. I also thank them for the best food on the island... I did not learn all the steps to cook the Azagoa, although I now have a list of 54 out of the 80 different kinds of leaves it needs...

~~~

On my stopovers in Gabon, Patrice Christy was always available to help and spend some time together for some good conversation.

~~~

That any fieldwork at all has taken place in Equatorial Guinea is the sole merit of Noélia Zafra Calvo, technical assistant from the University of Alcalá de Henares (Spain) in charge of the implementation of an institute of biodiversity and environment in the University of Equatorial Guinea (IUBioma). She established all the contacts with the University required to make the visit possible. Once in the country Noélia went out of her way to ensure that everything worked. It was an awful lot of work required to allow me to capture three species. I am extremely grateful to her. At the Universidad Nacional de Guinea Ecuatorial (UNGE) I would like to give my sincere thanks to the chancellor Carlos Nse Nsuga, and to the Director of the Escuela de Estudios Agropecuarios, Pesca y Forestal "Obiang Nguema Mbasogo", José Manuel Esara Echube. José Ondo Nguema, Director of the Department of Forestry of UNGE, joined us for the entire fieldwork. I am sure this must have strained his personal life and I am very thankful to him: his interest in the study and his guidance in his country were essential for carrying out the work. Apolinar Abaga Obiang, lecturer at UNGE, helped us in Pico Basilé and Moka. Estebán was our guide in Moka (Bioko) and Gervásio in Annobón. Javier García Francisco provided good advice for fieldwork in Bioko and gave me very useful maps. Mama Rufina welcomed Rita and me into her house in Annobón without knowing us from anywhere. I'll never forget that she kept enough flour to make us bread rolls for breakfast each morning. If you manage to see Annobón on a map, you will understand why this is special. If you don't manage to see it on a map, you will understand it even better.

~~~

It's a bit redundant to thank my friends for… being my friends, so I'll make it short. Thank you for keeping me happy. In Edinburgh: Mathieu, Greg, Marianne, Sasha, Elise, Cedric, Alex, Roberta… In Portugal: Alexandre, Júlia, João, and vocês todos, and my sister Helena and my brothers Vicente and Bernardo. In France: Mixa e Zé. In Australia: Trish. In South Africa: Carlos & Wendy and everyone else…

~~~

When I told my father I was very interested in birds (I was 9-years old) he started taking me bird watching. When I was 14-years old, my mother overrun my timidity and 'pushed' me into the Nature Conservation Institute: the next day I was ringing waders overnight in the Tagus estuary (an absolute dream). Since then things never stopped. I thank very much my parents for their support and for trusting me in my explorations – the fundamental condition for any exploration to happen.

~~~

It follows that I must also thank everyone at the Nature Conservation Institute that took on board that 14-year old kid: Rui Rufino, Renato Neves and António Araújo, and later on Júlia Almeida and José Pedro Granadeiro. I could not believe I was doing all those things. Capturing rare grosbeaks in 2003? The art needed to achieve such 'feat' was learnt back then.

~~~

Finally – Rita. Rita was with me during all stages of this project. It was great to make this long road together. Together, the great times were fabulous, the shared adventures amazing, and the difficult times soon overcome. I am looking forward for the next adventure, this time with Francisca - an adventure in herself - coming along.

~~~

CHAPTER COLLABORATIONS

Chapter 2 – Diversification of the white-eyes

Ben Warren (University of Reading) provided me with the unpublished sequences from his study on the Indian Ocean white-eyes.

Chapter 3 – Origins of the São Tomé grosbeak

The sequences used for the phylogeny of the genus *Serinus* were obtained by Rauri Bowie (University of Stellenbosch).

Chapter 4 – The Gulf of Guinea thrushes

This chapter used a previously published thrush phylogeny (Bowie *et al* 2005) and Rauri Bowie performed the maximum parsimony and maximum likelihood searches. Measurements from museum specimens were provided by Nigel Collar.

Chapter 5 – The Gulf of Guinea kingfishers

Jérôme Fuchs (Muséum National d'Histoire Naturelle, Paris) sequenced the second intron of the myoglobin and ran the Bayesian and maximum likelihood searches.

Chapter 6 – Population divergence in allopatry

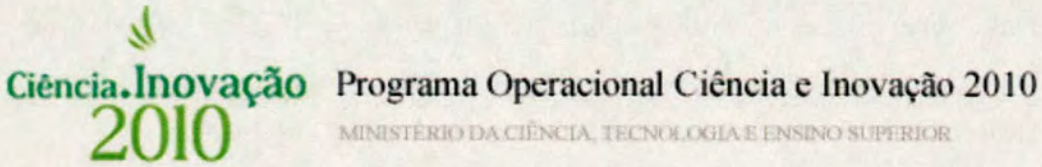
Bengt Hansson (University of Lund and University of Edinburgh) tested 74 microsatellite primer pairs on the Príncipe seedeater *Serinus rufobrunneus*. He also taught me everything I know about microsatellite-related lab work. Geoffroy Colin (Université de Montpellier II) measured and analysed the colour of the feathers for his 'Thèse de Maîtrise' under the supervision of Claire Doutrelant (CEFE-CNRS) and me. The study of song was done in collaboration with Claire: she taught me how to get field recordings, how to use the analytical software, and helped me in the painful and slow task of sorting song types by eye. She also obtained field recordings for this study. Rita Covas helped in the statistical analyses of the morphological data.

Chapter 7 – Playback experiments

Part of the song collaboration with Claire, who taught me how to design and carry out the experiments.

FINANCIAL SUPPORT

Most financial support for this thesis was provided by a PhD scholarship (SFRH / BD / 6396 / 2001) from the Portuguese research foundation (Fundação para a Ciência e a Tecnologia) sponsored by the program POCI2010 and FSE. This scholarship covered my stipend, university fees, and laboratory expenditures.
The collaboration with Claire Doutrelant was financed during two years under the Alliance program co-financed by Égide and the British Council-France.
The Genetics Society (UK) awarded me a £1,000 field expedition grant.



CONTENTS

1

SPECIES, SPECIATION, AND THE GULF OF GUINEA SYSTEM

WHY THIS THESIS?

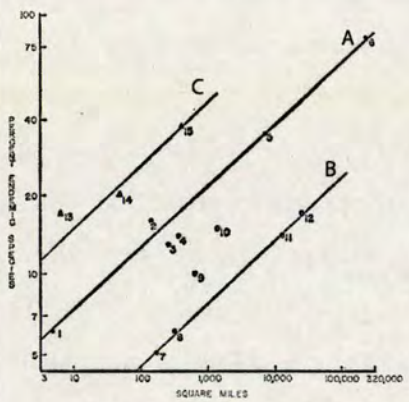
Understanding how new species arise – speciation – has been one of the most fascinating problems in biology ever since Darwin's theory of evolution. This problem was coined the 'mystery of mysteries' by Ernst Haeckel, an assessment subscribed to by Darwin both in his journal of the expedition on the Beagle (Darwin 1860) and in the preface of *The Origin of Species* (Darwin 1859), and indeed it has remained until today one of the most elusive problems in evolutionary biology. Despite almost 150 years of interest in speciation, only in the last two decades has a significant boom in speciation research occurred (Howard & Berlocher 1998; Magurran & May 1999; Barton 2001; Coyne & Orr 2004). In this period there was more work done on the subject than in the previous 120 years (Coyne & Orr 2004). Today it is one of the most dynamic and exciting areas of evolutionary biology, with the possibility of a real understanding emerging during the current decade (Mallet 2001). After a wealth of theoretical studies, significant advancements are expected to come from studies of natural populations (Turelli *et al.* 2001; Via 2002).

The aim of this thesis is to investigate the speciation processes in the origin of the bird diversity of the Gulf of Guinea island system. This system constitutes a spectacular centre of endemism (Jones 1994; Gascoigne 2004) that has been neglected by speciation researchers. It comprises three oceanic islands (Príncipe, São Tomé, and Annobón), one land-bridge island (Bioko) and one ecological island (Mt. Cameroon), all part of the Cameroon line of volcanoes. Birds are one of the groups for which endemism levels are truly impressive, with 33 endemic species present (Table 1.1). This represents almost one half of the endemics of the very large Guinean Forests' hotspot (Bakkar *et al.* 1999). As would be expected, most of the endemism is restricted to the oceanic islands. The three oceanic islands, with a total area of about 1,000 km^2, support 29 endemic species, with up to five endemic genera (Stattersfield *et al.* 1998; Jones & Tye 2006). For comparison, the Galápagos archipelago with 13 islands totalling c. 8,000 km^2 holds 22 endemic species, whereas the six main Hawaiian islands, with an area over 16,000 km^2, have currently 30

endemic species, with 19 extinctions documented in historical times (Stattersfield *et al.* 1998). The level of bird endemism in relation to area of the Gulf of Guinea oceanic islands appears to have no parallel, following a very distinct trajectory to the global pattern (Fig. 1.1). The average number of endemic birds of the 45 small islands of the world (area < 10,000 km^2) that have at least one endemic bird species is two, with the mode being one (Coyne & Price 2000). São Tomé, 857 km^2, has 16 single-island endemics and Príncipe, 139 km^2, has six (between them they share an additional 5 endemics). Hawai'i (10,000 km^2) has 16 described single-island endemics, nine of which went extinct in historical times. Mayr (1965) appears to have acknowledged the uniqueness of the Gulf of Guinea oceanic islands when he used three categories of islands in a global analysis of species accumulation with area: i) large isolated islands, ii) small islands within archipelagos, and iii) Gulf of Guinea islands (Fig 1.3). Interestingly, the identification of a single island system as a category apart, distinguished by its higher species richness, did not seem to elicit curiosity. An additional feature of the Gulf of Guinea system is that the impressive level of bird endemism is spread across different families rather than being concentrated within species-rich genera, as is the case with the most studied systems. Therefore, the Gulf of Guinea does not offer spectacular radiations such as those of the Darwin's finches in the Galápagos (Grant 1986) and the Hawaiian honeycreepers (Pratt 2005) – and this is likely to be one of the reasons why this system has been overlooked by speciation researchers. Instead, the Gulf of Guinea does offer many valuable phylogenetically independent replicates, creating a unique opportunity to test hypotheses about evolutionary processes at the levels of both speciation and adaptation.

Considering the evolutionary significance of this region this project will provide crucial information to support conservation efforts and promote awareness of a system of global importance. Each of the oceanic islands is listed by BirdLife International as an independent Endemic Bird Area (EBA) of Global Conservation Significance, whereas Bioko and Mount Cameroon are part of the Cameroon mountains EBA which, with another 29 endemic species, is the third most important in Africa (Stattersfield *et al.* 1998). São Tomé and Príncipe are the only small islands

in the world to make it to the top 25% of EBAs. The southwest forests of São Tomé, where all endemics can be found, were considered the second most important forests in Africa, including Madagascar, for bird conservation (Collar & Stuart 1988).



Double logarithmic plottings of area against percentage of endemic species of birds on three kinds of islands. (*A*) Solitary, well-isolated islands: Lord Howe, 1; Ponape, 2; Rennell, 3; Chatham, 4; New Caledonia, 5; Madagascar, 6. (*B*) Single islands near mainlands or large archipelagos: St. Matthias, 7; Pemba, 8; Manus, 9; Socotra, 10; Timor, 11; Tasmania, 12. (C) Islands in the Gulf of Guinea: Annobon, 13; Principe, 14; São Tomé, 15.

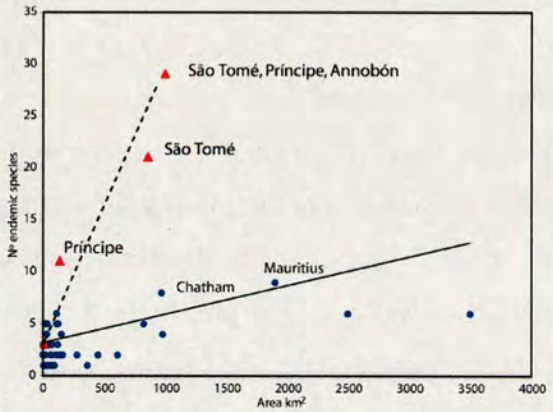


Fig. 1.1. The levels of endemism among the birds of the oceanic islands of the Gulf of Guinea appears to have no parallel.

Left: In a study of bird turnover on islands Mayr (1965) used the Gulf of Guinea oceanic islands as a category in its own right: the one with the highest proportion of endemism relatively to area.

Top: Coyne & Price (2000) compiled a list with the current number of endemic bird species of all small islands in the world (with at least one endemic). In their dataset the Gulf of Guinea oceanic islands are outliers following a very distinct trajectory.

This thesis uses genetic, morphological and behavioural data in an attempt to understand the factors behind the exceptionally high levels of diversification of the birds of this region. The guiding themes of this research programme are:

1) What is the role of isolation in speciation?

2) Does the current 'ecological speciation model' apply to isolated populations?

3) What is the link between divergence at the population level and speciation?

To understand the relevance of these research themes I first provide a background to the speciation problem. This is followed by a description of the Gulf of Guinea system. Finally, I introduce the specific aims of the thesis and a brief synopsis of each chapter contribution in addressing them.

Table 1.1. Endemic bird species of the Gulf of Guinea island system.

Sequence follows Jones & Tye (2006). *Alcedo* kingfishers are treated as full species by BirdLife International (Stattersfield *et al.* 1998). Islands - A: Annobón, B: Bioko, MC: Mount Cameroon, ST: São Tomé, P: Príncipe. Boné is a 30 ha islet 3 km from Príncipe. Status: global threat status (BirdLife International 2000) - nt: near threatened; CR: critically endangered; VU: vulnerable; remaining species are of 'least concern'. *: taxa included in this study. §: endemic genus. (§): *Ploceus sanctithomae* was formerly placed in endemic genus *Thomasophantes.*

Common name	Scientific name	Island	Status
Dwarf olive ibis	*Bostrychia bocagei*	ST	CR
São Tomé green pigeon	*Treron sanctithomae*	ST	
São Tomé bronze-naped pigeon	*Columba malherbii*	ST, P, A	
Maroon pigeon	*Columba thomensis*	ST	VU
São Tomé scops owl	*Otus hartlaubi*	ST	nt
São Tomé spinetail	*Zoonavena thomensis*	ST, P	
Príncipe kingfisher	*Alcedo (leucogaster) nais* *	P	
São Tomé kingfisher	*Alcedo (cristata) thomensis* *	ST	
Gulf of Guinea thrush	*Turdus olivaceofuscus olivaceofuscus* *	ST	nt
	xanthorhynchus *	P	nt
São Tomé prinia	*Prinia molleri*	ST	
São Tomé short-tail	*Amaurocichla bocagii* §	ST	VU
Annobón paradise flycatcher	*Terpsiphone smithii*	A	
São Tomé paradise flycatcher	*Terpsiphone atrochalybeia*	ST	
Bioko batis	*Batis poensis*	B	
Dohrn's thrush-babbler	*Horizorhinus dohrni* §	P	
Príncipe sunbird	*Anabathmis hartlaubii*	P	
São Tomé sunbird	*Anabathmis newtonii*	ST	
Giant sunbird	*Dreptes thomensis* §	ST	VU
Annobón white-eye	*Zosterops griseovirescens* *	A	
Príncipe white-eye	*Zosterops ficedulinus ficedulinus* *	P	VU
	feae *	ST	VU
São Tomé speirops	*Speirops lugubris* §*	ST	
Príncipe speirops	*Speirops leucophaeus* §*	P	VU
Fernando Po speirops	*Speirops brunneus* §*	B	VU
Mount Cameroon speirops	*Speirops melanocephalus* §*	MC	VU
São Tomé fiscal	*Lanius newtoni*	ST	CR
São Tomé oriole	*Oriolus crassirostris*	ST	VU
Príncipe drongo	*Dicrurus modestus*	P	
Príncipe glossy starling	*Lamprotornis ornatus*	P	
Príncipe golden weaver	*Ploceus princeps*	P	
Giant weaver	*Ploceus grandis*	ST	
São Tomé weaver	*Ploceus sanctithomae* (§)	ST	
Príncipe seedeater	*Serinus rufobrunneus rufobrunneus* *	P	
	thomensis *	ST	
	fradei *	Boné	
São Tomé Grosbeak	*Neospiza concolor* § *	ST	CR

BACKGROUND

Species and speciation

The speciation problem, and its mystery, originates directly from the theory of evolution, which established that all living organisms that ever lived on earth are connected by unbroken chains of descent. The problem of speciation consists therefore in explaining how one species can split into different ones, i.e, how can a continuous lineage be apparently 'broken' into discrete entities. It follows that one needs to define what a species is in order to be able to know when one species is formed, and therefore how it was formed. This could lead into endless digressions that could bring more obfuscation than clarity. A critical review of this issue is provided by Hey (2001), whereas detailed treatments of the different species concepts explained by their proponents can be found in Howard & Berlocher (1998).

The most fundamental theme underlying the debate is whether species are real entities or just artificial clusters of similar looking individuals. If species were artificial classes the field of speciation research would be meaningless. Obviously, I view species as real entities, or this PhD would not exist. Admittedly, my subscription to this view probably has components of personal belief and common sense (most people agree on what constitutes different species). I find particularly convincing the fact that indigenous knowledge systems and current systematics show a high degree of concordance in the definition of species, whereas little or no concordance occurs for other levels of biological organisation, like genera, families and orders (Coyne 1994). My perception of species as real entities was made much easier by the study group with which I am most familiar, birds, where the identification of 'different species' generally poses no problems: in any given region many discrete groups (the 'species') can be found without any intermediates in between. From a more scientific perspective, my main argument to support the reality of species derives from the concept of speciation. This might be an ontological sin, but I might be redeemed from it below. It might also explain why the proponents of species as artificial classes can also claim to study 'speciation'.

All species concepts view species as clusters of organisms that are more similar to each other (in whatever the dimension studied) than to individuals from other clusters, and with few or no intermediates when in contact. Speciation is defined as the process by which such independent clusters are formed. Such process has two components: structuring of variation within a cluster into different sub-clusters and maintenance of the integrity of the sub-clusters so that they do not form intermediates when in contact. Speciation involves the evolution and maintenance of polymorphisms. For some researchers, and possibly for all who maintain that species are artificial entities, the speciation process is limited to the evolution of polymorphisms. But evolution of polymorphisms *per se* does not lead to formation of stable discrete entities. Intermediates between polymorphisms can be continuously formed if polymorphisms are under frequency-dependent selection – for example, when selective pressures favouring each polymorphism are not present simultaneously but are linked to environmental oscillations (Grant & Grant 1995, 2002c), or when the selection pressure behind the polymorphisms is restricted to a small period of the year (Smith 1993, 1997; Smith & Girman 2000). Polymorphisms can also simply represent local adaptations, which might not persist if organisms disperse readily. In the extreme case, persistent mixing between polymorphisms can lead again to a monomorphic cluster. Therefore, the defining feature of the speciation process is the maintenance of polymorphisms. This can only be achieved by the evolution of reproductive isolation. This was the view put forward by the Modern Synthesis (Dobzhansky 1937; Mayr 1942) and, despite much debate, the evolution of a level of reproductive isolation necessary to keep the integrity of polymorphisms is still viewed as the crucial element for the speciation process. The most recent and exhaustive treatment of speciation equates the study of the speciation process to the study of the evolution of reproductive isolation (Coyne & Orr 2004). In this sense the Biological Species Concept (Mayr 1942, 1963), which defines species as independent breeding communities, is still the one used in practice by most speciation researchers.

Most of the 'species concept' debate is concerned with finding precise criteria or rules for species delimitation (see also: Mayden 1997; de Queiroz 1998, 2005; Hey

2006). If we agree with the above definition of speciation, this is an impossible task: because all species are 'connected by descent', one should expect 'grey areas' between closely related species, where polymorphisms are in place but reproductive isolation might still be weak. It is not because there are cases where we are unable to decide on what constitutes a species that species are not real. It is in this sense that I say that my best argument for the reality of species is the reality of the process of speciation: a process that generates independent evolutionary lineages (Mayden 1997; de Queiroz 1998, 2005).

A paradigm shift

The current boom in speciation research is associated with a paradigm shift, or at least with a change in perspective in the approach to the problem (Schilthuizen 2000). The two major speciation paradigms differ in the importance given to isolation for the speciation process to take place.

The paradigm that dominated the 20th century was the 'isolation paradigm'. This paradigm was already clearly developed by the beginning of the century (Jordan 1905), but only came into prominence with the Modern Synthesis. It was put forward by Dobzhansky (1937) and Mayr (1942), and developed by Mayr during the course of his career (e.g., Mayr 1963, 1992, 1999, 2000c, 2000b, 2000a). In this model, isolation, and not selection, is the crucial factor for the speciation process. Species are viewed as 'co-adapted gene complexes'. Interbreeding destroys each species' favourable gene combinations because recombination disrupts the epistatic and pleiotropic relationships of the genes within each species. For speciation to take place, the cessation of gene flow between two diverging populations must be complete, and this can only happen in geographic isolation. The isolation model is therefore also known as the 'geographic' or 'allopatric' model. The second important component of the isolation model is the importance of drift for the evolution of reproductive isolation. According to Mayr (1954, 1963), new gene complexes can arise only if there is a reorganization of the genome, and this requires the action of drift in the initial stages of speciation. Selection acts on the reorganized genome, leading eventually to a new 'co-adapted gene complex'. In this model, the need for

drift made speciation more likely to occur in small peripheral populations. This was supported by the observation that large populations separated by a geographic barrier show less differentiation than small isolated populations, like those on oceanic islands. Accordingly, Mayr developed the 'founder event speciation model' – where strong drift acting on bottlenecked populations could produce 'genetic revolutions' and hence reproductive isolation (Mayr 1954). This specific 'bottleneck model' has now been confidently refuted by both laboratory and fieldwork studies (Coyne & Orr 2004), but the importance of isolation for speciation to occur remains an open issue. The isolation model of speciation dominated the 20th century partially because it was in a sense 'trivial' and therefore did not elicit much interest (given enough time two isolated populations are expected to diverge to the species level), but also because evolutionary biology was mostly focused on understanding the origin of genetic variation within populations (Berlocher 1998).

The work of Guy Bush with *Rhagoletis* flies was the trigger for the current research boom (Bush 1969). His work suggested that speciation could occur in the face of gene flow. For that to be possible, the homogenising effects of recombination must be overcome (Felsenstein 1981b). A huge theoretical and empirical effort followed to investigate whether speciation could occur in sympatry (Via 2001). As Darwin hypothesised (Darwin 1859), a solid body of work demonstrated that natural selection can commonly create polymorphisms without isolation, when intra-specific competition is high and inter-specific competition is low (Smith 1987; Smith & Skúlason 1996; Langerhans *et al.* 2003). Demonstrating the evolution of reproductive isolation in sympatry (something that Darwin did not address) has proved more difficult. Reproductive isolation between sympatric morphs requires the evolution of strong assortative mating which is dependent on the evolution of a strong association between the adaptive and assortative mating traits. Several empirical studies have supported this 'divergence-with-gene-flow' model (Rice & Hostert 1993; Schliewen *et al.* 1994; Smith *et al.* 1997; Gíslason *et al.* 1999; Schneider *et al.* 1999; Hendry *et al.* 2000; Rundle *et al.* 2000; Ogden & Thorpe 2002; Richmond & Reeder 2002; Smith *et al.* 2005), but the debate is still open in part because one cannot always exclude the possibility of past allopatric phases

during the speciation process (Coyne & Orr 2004). Even if the question of 'sympatric speciation' is not settled, the research needed to tackle it was a major contribution to the emergence of a new speciation paradigm. The current paradigm views ecological factors as the primary cause of speciation. Selection (both natural and sexual) is now considered a much more powerful driver of speciation than random factors, even in the case of isolated populations (Rice & Hostert 1993; Orr & Smith 1998; Morell 1999; Mallet 2001; Schilthuizen 2001; Schluter 2001; Wu 2001; Fitzpatrick 2002; Kirkpatrick & Ravigné 2002; Coyne & Orr 2004; Rundle & Nosil 2005). This view can be labelled the 'ecological' or 'selection' paradigm.

A consequence of the emphasis on ecology rather than isolation was that it re-established the possibility of a link between micro- and macro-evolutionary processes. In the 'isolation paradigm' adaptation and speciation are decoupled, whereas in the 'selection paradigm' they are part of a continuum: selection that is responsible for anagenetic changes (gradual changes within a lineage, i.e., adaptation) is also responsible for cladogenetic changes (branching of lineages, i.e., speciation). While Mayr's model led to a geographic approach to the speciation problem, the current approach focuses on processes.

The road ahead

The ultimate aim of speciation research is to determine whether there is a general speciation model, applicable to all geographical contexts, or whether multiple speciation models are appropriate. The two current views were developed in distinct geographic situations: the 'isolation model' is restricted by definition to cases of allopatry, while the 'ecological model', was developed for parapatric and sympatric situations, as the objective was to demonstrate the strength of selection even in the face of gene flow. One needs to identify the major unsettled issues in each model, develop hypotheses, and test them in the different geographic contexts.

The two speciation models are not opposite sides of the same coin. The current acceptance of the 'selection' model does not allow for a total rejection of the 'isolation model', and even Mayr postulated natural selection as the major force

driving the divergence of isolated populations (Mayr 1963, p. 257). The isolation model requires full cessation of gene flow and the combined effect of drift and selection for speciation to occur. The 'selection school' established with confidence that selection is a much more important force in speciation than genetic drift. This conclusion holds even for speciation of allopatric populations. It has been shown that reproductive isolation of allopatric populations results in the majority of cases from the evolution of adaptations to different environments or of different adaptations for similar environments (Rice & Hostert 1993; Grant *et al.* 2000; Schluter 2001; Boughman 2002; Wong *et al.* 2004; Rundle *et al.* 2005; Tonnis *et al.* 2005; Baker 2006; Vines & Schluter 2006). In this latter case, speciation is expected to proceed more slowly. The Asian and American sycamore trees *Platanus* spp. do not show reproductive isolation after millions of years of geographical isolation (Stebbins 1950 *in* Schilthuizen 2000). Similarly, the two Galápagos warbler finches *Certhidea* spp. at the base of the radiation of Darwin's finches, have been separated for up to 2 Ma, but have not evolved pre-mating isolation, and post-mating isolation may also not be present (Grant & Grant 2002a).

The fundamental question is therefore whether isolation is required for speciation to occur. Evidence for niche conservatism, the tendency for species pairs separated by a geographic barrier to occupy similar niches, has been put forward as evidence that speciation is driven primarily by isolation rather than ecology (Peterson *et al.* 1999). This important argument has nevertheless two limitations: i) evidence for niche conservatism can only be demonstrated in cases where a geographic barrier separates similar habitats - the possibility of radiation into new niches is therefore limited; ii) the fact that isolation led to speciation does not invalidate the possibility that the evolution of reproductive isolation was driven by selection – the prediction being that the more similar the habitats, the longer the time for divergence and reproductive isolation to evolve. Another interpretation of niche conservatism is that it reflects an inherent inertia of species to adapt to new environments, which multiplies the opportunities for speciation in allopatry – since different habitats act as barriers (Wiens 2004; Wiens & Graham 2005). This perspective does not go against the 'ecological model' – rather it gives an extra role to ecology for speciation in

allopatry. In the first stage, adaptation to a given habitat isolates populations; in a second phase, selection drives the divergence of the isolated lineages (Wiens & Graham 2005).

The 'selection school' put ecology at the centre of the speciation process but did not settle the isolation problem, contrary to the commonly held view. It showed that polymorphisms can *evolve* with gene flow and that species integrity can be *maintained* with gene flow. What has not been demonstrated is whether species integrity (i.e. reproductive isolation) can *evolve* without isolation (cessation of gene flow) occurring at some stage. Hybridization between species might not destroy the polymorphisms if hybrids are less fit than the parental populations. Reduced hybrid fitness can result from intrinsic genetic incompatibilities or from extrinsic ecological or behavioural causes (hybrids being maladapted). In this situation, natural and sexual selection can lead to the evolution of assortative mating. It has been difficult to rule out whether the genetic differences underlying the incompatibilities (intrinsic or extrinsic) can evolve fully under gene flow, or if they need to diverge initially in allopatry (Coyne & Orr 2004). In this latter case, the evolution of assortative mating in sympatry is based in differences acquired in allopatry, under the process of reinforcement (Servedio & Noor 2003). Even the founder and most emblematic case of sympatric speciation, speciation by host-switching in the *Rhagoletis* flies, probably had an allopatric phase in its origin (Feder *et al.* 2003; Jiggins & Bridle 2004).

Currently, the weight of evidence supports a general speciation model where ecology is the main force driving population divergence, but where a period of isolation is required to trigger the speciation process by allowing some divergence to evolve unhindered by the homogenizing effects of gene flow. Speciation research must attempt to clarify whether isolation is necessary, even if temporarily, for speciation to occur and, because of the primary role of ecology, further explore the links between micro- and macro-evolutionary processes.

Oceanic islands and bird speciation

Oceanic islands are one of the most useful natural situations for the study of the evolutionary process, as acknowledged long ago by Charles Darwin and Alfred Russell Wallace (Whittaker 1998). Their usefulness arises from their inherent simplicity (systems with well-defined boundaries, generally small and with a depauperate biota) together with often being striking centres of evolutionary change and diversification (Grant 1998; Emerson 2002). Both these characteristics are particularly evident in the most isolated island systems, which are virtually independent from continental areas and where spectacular radiations have taken place (e.g., up to 1000 *Drosophila* species in the Hawaiian archipelago evolved from one or two species: Kaneshiro *et al.* 1995). Most of our understanding of evolution on islands is therefore based on studies from remote islands such as the Hawaiian (Wagner & Frank 1995; Craddock 2000) and Galápagos (Grant 1986; Arbogast *et al.* 2006) archipelagos. While studies in very isolated island systems will always be important, significant advancements in our understanding of the processes driving diversification are more likely to come from island systems closer to the mainland. The defining feature of such systems is the increase in the likelihood of secondary contact between insular and mainland populations. These more complex systems can play an important role in the clarification of the role of isolation in the evolutionary process, but still retain the advantage of being simpler than mainland systems. They also have a very significant advantage over isolated systems. Whereas isolated systems are dominated by a few species-rich genera, in those closer to the mainland species diversity is distributed across more families, because the likelihood of colonisation is higher. In this way, such systems offer many independent replicates for testing evolutionary hypotheses, something lacking from the study of extensive radiations. In the Gulf of Guinea, the moderate distances between the oceanic islands and the mainland (average 270km) have allowed colonisation by a wider array of species than is normally the case for very isolated archipelagos. The 29 endemic bird species of the three oceanic islands belong to c. 18 families (Jones & Tye 2006). In contrast, the 22 endemics species of the Galápagos belong to 6 families, with the 14 species of Darwin's finches descended from a single species (Sato *et al.* 2001).

A model for explaining the diversification of birds on oceanic islands was first developed for the Hawaiian honeycreepers (Drepanidinae), more than 100 years ago (Perkins 1901, 1903, 1913 in Grant 2001). This 'archipelago radiation model' proved to be correct in its main lines and has been refined ever since as more empirical data were gathered (Lack 1947; Grant & Grant 1998; Blondel 2000). Not only did it became the general model for bird speciation in general (Grant & Grant 1997a; Edwards *et al.* 2005) as, most interestingly, it also represents well the current state of knowledge on the speciation process as assessed above: selection is the driving force, but some temporary isolation is needed. Therefore, the model initially developed for Hawaiian honeycreepers may actually describe the most general sequence of events leading to speciation (Rundle & Nosil 2005). In this model, speciation is initiated by population subdivision (allopatry), which leads to selection-driven divergence. Cultural drift may also lead to divergence of mate recognition systems. In a second stage, the populations diverging in allopatry meet. They may either mix, ending the process of speciation, or they may co-exist. In this case, competition between the two populations will lead to further divergence resulting in ecological character displacement, i.e., the most similar individuals of each species suffer the strongest competition, such that extreme phenotypes are favoured by selection. Reproductive isolation evolves by the evolution of assortative mating driven by negative selection on intermediate phenotypes. This process is accelerated if resource use traits are also used for mate recognition (for example: individuals with similar body sizes tend to breed together, or individuals tend to breed close to the resources they use).

Considering the generality of the 'archipelago radiation model', it will be used in this thesis as a template against which the results are interpreted.

THE GULF OF GUINEA SYSTEM

A good introduction to the Gulf of Guinea system can be found in a special issue of Biodiversity & Conservation (Fa & Juste 1994). An excellent synthesis is provided by Jones & Tye (2006), who also summarise all current knowledge on the birds of the region. The Gulf of Guinea Conservation Group maintains a website pooling together all recent scientific work conducted in the region (http://www.ggcg.st). A photographic overview of the islands in this system is provided in Appendix 1.1.

Location and geographical limits

The Gulf of Guinea island system, as it is considered here, includes five islands: one ecological island, one land-bridge island, and three oceanic islands (Fig. 1.2). They are, from north to south:

Mount Cameroon: (4°13'N, 9°11'E). Ecological island: very large and isolated. Located in the Southwest province of the Republic of Cameroon, Mount Cameroon, with an area of c. 1750 km^2 (50 x 35 km), is the highest mountain in West Africa, reaching 4,095 m.

Bioko: (3°12'N - 3°47'N, 8°25'E - 8°56'E). Land-bridge island: lying on the continental shelf in shallow seas only 60 m deep, and only 32 km from the coast of Cameroon, it was connected to the mainland during glacial periods (Rohling *et al.* 1998), being last connected c. 11,000 years ago (Einsentraut 1965; Lambert & Chappel 2001). Formerly the Spanish possession of Fernando Pó, Bioko is now part of the Republic of Equatorial Guinea. It is some 2027 km^2 in area (roughly 35 km x 72 km) and its highest peak is Pico Basilé, at 3,011 m.

Príncipe: (1°32'N - 1°43'N, 7°20'E - 7°28'E). Oceanic island in seas over 1,800 m in depth. Formerly Portuguese and now part of the Democratic Republic of São Tomé and Príncipe. It lies 210 km SSW of Bioko and 220 km W of the African mainland. Its total area is 139 km^2 (c. 17 km x 8 km). The highest point is at 935 m (Pico do Príncipe).

São Tomé: (0°25'N - 0°01'S, 6°28'E - 6°45'E). Oceanic island in seas over 1,800 m in depth. Formerly Portuguese and now part of the Democratic Republic of São Tomé and Príncipe. It lies 150 km SSW of Príncipe and 255 km W of Gabon, and the Equator passes through the Ilhéu das Rolas at its southern tip. Its total area is 857 km^2 (47 km x 28 km), and its highest peak rises to 2,024 m (Pico de São Tomé).

Annobón: (1°24'S - 1°28'S, 5°36'E - 5°38'E) now a province of the Republic of Equatorial Guinea. It is the smallest and most remote, lying 180 km further to the SSW of São Tomé and 340 km from the mainland. Its total area is 17 km^2 (6 km x 3 km), and it rises to 700 m at Santamina.

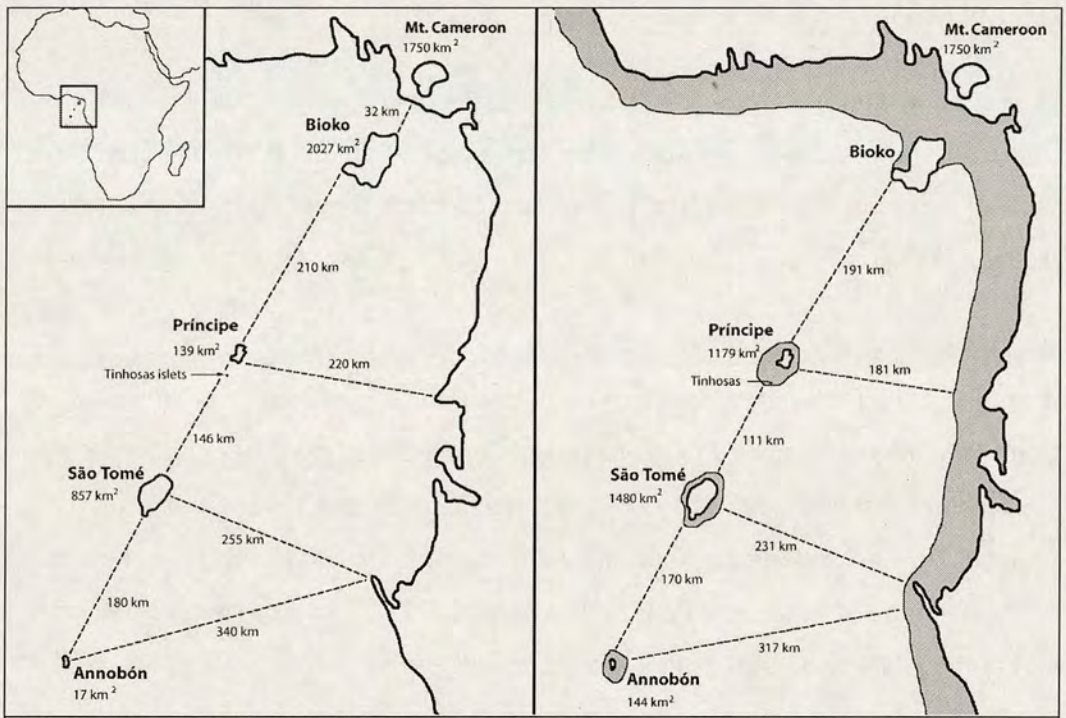


Fig. 1.2. The Gulf of Guinea island system, West Africa, showing island areas and distances between them and the mainland at the present day (left) and during glacial periods (right), This system includes an ecological island (Mt. Cameroon), a land-bridge island (Bioko) and three oceanic islands. Most endemism is concentrated in the two largest oceanic islands, São Tomé and Príncipe. Adapted from Jones & Tye (2006).

Climate

The region has an oceanic equatorial climate determined by the interaction between the southern monsoon winds from the Atlantic Ocean and the northern dry harmattan winds from the Sahara, which defines the Intertropical Convergence Zone (or meteorological equator). Mean temperatures are above 25°C at sea level, but decrease with altitude, falling below zero on Mount Cameroon. The predominant warm and moist southwesterly winds are intercepted by the high relief of the islands, creating a north-south divide in precipitation. Annual precipitation on the southwestern slopes of Mount Cameroon may be over 10,000 mm, compared with 1,500 to 2,000 mm in the northern slopes. On São Tomé annual rainfall on the southern forests may be over 7,000 mm, whereas it may not reach 600 mm in the north; maximum rainfall in Príncipe is 5,000 mm. No data are available for rainfall on Annobón, although it is probably less in accordance with its smaller size (Jones & Tye 2006). Precipitation over 11,000 mm is typical for the south of Bioko (Juste & Fa 1994). The year is divided into rainy and dry seasons. On Mount Cameroon and Bioko the main dry season is from December to March, and a shorter one occurs from July to August (Juste & Fa 1994; Borrow & Demey 2001). In São Tomé and Príncipe, the long dry season is from June to August and a shorter one extends for a few weeks in December and January; Annobón, south of the equator has a single dry season from mid May to the end of October (Jones & Tye 2006).

Habitats

The natural habitat type of the Gulf of Guinea islands has been described as rainforest (Exell 1944, 1973) or tropical moist broadleaf forests (Gascoigne 2004). The volcanic origin of the islands created a spectacular mosaic of rugged mountains with steep slopes, deep valleys, volcanic chimneys, tables, and huge waterfalls, which led to habitat diversification. The forest is stratified with altitude into lowland rainforest (0-800m), montane forest (800-1400), and mist forest (1400-2500), the latter being absent from Príncipe (Exell 1944). Curiously, Annobón appears to have the three forest types compressed in its small altitudinal range (max. 700m). A similar phenomenon occurs in the smaller southern peaks of São Tomé (Ogonovszky 2003). On Pico Basilé (Bioko) and on Mount Cameroon, subalpine meadows are

present over the forest limits. The rain-shadow effect over the northern parts of the islands most likely led to the establishment of a distinct habitat, probably resembling the dry forests currently confined to the northern gullies of São Tomé.

Geological history

The islands of the Gulf of Guinea system are the southern part of the 1000 km Cameroon Line of Tertiary to Recent volcanoes, which extends from Annobón to the Mandara Mountains on the Nigeria-Cameroon border (Burke 2001). This line runs in a NE-SW direction and includes six offshore swells (the four islands and two seamounts), and four continental swells (Mt. Cameroon, Manengouba, Mambotu, and Oku). Often, the Ngaoundéré and Biu swells are considered part of the line, which in this case is Y-shaped and 1,600 km long (Fitton 1987; Lee *et al.* 1994).

Volcanic activity in the continental and oceanic sector has been contemporaneous and more or less continuous since the Cretaceous (Fitton 1987; Lee *et al.* 1994; Burke 2001). Age progression in the line is only present in the oceanic sector: with the sub-aerial origins of Príncipe, São Tomé, and Annobón estimated at 31, over 13 and 4.8 Ma (Lee *et al.* 1994). The lack of a general age progression in the line and the long periods of activity of the different volcanic centres rules out a 'classic' hotspot origin for the Cameroon line (Lee *et al.* 1994), such as happened in the Hawaiian archipelago (Carson & Clague 1995). Also in contrast to the Hawaiian Islands, the elevation of the swells is determined more by lithospheric uplift than by erupted material (Burke 2001). The origin of all swells of the Cameroon line is linked to the same long-lived plume that feeds periodically a sub-lithospheric hot zone that has been expanding into the ocean. This plume may have been involved in the origin of the much older Nigerian granitic complexes (260 to 141 Ma) such as the Jos Plateau (Lee *et al.* 1994; Burke 2001). The activity that led to the origin of the Cameroon line of volcanoes began after the African plate rotated counter-clockwise c. 30-35 Ma, establishing a new platewide shallow-mantle convection system (Lee *et al.* 1994; Burke 2001). The rotation of the African plate probably displaced the Y-shaped hot-zone that would have underlain the Benue trough (or rift) into the current location of the Cameroon line (Fitton 1980). This would explain the peculiar

geological pattern in the region of having side-by-side 'a rift without volcanoes, and a volcanic line without a rift'. This complex history makes the Cameroon line a unique feature of the Earth (Burke 2001).

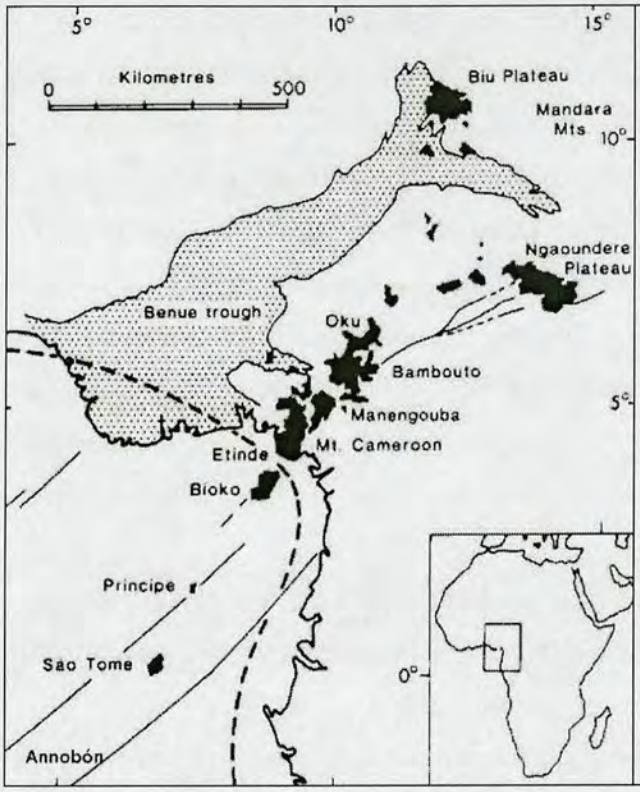


Fig. 1.2. The Cameroon line of volcanoes - from Lee *et al.* (1994). Notice the similarity in shape between the volcanic line and the Benue trough (or rift). This was likely due to the rotation of the African plate c. 30-35 Ma that displaced the athenospheric hot-zone underlying the Benue trough to its current position. All volcanoes derive from the activity of a single long-lived plume currently located in the Oku area. This plume periodically feeds a 'hot-zone' in the lithosphere that has been expanding into the ocean. The broken line marks the boundary between the continental and oceanic crust.

The ages of the oldest lava flows only give information on the minimum age when each island was sub-aerial since older rocks may be buried under the most recent ones. For example all the exposed lavas on Mount Cameroon are less than 1 million years old, with the mountain built upon much older lava flows (Fitton 1987). Furthermore, because volcanic activity persisted until recently on all islands, being still present in Mount Cameroon and to a lesser extent in Bioko, the impact of volcanic activity on the biota of these islands is unknown (but see: Jesus *et al.* 2005c). Therefore, when and how the land surface area (and hence colonisable habitats) changed through time is only hinted at by the ages available.

Beyond birds

On the oceanic islands, isolation led to the evolution of a highly original flora, characterised by the co-occurrence of neo-endemics and of Afromontane paleo-endemics (Figueiredo 1994). Príncipe has 26 single-island endemic angiosperms, São Tomé has 81 (including the endemic genus *Heteradelphia*, Acanthaceae), and Annobón has 14. From the 176 endemic species of the Gulf of Guinea islands, including the 40 single-island endemics of Bioko, only 16 are shared between more than one island, with only two shared with Bioko (Figueiredo 1994). This indicates that the floras of each island evolved mostly independently from each other. A recent biogeographic study on the begonias (Begoniaceae) suggested that São Tomé functioned as an important pre-Pleistocene refuge for these and possibly other plants (Plana *et al.* 2004). Ferns have been a successful group, with the forests of São Tomé and Annobón having the highest fern diversity and density in Africa (Exell 1944).

The Gulf of Guinea oceanic islands are also a striking centre of endemism for reptiles and amphibians. São Tomé and Príncipe harbour seven endemic amphibians – six frogs (Anura: five Hyperoliidae, one Ranidae) and one caecilian (Gymnophiona: Caecilidae) – and eight endemic reptiles (Squamata) – three burrowing-snakes (Typhlopidae), two geckos (Gekkonidae), two skinks (Scincidae), and one snake (Colubridae). Annobón has two endemic geckos (one shared with São Tomé and Príncipe), and one endemic skink. The reptiles and amphibians of the Gulf of Guinea islands pose some of the most difficult questions with respect to colonisation since many reptiles and nearly all amphibians are considered to have low tolerance of seawater (Balinsky 1981). Endemism for most species has been recently confirmed with molecular phylogenies (Drewes & Stoelting 2004; Drewes & Wilkinson 2004; Jesus *et al.* 2005a, 2005b), ruling out the possibility of anthropogenic introductions. Even shared species in different islands may represent independent colonisations from the mainland, rather than human-mediated dispersal events (Jesus *et al.* 2005a, 2005b; Jesus *et al.* 2005c). The endemic caecilian of São Tomé is the only known case of overseas dispersal in this group: caecilians are

largely subterranean amphibians that are considered highly unlikely to disperse over major marine barriers (Nussbaum & Pfrender 1998; Gower *et al.* 2002).

As is expected for oceanic islands, most endemic mammals are bats (Pteropodidae, Vespertilionidae, Molossidae, Hipposideridae: Juste & Ibañez 1994), but São Tomé has an endemic shrew (Soricidae), the São Tomé shrew *Crocidura thomensis*, and Príncipe has an endemic subspecies of Fraser's musk shrew *Crocidura poensis*. The shrews represent another 'zoogeographical mystery' (Dutton & Haftt 1996), because their metabolism and surface to volume ratio necessitates that they eat constantly, making them unlikely marine dispersers (Heim de Balsac & Hutterer 1982).

The Gulf of Guinea is the richest centre of endemism for coral-reef fish, supporting 20% of the world's endemic species and twice as many as New Caledonia, the second most important endemic area for this group (McAllister *et al.* 1994). On the islands most fish occurring in estuarine habitats are primarily marine, and all known freshwater fish are secondarily adapted.

Less data are available for invertebrates. Endemism is high among land snails, with c. 60 endemic species in the three oceanic islands, including one endemic family and six endemic genera (Gascoigne 1994). Among butterflies, São Tomé has 64 endemic species, and Príncipe 45, 28 of which are shared between them (Pyrcz 1992).

Human impact

Physical factors like isolation and area, and ecological factors such as habitat diversity, determine the potential of an island in terms of the biodiversity it can hold. Often it is the 'human factor' that determines the biodiversity levels currently observed – its impact ranging from causing the extinction of endemics to the introduction of new species. Whereas Bioko and Mount Cameroon have a long history of human settlement, the oceanic islands were uninhabited at the time of discovery. São Tomé was discovered by Portuguese navigators at the end of 1470. Príncipe and Annobón followed in early 1471. Colonisation started in the 1490s, mostly by slaves from the neighbouring mainland (today's Nigeria, Cameroon, and

Gabon), but also with Portuguese prisoners and, most infamously, with Jewish children separated from their parents in 1493 during the Inquisition (Usque 1553/1965).

The destruction of the lowland forest was fast: by 1529 São Tomé was the prime producer of sugar cane in the world. With the decline of sugar cane production in the 17th century, the islands' economy was based on supplying slave trading ships, en route from West Africa to America. The introduction of coffee and cocoa at the beginning of the 19th century led to further destruction of the forest, this time at altitudes up to 1200m. From 1908 to 1919 São Tomé and Príncipe were the world's largest cocoa producers. Production then declined and, after independence in 1975, a great number of plantations were abandoned, leading to the growth of secondary forest.

We will never know the full impact that the destruction of the lowland forest had. Nevertheless, no extinctions of endemic species have been documented and the current situation seems to be beneficial for biodiversity. Approximately 90% of São Tomé and Príncipe are still covered with forest, equally divided into primary forest, secondary forest and plantations (Jones & Tye 2006). Cocoa remains the most important plantation crop, which is particularly good considering its 'environmental-friendliness'. Requiring shade to grow properly, the small cocoa trees are planted below a canopy of tall trees to make a 'shade forest' whose structure mimics closely that of the original rainforest (Jones & Tye 2006). Annobón is mostly covered in forest except for the northern flat area where the only permanent human settlement is located. Nevertheless, agriculture differs from that of São Tomé and Príncipe in that it is often done right within the forest, following a rotational scheme of plots. Although this appears to have achieved for some time an appropriate balance between biodiversity and human needs, the last years have witnessed a fast degradation of the forest cover (Heras *et al.* 2002).

In these conditions, many of the endemic birds of São Tomé and Príncipe have adapted to the shade and secondary forests and are therefore still common (Atkinson

et al. 1991; Jones *et al.* 1992; Peet & Atkinson 1994). Nevertheless, research is needed urgently to assess how well the endemic birds are doing in the human-altered habitats. It has been shown that pesticide use negatively affects most of the insectivores (Atkinson *et al.* 1991). Obviously, conservation of the primary rainforest, where all endemics evolved, is crucial. It is the stronghold for most species, with eight species on São Tomé totally dependent on it. These include the dwarf ibis *Bostrychia bocagei*, the São Tomé fiscal *Lanius newtoni*, the São Tomé short-tail *Amaurocichla bocagii* and the São Tomé grosbeak *Neospiza concolor*, four species long thought to be extinct but rediscovered in 1990 and 1991 (Atkinson *et al.* 1991; Sargeant *et al.* 1992). In Príncipe, the three most elusive species are dependent on the primary forest: the subspecies of the olive ibis *Bostrychia olivacea rothschildi* (likely to be extinct), the subspecies of the Gulf of Guinea thrush *Turdus olivaceofuscus xanthorhynchus* (a distinct species according to this study: Chapter 4), and possibly a yet undescribed scops owl (Naurois 1975a; Melo 1998). The two endemic species of Annobón are common all over the island, but the recent degradation of the forest may threaten not only these species but also the entire ecosystem, which is highly dependent on the forest to capture and maintain water. The two endemic speiropses of Bioko and Mount Cameroon have stable populations, but due to their very restricted range are classified as 'Vulnerable' under IUCN criteria (BirdLife International 2000).

Mount Cameroon is the only region lacking protected areas; in Bioko, the entire range of the Fernando Pó speirops *Speirops brunneus* is within the Pico Basilé National Park (330 km^2), which is a doubly-protected area because it is a zone whose access is restricted for military reasons (Blom & Schipper 2004). In São Tomé and Príncipe, about one third of each island has recently been declared as Natural Parks (total of 293 km^2) covering all primary forests and mature secondary forests and including a buffer zone of secondary forests (Gascoigne 2004). The entire island of Annobón was also declared a protected area (Gascoigne 2004). If the implementation of all these protected areas can be assured, so can the future of the unique biodiversity of this system.

THIS THESIS

This thesis addresses the speciation problem in a natural system, using oceanic island birds as the study model. My enquiry spans several bird groups in an attempt to identify general factors behind the high levels of diversification of the Gulf of Guinea birds. The work was carried out both at the macro- and micro-evolutionary scales in order to investigate the factors promoting population divergence and those promoting speciation and how they are (or are not) connected. Ecological factors are predicted to be the determinant for population divergence, while speciation is expected to occur via a combination of isolation and ecological divergence.

The thesis is organised therefore in three components:

1. *Macro-evolutionary patterns of diversity*. Molecular phylogenies were inferred for several groups with the objective of setting the basis for any further investigation of the speciation process, to identify the main biogeographic factors behind the spectacular endemism levels, and to assess the importance of isolation in speciation. (Chapters 2-5)
2. *Micro-evolution*. Phenotypic patterns of diversity (morphology, colour, song) were described for the Príncipe seedeater *Serinus rufobrunneus* with the objective of determining what is driving divergence of populations, within and between islands. Population genetics methods were used to estimate the structure of and gene flow between populations in order to assess the importance of isolation during the divergence process. (Chapter 6)
3. *Link between micro- and macro-evolution*. A link between population divergence and the evolution of reproductive isolation was investigated by testing whether the divergence of the song of allopatric populations could constitute a reproductive barrier in the case of secondary contact. (Chapter 7)

Investigation of the biogeographic factors that promoted speciation in the Gulf of Guinea starts in **Chapter 2** with the reconstruction of the evolutionary history of the most speciose bird group in the region, the white-eyes (Zosteropidae). This is a

fascinating group of birds, as they appear to be adapted to colonise islands and speciate once there: most of the species of the family are island endemics, and no other passerine group has colonised so many islands. In the Gulf of Guinea this family exhibits a wide variety of phenotypes, something rare for this very homogeneous group. Using a molecular phylogeny, trends of phenotypic change are described and hypotheses for their causative factors are proposed.

In **Chapter 3**, I tackle the origins of one of the most enigmatic birds of the world, the São Tomé grosbeak *Neospiza concolor*. This is a bird that 'disappeared' for 101 years after its discovery in the late 19th century, and has been seen only a handful of times since then. The phylogenetic affinities of this monospecific endemic genus are uncertain. Initially considered a weaver (Ploceidae) it is now believed to be a finch (Fringillidae) related to the Príncipe seedeater *Serinus rufobrunneus*, a species endemic to São Tomé and Príncipe and to a small islet 3 km from Príncipe. The origins and phylogeography of *S. rufobrunneus* are therefore also inferred in this chapter, setting a firm basis for investigation of the processes driving divergence of its populations (Chapter 6).

Phylogenies were also inferred in two situations of very recent or incipient speciation, as such cases are particularly useful in identifying factors associated with the speciation process. **Chapter 4** uses both genetic and morphological data to assess whether the two very differentiated races of the Gulf of Guinea thrush *Turdus olivaceofuscus* (Turdidae) might already constitute full species, while **Chapter 5** looks at the relationships between the two endemic kingfishers (Alcedinidae) of São Tomé and Príncipe. The former question has important conservation implications since the race of Príncipe Island *T. o. xanthorhynchus* is extremely rare, and could qualify for the IUCN Critically Endangered status if it were a full species.

The São Tomé and Príncipe seedeater, endemic to three islands, is used as a model to investigate the diversification process of allopatric populations in **Chapter 6.** Gene flow between populations was estimated with 14 microsatellite loci. Differentiation was measured for several adaptive morphological traits, and for traits associated with

mate recognition (colour and song). Morphological and neutral genetic patterns of diversification were compared to determine whether change could be accounted for by drift only, or if selective forces were likely to have been involved. The relative roles of ecology and isolation were also assessed by comparing the amount of change explained by habitat differences with that explained by isolation alone. In this system this is a very conservative test in favour of isolation, as the islands were covered in rainforest for most of their history. Only after human colonisation did the opening and clearing of the forests change this, with shade tree plantations being a characteristic feature of the environment for only the last 300-400 years. Finding more divergence in isolated forest populations could merely reflect the fact these populations (separated for thousands and million of years) had more time to diverge than those separated for the last few centuries.

Investigation of the link between population divergence and speciation was restricted to song, a trait with a high potential for constituting a reproductive barrier in birds as it is used for mate attraction. To determine whether divergence in the song of the allopatric *S. rufobrunneus* populations (Chapter 6) could lead to the evolution of reproductive isolation, playback experiments were carried as described in **Chapter 7**.

Finally, in **Chapter 8** a general discussion puts together all the lines of evidence collected during this work, in order to synthesize what was learned regarding the importance of isolation and ecology in the speciation process, and to propose an explanatory model for the uniquely high levels of bird endemism found in the Gulf of Guinea.

2

THE GULF OF GUINEA WHITE-EYES (ZOSTEROPIDAE): EVOLUTIONARY HISTORY AND PHENOTYPIC DIVERSIFICATION

The white-eyes (Zosteropidae), the most speciose bird family in the Gulf of Guinea island system, are used to investigate the factors behind the high levels of bird endemism in the region. Eight taxa are present, seven of which are endemic to single islands (including Mt. Cameroon). In four instances, two taxa co-occur on the same island, a rare situation in a family of mostly allopatric species. In each case one of the species diverged so much from the typical white-eye pattern (genus *Zosterops*) that they have been classified as a separate genus, *Speirops*. They are larger and have lost the green and yellow colours of *Zosterops*. Mitochondrial and nuclear sequences were used together with microsatellite genotypes to infer the evolutionary history of the group in the region. Morphological data were used to infer factors behind the patterns of phenotypic differentiation. A total of 1302 bp were obtained from the mtDNA (cyt *b*, ND3, ATP6). A total of 1952 bp were obtained from four nuclear introns (FIB7, G3PDH, MPP4, PEPCK9), but they had no informative sites within the Gulf of Guinea taxa. From 14 microsatellites described in other species, 12 amplified, and the six most variable were used to infer relationships between the Gulf of Guinea taxa. The Gulf of Guinea white-eyes did not constitute a monophyletic group but consist of two separate clades. One groups all oceanic island taxa, the other the Bioko and Mt. Cameroon ones. The genus *Speirops* with members in both clades is therefore invalid. The close affinity between Bioko and Mt. Cameroon is in agreement with their geographical proximity and the fact that Bioko was connected to the mainland many times in the past. The monophyly of the oceanic species supports an archipelago radiation model of diversification, where inter-island dispersal was much more important than colonisations from the mainland. The *Speirops* from the oceanic clade represent the most recent speciation events, indicating a very fast rate of phenotypic evolution. Character displacement was the most likely process behind the origin of the 'aberrant' traits of the oceanic *Speirops*, whereas the 'aberrant' characters of mainland *Speirops* represent typical adaptations to the montane environment where they occur. The sympatric speciation mode could not be safely rejected for the origin of *S. brunneus* (Bioko) and *S. lugubris* (São Tomé). The rapid radiation of the oceanic species led to a paraphyletic pattern for *Z. ficedulinus*, with subspecies on São Tomé and Príncipe. Clarification of the specific status of each population has important conservation implications, as Príncipe's population is very rare.

INTRODUCTION

The white-eyes (family Zosteropidae) constitute the most speciose bird group in the Gulf of Guinea and are therefore the obvious starting point to investigate the factors behind the high levels of bird diversification in the region. They constitute a very useful model to address the basic biogeographic questions of colonisation routes and modes of speciation, and to investigate the factors behind the evolution of phenotypic diversity in this system.

The Gulf of Guinea Zosteropidae

The Gulf of Guinea is a speciation centre for Zosteropidae. Of the 18 species occurring in Africa (Box. 2.1), seven are restricted to the Gulf region. Of these, six are endemic to the island system (five on the islands, one on Mt. Cameroon; Fig. 2.1). The forest white-eye *Z. stenocricotus* is restricted to the countries surrounding the Gulf, from SE Nigeria to Gabon (Fry 2000). The Gulf of Guinea island system is a place of particular interest for Zosteropidae as it has four cases where two species co-occur in the same place (Fig. 2.1) – something extremely rare for the family (three other examples occur in the Indian Ocean islands). In each of these four instances, one of the species is so differentiated from the typical white-eye pattern that all four have been placed in a separate genus *Speirops*, endemic to the Gulf (all other African species belong to *Zosterops*). *Speirops* is characterised by having lost the lipochromes (green and yellow colours) and being larger than *Zosterops*, with the São Tomé speirops *S. lugubris* being one of the largest zosteropids in the world together with other atypical species such as the Yap greater white-eye *Rukia oleagina* from Micronesia. Nevertheless, in many other characters they are as different between themselves as they are from *Zosterops* (Moreau 1957): colours vary from almost white to dark brown, and the eye-ring from prominent to absent. What gives unity to the genus is the fact that the constituent species are so different from all African species and that they occur together in a restricted geographical area. In this sense, the genus appears to be an arrangement of convenience to highlight the 'aberrant' characters of these four species. This seems to have been the position of (Moreau 1957, cf. note 28, Appendix 1; Hall & Moreau 1970). Amadon

(1953) suggested that the genus *Speirops* could be made synonymous with *Zosterops*, but still considered these 4 species to be closely related, also using 'aberrancy' as the unifying characteristic. More recently, morphological data were used to support the monophyly of the group (Feiler & Nadler 1992; Eck 1995). Together with the general acceptance of the genus *Speirops*, there is a consensus that the species of the genus must represent earlier invasions of the Gulf of Guinea due to their higher differentiation (Moreau 1957; Hall & Moreau 1970; Jones & Tye 2006). A point of debate has been the relationship between the São Tomé and the Mt. Cameroon species. The Mount Cameroon Speirops *S. melanocephalus* was initially described as a separate species (Gray 1862), but in later revisions it was considered conspecific with the São Tomé species, *S. lugubris* (Moreau 1957; Hall & Moreau 1970). Wolters (1983) separated the two species again, but this was considered unjustified by Dowsett & Dowsett-Lemaire (1993) and Eck (1995). Recently however, the separation of the two species has been widely accepted (Fry 2000; Borrow & Demey 2001; Christy 2001; Sinclair & Ryan 2003; Jones & Tye 2006). Considering the populations from Mt. Cameroon and São Tomé as the most closely related (independently of whether or not they are valid species) creates a biogeographical puzzle, as it seems to require a dispersal event over two islands. One hypothesis considers that the genus *Speirops* originated on São Tomé, and from there back-colonised the mainland by island-hopping; the populations that were established on Príncipe and Bioko would have diverged more than the Mt. Cameroon one because of their smaller effective population size (Jensen & Stuart 1986). The possibility of back-colonisation depends foremost on the species from São Tomé being the oldest of the pair, something for which Moreau (1957) did not find conclusive evidence. If this condition is met, back-colonisation could occur since the prevailing winds blow in the required direction. The second part of the hypothesis aiming at explaining the appearance of two very distinct species between the two sister populations is more unsatisfactory: it is based on the little supported founder-effect speciation model (Coyne & Orr 2004), and does not consider the fact that the sequential colonisations until reaching mainland would, if anything, decrease the genetic diversity of the Mt. Cameroon population (Clegg *et al.* 2002a). Additionally, the Mt. Cameroon species can be considered as much an insular population as the

others as its full range is restricted to a narrow altitudinal belt (1820-3000m; Fry 2000) of a single mountain.

In this study, I use mitochondrial and nuclear genetic data to: 1) infer the origins, time and sequence of colonisation, and 2) the diversification mode of the Gulf of Guinea white-eyes. Results will clarify the affinities of the species in the group, and in particular the origin (and validity) of the genus *Speirops*. Morphological data are used to: 3) investigate trends of phenotypic evolution, and how they relate to the co-existence of white-eyes in sympatry. These three lines of enquiry are of direct interest to the understanding of evolutionary history of the region, as developed next.

1. Origins and colonisation routes in the Gulf

The Gulf of Guinea white-eyes allow us to tackle one of the major biogeographic issues regarding bird diversification in the Gulf, namely determining the relative roles of mainland colonisations and inter-island dispersal events in promoting speciation on the three oceanic islands. This constitutes an unsolved issue as the distances between islands are very close to the distances between the islands and the mainland (Fig. 1.1): is each island a separate entity or do they make a linked unit? The fact that only 6 out of the 29 endemic birds are shared between islands has been used to support the primacy of independent colonisations from the mainland (e.g., Snow 1950). This argument is likely to hold for those cases where species from different families are present in different islands, but not for more closely related species, as is the case with the white-eyes, and several other groups such as the sunbirds (Nectarinidae), seedeaters (Fringillidae), thrushes (Turdidae), and pigeons (Columbidae). Here, the archipelago radiation model (Chapter 1; Grant 2001) is the alternative hypothesis. In this model, a mainland species colonises one island and then disperses to others; the different populations diverge in allopatry; further dispersal events lead to secondary contacts; when two populations that diverged in different islands meet in sympatry they might fully hybridize and merge or, instead, diverge even further as a result of competition, leading eventually to full reproductive isolation. Therefore movements between islands are the most important factor in generating diversity.

As already pointed out by Amadon (1953), the predominant winds in the Gulf of Guinea blow from a SSW direction which is more favourable to dispersal events up the island chain than to mainland-islands movements. On the other hand, during the glacial cycles the influence of the northerly dry harmattan winds extended further over the Gulf of Guinea as the meteorological equator moved southward (Lézine *et al.* 1994). This suggests that any colonization from the mainland was more likely to happen from the north, a prediction that has been supported with results for other aerial colonisers such as angiosperms (Exell 1973) and blackflies (Mustapha *et al.* 2006), which have their source populations in Nigeria and Cameroon respectively, rather than in Gabon, as well as for the only bird for which genetic data are available, the grey parrot, which colonised Príncipe from the Upper Guinea region (Melo & O'Ryan 2007). In this study, the phylogenetic hypothesis inferred from molecular data will be used to test amongst different scenarios of colonisation.

2. Speciation mode of sympatric species

One corollary of the archipelago radiation model is that the co-existence of closely related species on the same island is a result of secondary colonisations and not of sympatric speciation. The high mobility of birds together with the observation that island species tend to be more generalist than their mainland counterparts (Scott *et al.* 2003, and references therein), makes conditions for sympatric speciation unlikely to be met in birds in general, and in island birds in particular (Newton 2003). Indeed, a meta-analysis could not find evidence for sympatric speciation of birds on small islands (Coyne & Price 2000). Molecular studies have also supported the double colonization model (Tarr & Fleischer 1995; Roy *et al.* 1998; Marshall & Baker 1999; Warren *et al.* 2003, 2006). Such is the strength of the double colonisation paradigm, that the rare exceptions where sympatric populations are closer genetically to each other than either is to allopatric populations are explained as resulting from introgressive hybridization (Grant & Grant 1979; Yang & Patton 1981; Grant 1986). By using both information from the mitochondrial and the nuclear genome, this study will test the null hypothesis that the white-eyes of the Gulf of Guinea constitute "beautiful and varied cases of 'double invasion' of small islands" (Moreau 1957, p. 408).

3. Causes of phenotypic evolution

Independently of the mode of speciation of closely related species occurring in sympatry, their co-existence is only possible if neither out-competes the other. This can be achieved by spatial or temporal segregation, or by evolving adaptations to exploit different niches in the same place and time (Lack 1971; Diamond 1973). That ecological differentiation is necessary for co-existence in the same area has been supported by evidence showing that congeneric species occurring in the same habitat are more distinct between themselves than they are in relation to congeners in different habitats (Mayr 1942; Lack 1947; Grant 1965b, 1968; Williamson 1981). In the allopatric speciation model, a minimum ecological difference must evolve before species meet in sympatry (Mayr 1940; Lack 1944). Upon secondary contact, competition can promote further phenotypic divergence between species so that each specializes in different resources, in what is designated 'character displacement' (Brown & Wilson 1956). Character displacement is one of the central tenets of the ecological theory of adaptive radiation but is difficult to demonstrate directly (Schluter & McPhail 1993; Schluter 2000). In island birds, besides the compelling indirect evidence provided by the comparison of sympatric and allopatric species pairs mentioned earlier, there are a few studies strongly supporting competition in sympatry as a driver of phenotypic evolution (Moulton 1985; Schluter *et al.* 1985; Diamond *et al.* 1989).

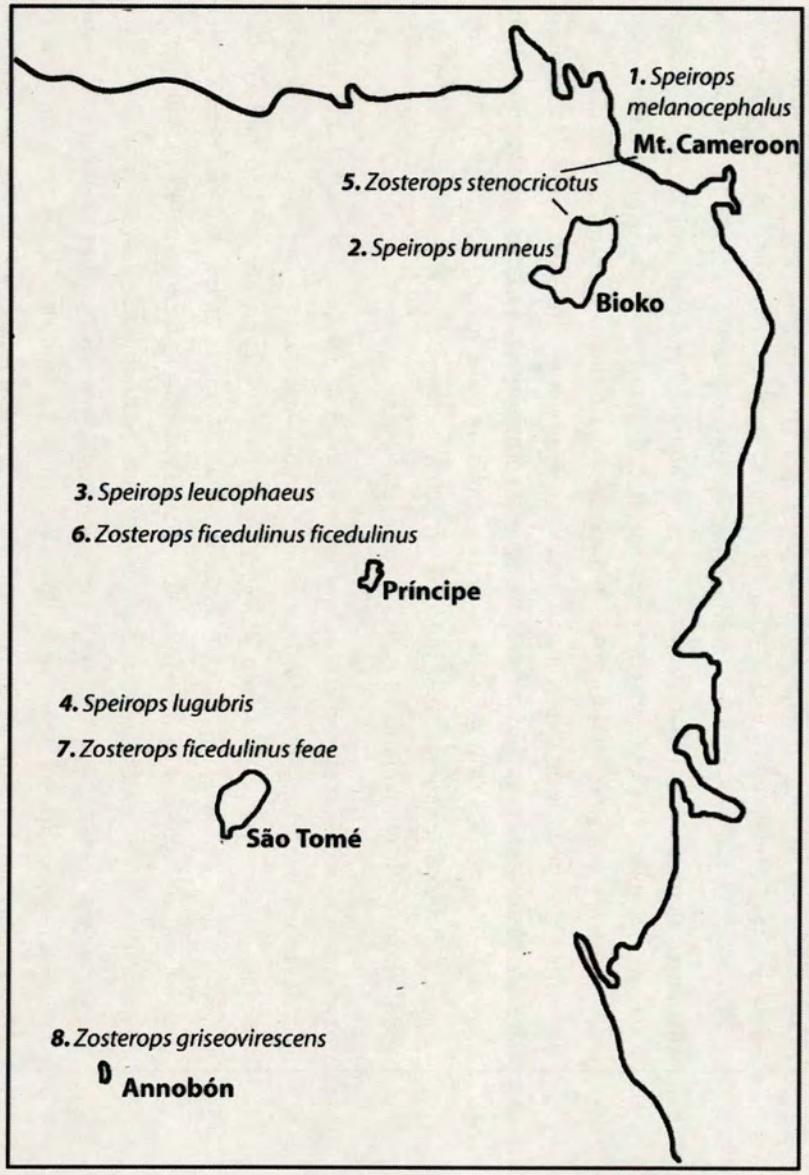
Here I use molecular phylogenies to describe trends of phenotypic change and infer the most likely causes of phenotypic divergence between sympatric species pairs: did current differences evolve fully in allopatry or did character displacement in sympatry play a role? The former case would be supported if the two sympatric species were derived from different ancestors to whom they are phenotypically closer. This pattern favours a model of size assortment: successful colonisers are those that are different from the species already present (Case & Sidell 1983). The character displacement hypothesis is favoured if the two sympatric species are derived from the same ancestor or from ancestors with similar sizes.

Box 2.1 - The white-eyes, family Zosteropidae

White-eyes (or silver-eyes) are an Old World tropical family comprising c. 100 species in 13 genera of small, gregarious, arboreal birds (Sibley & Monroe Jr 1990; Fig 2.2). The family is generally very homogenous, with phenotypic uniformity being particularly striking within the nominate genus *Zosterops*, which comprises 75% of the species. *Zosterops* are insect and fruit eaters and have a brush-tipped tongue used to suck nectar and other liquid foods such as insect secretions; typically, they have a yellow or green colour and a white-feather ring around the eyes (Fry 2000). The other 12 genera depart from the nominate genus mostly in colour and body size, but few have been well studied (Fry 2000). Only rarely do two or more species occur in sympatry. A very interesting feature of this family is the propensity of its species to colonise islands and speciate there: it is the passerine group that colonised more islands worldwide (Moreau 1957), with almost 90% of the species confined to islands, whereas wide continental areas are occupied by a few similar species (Hall & Moreau 1970). Even on the mainland, most diversity is restricted to habitat islands, such as montane forests. Hence, this group played an important role in the development of Mayr's peripatric model of speciation (Mayr 1942), that led later to the founder effect model of speciation (Mayr 1954).

The combination of phenotypic uniformity across most species together with the presence of a few species departing from the common pattern has made the establishment of taxonomic relationships based on phenotypic measurements extremely difficult, with "affinities of numerous taxa remain[ing] obscure" (Fry 2000, p. 305). The small range of variation makes instances of convergence common, leading often to cases where species living thousands of kilometres apart are much more similar than neighbouring populations of the same species. A typical example is the very close resemblance of *Z. griseovirescens* from Annobón Island to *Z. natalis* from Christmas Island in the Indian Ocean. Morphometric and colour characters have proved to be of little utility in understanding the relationships within the Zosteropidae (Moreau 1957). This problem can only be addressed by inferring accurate phylogenies from neutral molecular data. The first steps in this direction have just begun, with phylogenies being constructed for Micronesian (Slikas *et al.* 2000), Indian Ocean (Warren *et al.* 2006) and South-west Pacific (Phillimore, *in. prep.*) white-eyes.

The exact number of Zosteropidae occurring in Africa is an unsettled issue for the reasons just described. Modifying the classification of Fry (2000) to accept the currently dominant view that *Z. [pallidus] capensis* and *Z. [senegalensis] stenocricotus* are valid species (Sinclair & Ryan 2003; Dillon & Fjeldså 2005), and adding the Madagascar and Indian ocean species (Warren *et al.* 2006), Africa and its associated islands currently hold 18 species of Zosteropidae, 10 of which are island endemics. One additional species, *Z. semiflavus* from the Seychelles, went extinct in recent times (Diamond 1984). Seven species occur in the Gulf of Guinea region, with six being endemic to the island system. Here there are four cases of two species co-existing on the same island. In each case one species differentiated enough to be in a separate genus, *Speirops*, endemic to the Gulf.



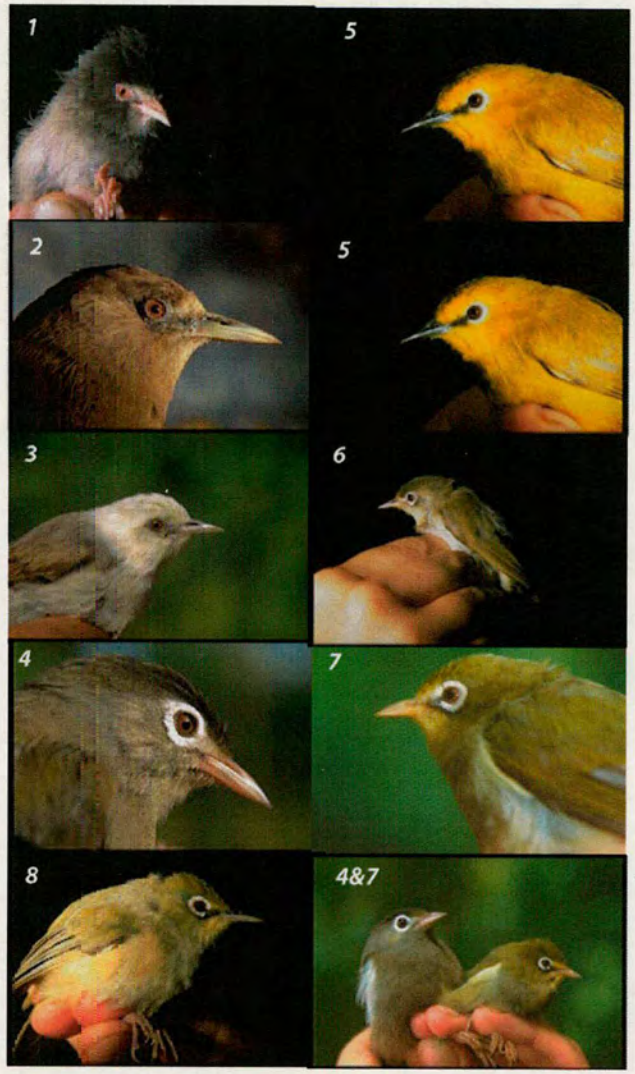


Fig 2.1. Gulf of Guinea Zosteropidae. Four 'aberrant' white-eyes making the endemic genus *Speirops* (1–4) co-occur with 'typical' *Zosterops* white-eyes (5–7). *Speirops* are bigger and lack yellow and green pigments.Annobón has a single *Zosterops* species (8).

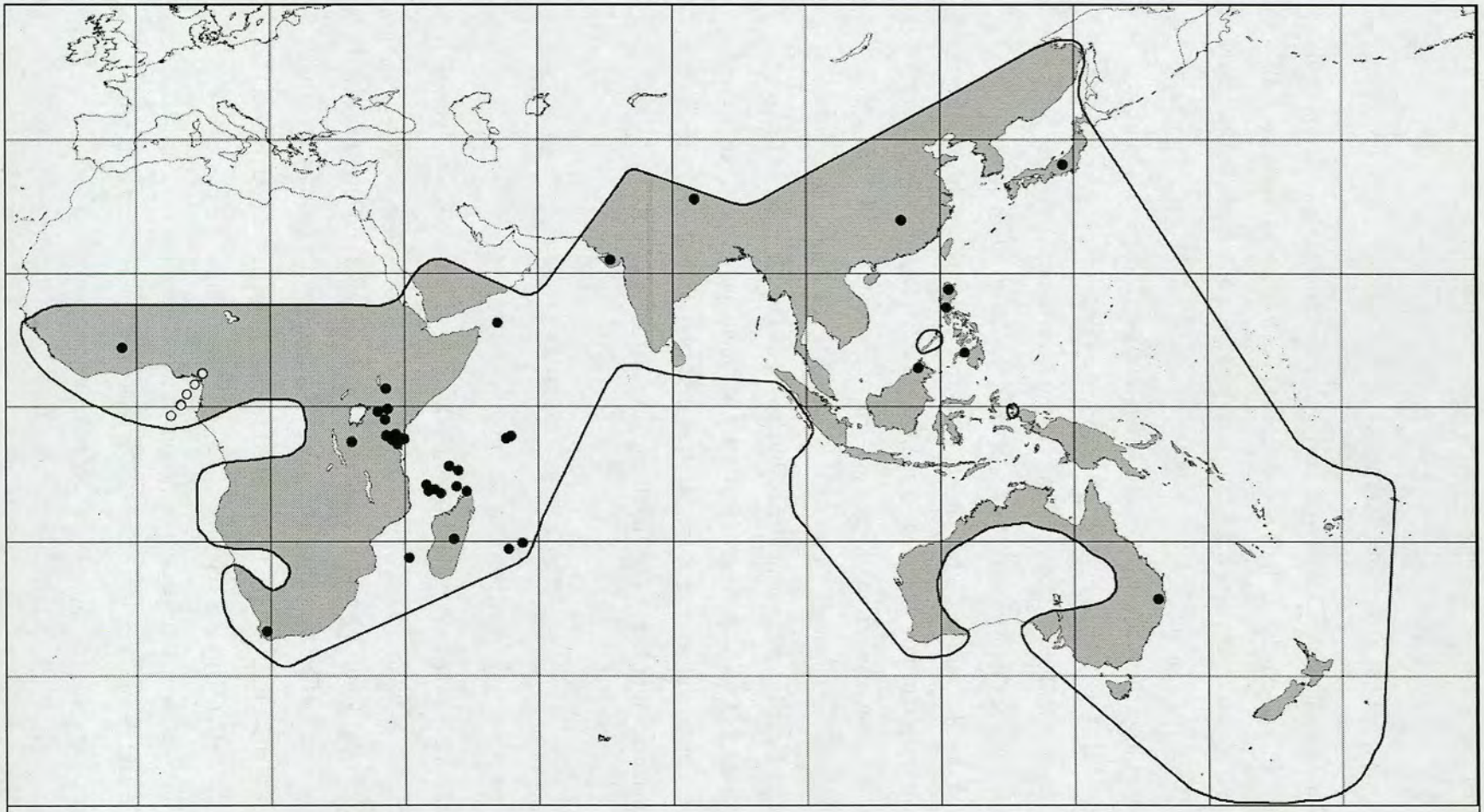


Fig 2.2. The global distribution of white-eyes (Zosteropidae) adapted from Warren *et al.* (2006). Sightings of white-eyes in southern Iran and eastern Oman have not been included as their actual distribution and status is unknown. Dots indicate sampling localities, with white-dots referring to the Gulf of Guinea localities enlarged in Fig. 2.1.

METHODS

Sampling

In total, 47 individuals from 32 taxa were analysed in this study (Fig. 2.1; Appendix 2.1). This sampling includes all the 9 populations of the 8 Gulf of Guinea taxa, all the south-west Indian Ocean taxa (including all island populations of the Madagascar white-eye *Z. maderaspatanus* with the exception of the Cosmoledo population), and representatives from the major African and Asian lineages (Mees 1957; Moreau 1957; Mees 1961, 1969). The phylogeny was rooted on the chestnut-faced babbler *Stachyris whiteheadi* (Timaliidae) whose suitability as an outgroup for the Zosteropidae has been shown in previous studies (Slikas *et al.* 2000; Warren *et al.* 2006). Blood samples for the Gulf of Guinea taxa were collected for this study during field expeditions that took place from 2002 to 2005. Blood samples were collected non-destructively from the brachial vein of mist-netted individuals, and were stored in absolute ethanol. Sampled birds were measured and ringed. Measurements recorded were: upper mandible length and width, wing and tarsus length, and weight. Two or three individuals per taxon and population were sequenced. Samples and sequences for the other taxa were previously obtained for a study on the radiation of Indian Ocean white-eyes (Warren *et al.* 2006).

Laboratory protocols

DNA extraction and sexing

Total genomic DNA was extracted from the blood using a DNeasy Tissue extraction kit (Qiagen), following the protocol for DNA extraction from animal tissues. Previous to the extraction, blood was removed from the ethanol and left to dry on the bench. It was then washed by adding 800 µl of blood lysis buffer (50mM Tris-HCl, 100mM NaCl, 5mM EDTA, 1% SDS), vortexing, centrifuging (> 8,000 rpm), and discarding the buffer. This was repeated once, before addition of the DNeasy lysis buffer. As blood preserved in ethanol dries considerably, the digestion step was carried overnight. Some samples required additional proteinase K and lysis buffer and an extension of the digestion time in order to be adequately digested.

As the Zosteropidae have no sexual dimorphism, for the morphological analyses birds were sexed with a molecular protocol (Griffiths *et al.* 1998). This PCR based approach amplifies homologous segments of the two sex-linked chromo-helicase-DNA-binding genes (CHD). Due to differences in intron sizes, copies in each sex chromosome, Z and W, also differ in size. Therefore, when viewed in a gel, the amplification product for females gives two bands (ZW), whereas in males only one band can be viewed (ZZ).

Sequence data - mitochondrial DNA

A total of 1302 bp of mitochondrial sequence data was obtained for all the Gulf of Guinea individuals from the complete ATP synthase 6 (ATPase6; 661 bp) gene and partial sequences of the NADH dehydrogenase subunit 3 (ND3; 351 bp) and the cytochrome *b* (cyt *b*; 290bp) genes. Overlapping sequence fragments were amplified via polymerase chain reaction (PCR) using the primers and conditions detailed in Appendix 2.2. PCR products were electrophoresed on 1% agarose gels stained with ethium bromide and visualised under UV light. Successful PCR amplifications were purified with a Qiagen PCR Purification Kit (cyt *b*) or a Qiagen Gel Purification Kit following the manufacturer's protocol. The amplification primers were used to sequence both DNA strands with the dye-labelled termination method using ABI BigDye 3.1 reagents (following the manufacturer's protocol for quarter reactions) on an ABI 3730 sequencer. Sequences were aligned with SEQUENCE NAVIGATOR (Applied Biosystems) and by eye. As expected for protein coding sequences, no gaps, insertions or deletions were detected, and all sequences translated appropriately into amino acid form. The possibility of having amplified nuclear copies of mitochondrial genes (Bensasson *et al.* 2001) could therefore be safely discarded. Sequences were deposited in GENBANK.

Sequence data - nuclear DNA

As all mitochondrial genes are inherited as a single linkage group (Ballard & Whitlock 2004), independent molecular information must come from the nuclear genome. Introns are a nuclear marker of choice for phylogenetic studies because: i) they are non-coding and therefore evolve at relatively fast rates; ii) they are flanked

by protein-coding exons that, being much more conserved, are a very adequate target to design primers that work over large evolutionary expanses; iii) they are abundant and of convenient lengths for sequencing (Prychitko & Moore 1997). As no introns have ever been used in Zosteropidae, the amplification success and levels of variability of five introns were tested in the sample of the Gulf of Guinea taxa (n = 18: 9 populations, 2 individuals per population). The introns tested were: introns 5 and 7 of the β-fibrinogen gene (FIB5, FIB7), intron 11 of the glyceraldehyde-3-phosphate dehydrogenase (G3PDH), intron 4 of the Myelin proteolipid protein (MPP), and intron 9 of the phosphoenolpyruvate carboxykinase (PEPCK9). The choice of introns was guided by the usefulness demonstrated in other studies (e.g. (Prychitko & Moore 1997, 2000; McCracken & Sorenson 2005). PCR primers and conditions are detailed in Appendix 2.2. PCR amplifications were purified with a Qiagen PCR Purification Kit (FIB7) or a Qiagen Gel Purification Kit (G3PDH, MPP). PEPCK9 was amplified by nested PCR: the product of a first amplification was cut from the gel, put inside a filter tip which was cut to fit a 1.5 ml microtube, and centrifuged for 1 min. One μl of the recovered product was used as the template for a second amplification of a smaller inner fragment. This final product was purified with a Qiagen PCR Purification Kit. FIB5 could not be amplified. Sequencing was done as for mtDNA. Because none of the amplified introns provided informative sites for phylogenetic analysis no further efforts were made in the optimization of PCR conditions for FIB5. For FIB7 only the forward primer provided an adequate sequence (a minimum of 640 bp for all individuals). Internal sequencing primers (FIB-BI7U2 and FIB-BI7L2; (Prychitko & Moore 1997) were also inadequate for sequencing in the Zosteropidae. No further efforts were made to optimize the sequencing reactions as the 640 bp obtained (out of *ca.* 1000 bp) only had six variable sites, none informative. A sequence of 330 bp was obtained for the G3PDH, from both strands, which had only three variable sites. A sequence of 310 bp was obtained for the MPP, from both strands, with no variable sites. A sequence of 672 bp was obtained, from both strands, for PEPCK9, but it contained only three variable sites, none informative. Sequences were deposited in GENBANK.

Microsatellites

Since introns did not provide enough variation to elucidate relationships within the Gulf of Guinea white-eyes, a search for potentially useful microsatellites was carried out. Because of their very high mutation rate, microsatellites are the preferred marker for the analysis of genetic structure within species, but have also been used to infer relationships between closely related species (Goldstein & Schlötterer 1999). Microsatellite loci are difficult to find and birds are the vertebrate group with the lowest microsatellite density (Ellegren 2005). Nevertheless, amplification of microsatellites across species boundaries (cross-amplification) is possible, with the probability of success inversely related to the genetic distance between species (Primmer *et al.* 2005). Therefore, rather than building a microsatellite library for the study species, 14 microsatellite primers developed in other species were tested for cross-amplification and polymorphism on 12 individuals from five Gulf of Guinea taxa. Of the 15 microsatellite loci described for the Capricorn silvereye *Zosterops lateralis* (Degnan *et al.* 1999; Frentiu *et al.* 2003), 10 have been shown to be polymorphic across several *Zosterops* species (D. Dawson's BIRDMARKER webpage). Of these, nine were tested for cross-amplification as two (ZL44 and ZL45) have recently been shown to be linked (Hansson *et al.* 2005). Additionally, five loci described for the Seychelles warbler *Acrocephalus sechellensis* that have been shown to amplify successfully in other species were also tested (BIRDMARKER; Richardson *et al.* 2000). PCR primers and conditions are detailed in Appendix 2.3. The fluorescent-labelled PCR products were separated and alleles were detected in an ABI PRISM 3730 capillary sequencer (Applied Biosystems). Only one locus did not amplify (ZL35), whereas another gave a product with too much noise from stutter bands (Ase48). Of the remaining 12, three were monomorphic and three had only two alleles (Appendix 2.3). The remaining six were used in all populations of the Gulf of Guinea taxa and in the Cape white-eye *Zosterops capensis*, which was used as an outgroup. Ten individuals were genotyped for six out of the eight Gulf of Guinea taxa, and for the outgroup. For the Mount Cameroon Speirops *Speirops melanocephalus* and for the Príncipe subspecies of the São Tomé and Príncipe white-eye *Z. f. ficedulinus* only three and four samples were available, respectively. For the

Western yellow white-eye *Zosteropus stenocricotus*, five samples were from Mount Cameroon and five were from Bioko Island.

Phylogenetic analyses - mitochondrial sequence data

Sequence characteristics

The evolutionary dynamics of the three mitochondrial genes and of their codon positions were investigated in PAUP* version 4.0b10 (Swofford 2003), using all ingroup sequences. The following parameters were estimated by maximum likelihood on a neighbour-joining (NJ) tree inferred from LogDet transformed distances: percent base composition, relative substitution rates, gamma shape parameter (α), proportion of invariant sites, and transition to transversion ratio. Whenever the proportion of invariant sites was close to zero, it was set to zero and parameters were re-estimated. Uncorrected pairwise genetic distances within the ingroup taxa were calculated to assess the potential for loss of phylogenetic signal due to superimposed substitutions leading to homoplasy (when a character is shared not by descent but by convergent or parallel evolution: Swofford *et al.* 1996). To explore this possibility further, pairwise plots of corrected and uncorrected distances for each gene and each codon position were constructed (saturation plots). Homogeneity of base composition across ingroup taxa for each gene and each codon position was tested in PAUP* using a chi-square analysis using only the informative data, as taxa with biased base composition may affect phylogenetic inference (Sanderson & Shaffer 2003).

Mitochondrial genes are inherited as a single linkage group and share therefore the same history (Ballard & Whitlock 2004). Nevertheless, the phylogenetic signal of the different genes could still be incongruent if any (or all) of the genes are behaving in a way that cannot be properly handled by the phylogenetic inference method (Bull *et al.* 1993). Even if they share the same history, combining the genes in this case could either dilute or enhance the problem, depending on the strength of the signal of the gene(s) that allow proper phylogenetic inference. To assess whether the different genes had a compatible phylogenetic signal a parsimony based incongruence-length differences test (ILD; Farris *et al.* 1995) was performed in PAUP* with the invariant

sites removed (Cunningham 1997), for 100 replicates and 5 iterations per replicate. Because the utility of the ILD to detect phylogenetic congruence has been seriously criticised (e.g., Barker & Lutzoni 2002), a Bayesian test of incongruence (BTI; (Irestedt *et al.* 2004) was also performed. The BTI uses Bayes factors to compare trees inferred by constraining all genes to share the same topology (but allowing each one to follow its own substitution model) and trees where each gene is allowed to have an independent topology. Bayesian inference (BI) as implemented in MRBAYES version 3.2 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003) was used to obtain the trees under each model. Details of the search strategy and the use of Bayes factors are presented below. Both tests supported congruence of the phylogenetic signal, and data from the three genes were combined for the analyses.

Phylogenetic inference

Phylogenetic analyses were performed using maximum parsimony (MP) and maximum likelihood (ML) in PAUP*, and Bayesian inference (BI) in MRBAYES. MP cannot deal with homoplasy (Swofford *et al.* 1996), whereas both ML and BI deal explicitly with homoplasy by incorporating statistical models of sequence evolution during phylogenetic inference (Swofford *et al.* 2001; Holder & Lewis 2003; Felsenstein 2004). For this reason, if there are indications of homoplasy, and the different methods give different results I decided *a priori* to consider the model-based estimates as the most reliable. Nevertheless, because the debate on the most appropriate phylogenetic inference method is still open, a MP inference can still be useful as: i) congruence of relationships inferred by different methods can be interpreted as providing strong support for those relationships; ii) MP can provide accurate estimates when sampling is dense so that branches are short and the potential for superimposed substitutions is reduced (Kim 1996). As this study looks at the relationships within a genus of closely related bird species, and taxon sampling is dense, one can expect that MP will perform reasonably well.

An unweighted parsimony analysis was conducted using a heuristic search implementing a stepwise addition with 100 random sequence replicates and TBR branch swapping. Clade support was estimated using the non-parametric bootstrap

(Felsenstein 1985) with 1000 pseudo-replicates, each with 100 random addition replicates.

A ML search was conducted using a successive-approximations approach (Swofford *et al.* 1996). In 'classical' ML phylogenetic inference the parameters of the substitution model must be optimized in all the trees visited during a search, making this an extremely time-consuming method. The realization that model parameters can be accurately estimated in any topology that includes the bipartitions strongly supported by the data (Sullivan *et al.* 1996) led to the development of the successive-approximations approach. A 'reasonable' first tree is estimated with a fast method such as parsimony or neighbour-joining. Parameters of the model of sequence evolution are estimated in this tree. These parameters are held constant during a first ML search. They are then re-estimated on the new tree found. These new parameters are held constant in a new ML search, and this iteration stops when no improvement in the tree (and change in the parameters) occurs. Sullivan *et al.* (2005) showed that if a rigorous heuristic search of the tree space is implemented (starting from several replicates of stepwise-addition trees obtained by random sequence addition) this method is as accurate as a full-optimization one. In this study, a LogDet NJ tree was used as the first topology for the estimation of model parameters by ML. ML searches used a heuristic search implementing 10 random addition sequence replicates. Each search was stopped when the likelihood score of a tree remained the same for more than 5,000 rearrangements, strongly suggesting that it had reached a plateau. Using non-parametric bootstrapping for ML proved extremely time-consuming. It was therefore decided to restrict the assessment of clade support to the support estimated by BI (for the clades recovered by both methods).

Bayesian inference of phylogeny is also based on likelihood, but differs from ML methods in that it incorporates priors, i.e., probability distributions summarising the knowledge about a model and its parameters before the analysis (Huelsenbeck *et al.* 2001; Huelsenbeck *et al.* 2002; Holder & Lewis 2003). The incorporation of *a priori* knowledge is sometimes criticised but it is what makes BI so powerful. By including priors, all parameters except for the one of interest (the tree for example) can be

integrated out and a posterior probability obtained. The posterior probability of a parameter gives the probability that the parameter is true given that the model, priors and data are correct. It can be interpreted as a true *p* value, so that for $\alpha = 0.05$, a posterior probability is only significant if > 0.95. Because it is impossible to integrate over all possible combinations of models and parameters, Markov chain Monte Carlo (MCMC) techniques are used to sample this space and obtain an estimate of the posterior probability. BI of phylogeny therefore estimates the tree and the support for its nodes at the same time. The choice of priors is generally not a critical issue in BI: non-informative priors can be used; the posterior probability from one analysis can be used as a prior for a new analysis; the effect of the priors on the posterior probabilities decreases the more data there are (Huelsenbeck *et al.* 2002). On the other hand, estimation of posterior probabilities is sensitive to model choice, with oversimplified models leading to strong support for the wrong phylogenies (Huelsenbeck & Rannala 2004; Lemmon & Moriarty 2004). This issue can be properly addressed in BI, since one of the strengths of the Bayesian MCMC approach is its ability to handle complex models: the data can be divided into separate partitions and each can follow independent evolutionary dynamics, sharing or not some of the model parameters (Ronquist & Huelsenbeck 2003; Nylander *et al.* 2004). For example, for protein coding genes it is more realistic to infer different substitution models for different codon positions, rather than estimating a single model for the entire sequence (Shapiro *et al.* 2006). The choice of the best partition scheme and models of evolution for each can be guided by the estimation of Bayes factors (Huelsenbeck *et al.* 2004; Nylander *et al.* 2004). Bayes factors can be used to compare any set of models (nested or not) by weighting the strength of evidence in favour of one model M_1 in relation to another, M_2, given the data X. The Bayes factor in favour of M_1 over M_2, B_{12}, is the ratio of the model likelihoods, $f(X \mid M_i)$, which are approximated by the harmonic mean of the likelihood values of the phylogenies sampled after convergence of a MCMC run. Interpretation of Bayes factor values follows the recommendations of Kass & Raftery (1995): $0<2\ln(B_{12})<2$, equivalent models; $2<2\ln(B_{12})<6$, positive evidence against M_2; $6<2\ln(B_{12})<10$, strong evidence against M_2; $2\ln(B_{12}) > 10$, very strong evidence against M_2.

The strategy used for model choice was: i) define logical data partitions; ii) choose the best-fit model of sequence evolution for each partition using the Akaike Information Criterion (AIC) as implemented in MRMODELTEST version 2.2 (Nylander 2004) letting the specific model parameters being those estimated during the MRBAYES analysis; iii) use Bayes factors to identify the best supported model. The partition schemes considered were: i) no partitions (3 genes combined, one single model); ii) 3 partitions = 3 genes; iii) 3 partitions = 3 codon positions (3 genes combined); iv) 2 partitions = codon position 1 and 2 + codon position 3 (3 genes combined). 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 the Bayesian Incongruence Test (see above), partitioning by gene was used (scheme 2) in the two models to be compared: one where all genes shared the same topology and one where each gene was allowed to have its own topology. For each model tested, two independent MRBAYES runs were carried out simultaneously, starting from random trees and using the default priors. Each run consisted of four simultaneous Metropolis-coupled MCMC (MC^3) chains with a length of 2 million generations. Trees were sampled every 100 generations with the first 10% discarded as burn-in. Proper mixing of the chains was assessed by the acceptance rates of the Metropolis proposals. Log-likelihood plots were checked to confirm that stationarity was reached after the 10% burn-in. Additional confirmation that the analyses had properly explored the tree space was done by estimating the effective sample size of the log-likelihoods (ESS: estimate of the number of samples that can be considered independent) using TRACER version 1.2 (Rambaut & Drummond 2003). Convergence of the two independent runs was confirmed by checking the frequency plots of the log-likelihoods in TRACER and the value of the average standard deviation of split frequencies in MRBAYES. Bayesian posterior probabilities were estimated for each clade from the 50% majority-rule consensus tree of the 36,000 sampled trees. To confirm that the convergence to the global optimum was reached, a new search was conducted for the best model of sequence evolution with two independent runs with a length of 5 million generations, a sample frequency of 500, and a 25% burn-in. Bayesian posterior probabilities were estimated from the 50% majority-rule consensus tree of the 15,000 sampled trees.

Bayes factors and Shimodaira & Hasegawa (1999) tests were used to test alternative phylogenetic hypotheses. These included: i) relationships from traditional taxonomy not supported by the molecular phylogeny; ii) sympatric versus allopatric modes of speciation in species co-occurring on the same island. In each case, one constrains the phylogenetic inference to recover the hypothesis to be tested, and then tests if the difference in log-likelihood scores is significant (the alternative hypothesis will always be worse by definition: it was not the best tree found in the first place). The Bayes factor based test was conducted as above. The Shimodaira and Hasegawa (SH) test (with the RELL approximation) was conducted using ML trees. Each alternative topology was inferred using the successive-approximation method described above, starting with the parameters of the ML tree inferred without constraints.

Origin and colonisation routes of the Gulf of Guinea oceanic white-eyes

To infer the area of origin of the Gulf of Guinea oceanic island species and the direction and sequence of colonisation of the archipelago, two parsimony-based and two Bayesian-based methods of ancestral character state reconstruction were used. The use of this array of methods had the heuristic purpose of assessing the performance of the different approaches.

Parsimony methods find the ancestral states that minimize the number of steps of character changes onto the phylogenetic tree of interest. It is the classical approach of historical biogeography and different sets of rules and algorithms have been proposed to count the changes (Morrone & Crisci 1995). For example, different events – such as dispersal, vicariance or extinction – can be given different costs according to their likelihood (Ronquist 1998; Sanmartín & Ronquist 2004). The two parsimony methods used were: i) parsimony reconstruction of the distribution areas mapped on the ML tree, as implemented in MESQUITE version 1.06 (Maddison & Maddison 2005); ii) the ancestral areas method of Bremer (1992). The former is the basic parsimony method used here with an ordered parsimony model described on a stepmatrix. This model attributed more weight (steps) to a change the farther apart any two areas were. Colonization of São Tomé and Príncipe from the mainland were

given the same weight, with the colonization of Annobón given an extra step (being the smallest and most isolated island). Bremer's (1992) ancestral areas method combines two parsimony analyses, one from the tips to the root and another from the root to the tips. The former assumes that the ancestral area was the same as the present area so that any area missing from a branch is considered a loss; the latter makes the alternative assumption, counting all area presences as gains. The gain/loss ratios for all areas are re-scaled, so that the maximum value is one, in order to provide the relative probabilities of each area belonging to the ancestral area for the group. This method does not penalize dispersal in relation to vicariance, making it appropriate for an oceanic island setup.

The major limitation of parsimony-based methods is that they rely on one single topology, which is rarely free of error. A second limitation is that often several alternative hypotheses for the reconstruction of ancestral areas can be found with very close parsimony scores (Ronquist 2004). Bayesian inference of ancestral character states can take into account both types of uncertainty (Ronquist 2004). As for any other parameter of interest, the posterior probability of an ancestral character state in a particular node can be obtained by integrating over all other parameters during sampling of the tree space. Because not all sampled trees will necessarily contain the node of interest, Huelsenbeck & Bollback (2001) suggested restricting the analyses to the subset of trees where the node is present. For the white-eye data, this was done by setting a constraint in MRBAYES with the node of interest and asking for the inference of the character state for that particular node. Only one node can be constrained (and its ancestral character state estimated) during a MC^3 search in order to not exclude trees that only have a subset of the nodes of interest. MC^3 runs were conducted under the same conditions as used for phylogenetic inference. One run estimated the ancestral state of the white-eyes from the Gulf of Guinea oceanic islands; a second run estimated the ancestral state of the node immediately above.

Pagel *et al.* (2004) showed that using only the sample of trees with the node of interest overestimates the posterior probability of finding a given character state in

that node, because a component of phylogenetic uncertainty (trees that do not contain the node) is excluded from the estimation. Pagel *et al.* (2004) devised a Bayesian and a ML approach that incorporates both the uncertainty about the phylogeny and the ancestral character state, and allows specifying different scenarios of character change that can be tested against each other (by constraining the different types of change that can occur). This approach, implemented in the program BAYESMULTISTATE, was used to infer ancestral areas and test different colonisation scenarios of the oceanic islands of the Gulf of Guinea. Bioko was treated as part of the mainland to which it was linked several times in the past. Nine different colonisation models were tested alongside the null model where any area could be the source of any other (Fig. 2.3). Model 1 describes the 'true' stepping-stone model for the region. Models 2-4 are variations of the stepping-stone model, which could arise due to the similar distances between the different areas (Fig. 2.1). Models 5-6 are two variations of the hypothesis of the origin of the genus *Speirops* on São Tomé followed by back colonisation to the mainland via Príncipe (see Introduction); on model 6, multiple colonisations from mainland allow for independent origins of the insular *Zosterops*. Model 7 considers that all insular species derive from mainland colonisations, whereas model 8 also allows for dispersal between neighbouring islands. Model 9 allows all changes between neighbour areas to occur. It only differs from the null, unordered, model (10) in that dispersal between Annobón and Príncipe is not allowed. A sample of 500 trees was obtained with BAYESPHYLOGENIES (Pagel & Meade 2004). Instead of defining partitions *a priori*, each site was fitted to one of three GTR rate matrices, with gamma-distributed rate heterogeneity. The search strategy consisted of one single chain that explored the tree space more thoroughly than the MC^3 chains by using a logarithmic cooling algorithm. The chain was heated with a temperature parameter of -100 and allowed to cool over the first 100,000 generations. After cooling, the chain was allowed to reach convergence and then 500 trees were sampled at intervals of 10,000 trees (5,000,000 iterations), to guarantee independence of the trees used for character reconstruction. As only the Gulf of Guinea oceanic white-eyes are of interest in this analysis and in order to speed up computing time, this search was performed for a sub-sample of the original data set (19 rather than 32 taxa, 31 rather than 47 individuals).

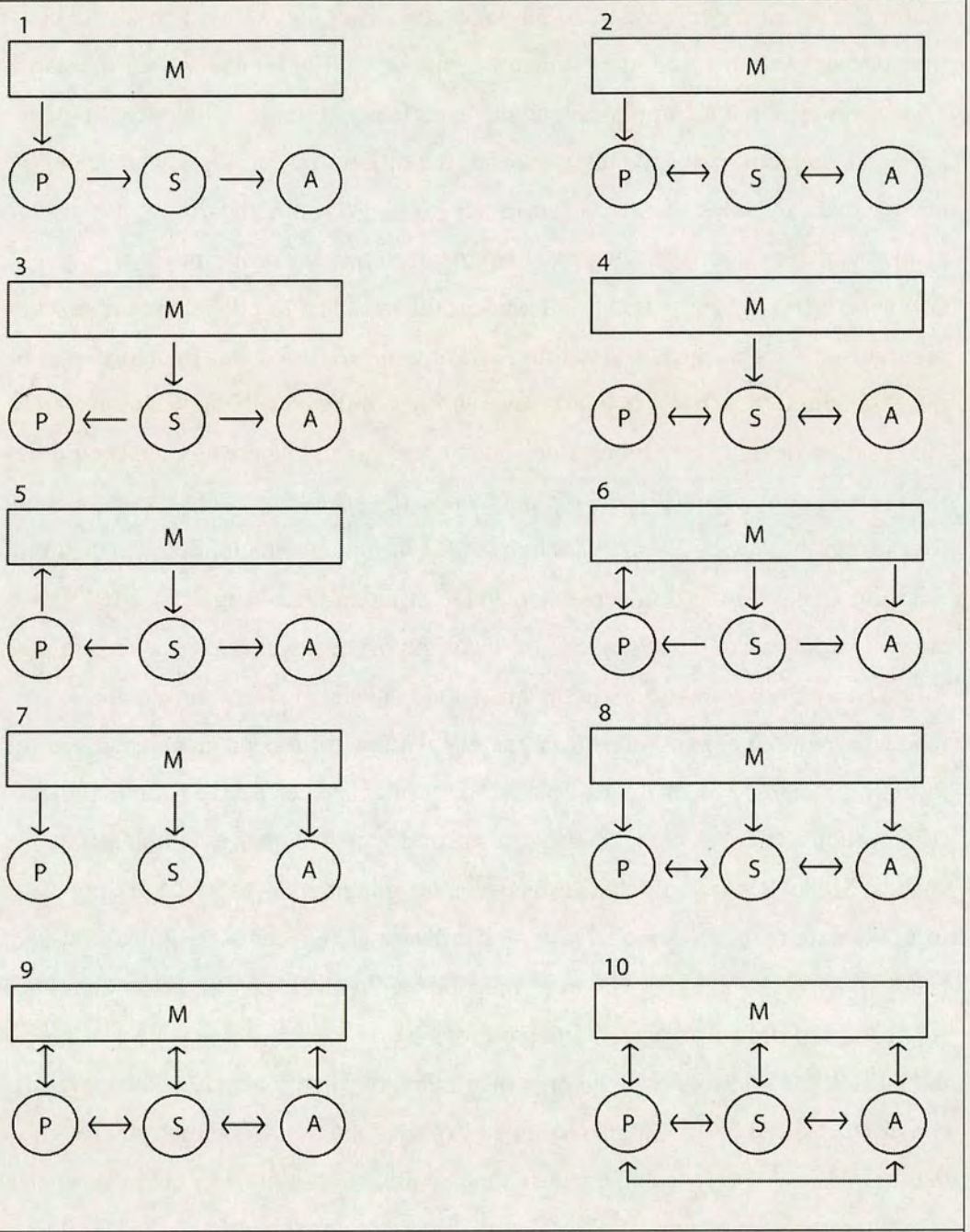


Fig 2.3. Colonization models tested with BAYESMULTISTATE.

1- Stepping-stone, entry Príncipe. **2**- Entry Príncipe + archipelago dispersal. **3**- Stepping-stone, entry São Tomé. **4**- Entry São Tomé + archipelago dispersal. **5**- Entry São Tomé, back colonisation to mainland via Príncipe. **6**- As 5, but multiple colonisations from mainland permitted. **7**- Multiple colonisations from mainland. **8**- Multiple colonisations from mainland + archipelago dispersal. **9**- All 1-step dispersal events allowed. **10**- Unordered: all dispersal events allowed. **M** - mainland (with Bioko); **P** - Príncipe; **S** - São Tomé; **A** - Annobón.

Relationships recovered with this reduced dataset were the same as with the complete dataset (Appendix 2.4). Ancestral character states were reconstructed on the sampled trees using the Bayesian method implemented in BAYESMULTISTATE. The MCMC search had a burn-in period of 100,000 generations, followed by a chain of 15,000,000 generations, with trees sampled every 300 steps. The Bayes factors approach was used to evaluate the different colonisation scenarios.

Estimation of divergence and colonisation times

The assumption of a molecular clock was tested in PAUP* with a likelihood-ratio test (LRT; Felsenstein 1981a). The null hypothesis of a constant rate of evolution is given by the ML tree generated in PAUP* with a clock enforced; the alternative hypothesis is the ML tree generated without the clock constraint. The test was carried out for the ingroup taxa (n = 47). Results supported a molecular clock (-lnL clock-enforced tree = 6603.22, -lnL unconstrained tree = 6577.65, $\chi^2 = 51.04$, d.f. = 45, P = 0.25), and divergence times were estimated from the ML tree with the clock enforced. The molecular clock rate used was calibrated from the 4.66% rate estimated by Warren *et al.* (2006) in the Indian Ocean white-eye radiation. The authors used the molecular clock method of Fleischer *et al.* (1998), which considers that when sister taxa occur on two islands, the age of the younger island represents the maximum age of the split between the lineages and therefore can be used to estimate the minimum rate of divergence.

Colonization date estimates for the Gulf of Guinea islands were considered to be within the range defined by the most recent common ancestor (MRCA) and the most ancient common ancestor (MACA) to the group (Fig. 2.4). The MRCA represents the beginning of divergence among the insular taxa. The MACA represents the time of divergence of the insular taxa from their nearest extant mainland relative. The observed MRCA constitutes the upper limit of colonization date because: i) divergence of island taxa does not necessarily start immediately after colonization and ii) earlier colonizers that diverged into the extant island species may have gone extinct (Vences 2005; Hayward & Stone 2006). In the same way, the MACA is the oldest possible date of colonization as it assumes that colonization was simultaneous

with the origin of the insular lineage and that there was no mainland taxon more closely related to the insular ones that went extinct (Vences 2005; Hayward & Stone 2006). Divergence dates of the nodes of interest were estimated using a Bayesian MCMC approach as implemented in BEAST version 1.3 (Drummond & Rambaut 2003) with the results and adequacy of the sampling examined in TRACER. The best-supported model of sequence evolution determined previously by the Bayes factors approach was used. MCMC chains were 10 million generations in length, with parameters sampled every 1000 steps, and the first 1 million steps discarded as a burn-in.

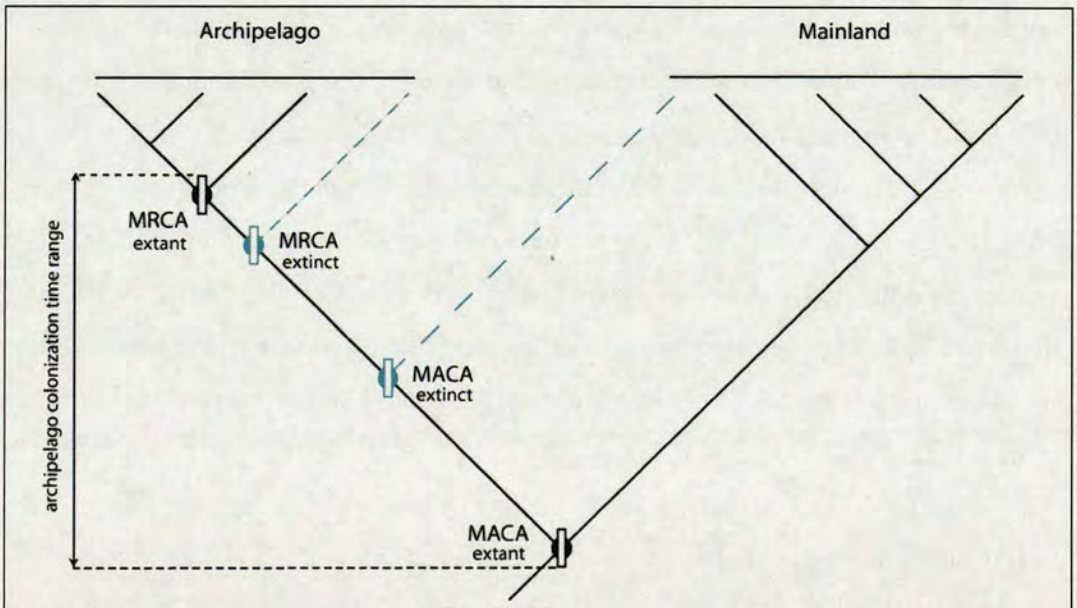


Fig 2.4. The colonization date of an archipelago will be located anywhere between the time of most ancient common ancestor (MACA) and the most recent common ancestor (MRCA). Time of MACA assumes that the appearance of the insular lineages was simultaneous with colonization. Time of MRCA does not take into account post-colonization demographic effects on genetic diversity. Even if the true MACA and/or MRCA went extinct (or were not sampled) they would be located within this interval. Adapted from Vences (2005) and Hayward & Stone (2006).

Phylogenetic analyses - microsatellite data

Microsatellites have the potential to offer many independent sources of phylogenetic information. Nevertheless there at least are two limitations for their use in phylogenetic inference. Their fast rate of evolution, which makes them so attractive for population genetics, is expected to lead to homoplasy problems because microsatellite repeat regions are bounded in size (Goldstein *et al.* 1995; Goldstein & Pollock 1997). Microsatellites are therefore more appropriate to reconstruct phylogenetic relationships among populations and closely related species complexes (Franck *et al.* 1998; Harr *et al.* 1998; Ritz *et al.* 2000; Richard & Thorpe 2001; Kankare & Shaw 2004; Chirhart *et al.* 2005; Krüger *et al.* 2005). In this study, microsatellites were used to elucidate the relationships among the Gulf of Guinea white-eyes, a priori a group of closely related species but with divergence levels probably substantially higher than those found within species or at the species boundary. In such situations, microsatellites might still perform well, as was the case for the Darwin's finches (Petren *et al.* 1999). A much more serious problem for the use of microsatellites in phylogenetic inference is that, contrary to the case of genomic sequences, there is no real understanding of their mode of evolution (Ellegren 2004; Felsenstein 2004). Therefore not only it is impossible to correct for homoplasy but it is impossible to tell how closely related two distinct alleles of a given locus are. Phylogenetic inference with microsatellites should therefore be approached with care, and results interpreted from a phenetic perspective.

All methods for inferring species relationships with microsatellites calculate pairwise distances from the allele frequencies which are then used as the input for clustering methods like neighbour-joining or UPGMA. Many estimators have been proposed, all falling into two categories: i) those derived from estimators used for sequence data, based on the infinite-allele model (IAM); ii) those that are based on the stepwise mutation model (SMM), specifically created as an approximation to an hypothesis of microsatellite evolution (Takezaki & Nei 1996). The measure that performs better under the SMM is the $(\delta\mu)^2$, as it is robust to departures from the SMM and to population fluctuations, and maintains linearity with time allowing the

estimation of divergence times when mutation rates are known (Goldstein *et al.* 1995). Nevertheless, even if $(\delta\mu)^2$ can be successful in recovering branch lengths it is generally inadequate in recovering the correct topology (Takezaki & Nei 1996). The best metrics for topology reconstruction are Nei's *et al.* (1983b) D_A and Cavalli-Sforza & Edwards' (1967) chord distance D_C, both IAM-based (result of simulations: Takezaki & Nei 1996).

Overall allelic diversity of the six polymorphic loci was characterised with Nei's (1987) estimators of heterozygosity calculated with FSTAT version 2.9.3.2. (Goudet 2002). The program MICROSAT version 1.5 (Minch *et al.* 1995) was used to summarise allelic diversity for each taxon (or population in the case of *Z. stenocricotus* from Bioko and Cameroon) and to estimate $(\delta\mu)^2$ and D_C distances. Bootstrap pseudo-replicates (c. 400, as only up to 462 can be made for 6 alleles) were used to estimate standard deviations. Distance matrices were used to construct NJ phenograms in PAUP*. From the arguments above, it was decided a priori that topologies based on D_C distances were more reliable than those based on $(\delta\mu)^2$ distances.

Morphological comparisons

Adaptive morphometric traits linked to diet and flight performance (Grant 1986; Smith & Girman 2000) were measured: wing length, tail length, tarsus length, upper mandible length, bill width, and mass. Mass was measured to the nearest 0.5 g with a 50-g Pesola spring balance. Wing and tail length were measured with a standard wing ruler to the nearest 0.5 mm. Bill and tarsus measurements were taken with a digital calliper to the nearest 0.1 mm. Measurements were taken as follows: wing length (flattened), from the carpal joint to the tip of the longest primary; tail length, from the uropygial gland to tip of the central rectrix; tarsus length, from the tibiotarsus joint to the distal end of the tarsometatarsus when the foot is held to the leg; upper mandible length, from where the culmen enters the feathers of the head to tip; bill width at the anterior end of nares. The different variables were compared between sympatric species pairs. Because morphological data were often scarce for at least one species of the pair, testing for significant differences was not possible.

RESULTS

Phylogenetic analyses - mitochondrial sequence data

Sequence characteristics

All sequences aligned without gaps or insertions and translated appropriately. The three mtDNA sequences used had similar levels of variability, with the cyt *b* being slightly less variable than ATPase6 and ND3 (informative sites: 17%, 23% and 22% respectively; Table 2.1). The combined dataset of 1302 bp had 375 variable sites of which 283 (75%) were informative. All genes show a deficiency of G nucleotides, which is more severe at third positions, and an excess of C nucleotides, typical of many avian mtDNA genes (e.g., Baker & Marshall 1997; Moore & DeFilippis 1997; Voelker & Spellman 2003; Klicka *et al.* 2005). Among-site rate heterogeneity (described by the α parameter of the gamma distribution) differed strikingly between codon positions in the ATPase6 and ND3 genes (Table 2.1).

The average uncorrected pairwise distances calculated from the combined dataset was 5.6%, with a maximum value of 8.1% (between *Z. modestus* and *Z. atricapillus*). The minimum distance between taxa was 0.6% between the subspecies *kikuyensis* and *kulawensis* of *Z. poliogaster.* The minimum difference between two taxonomic species was 1.4% between *S. lugubris* and *S. leucophaeus*. Within-species divergence varied from 0.08% (*S. brunneus*) to 1.4% in *Z. stenocricotus.* This relatively high within-species divergence was not found in any other species: the next highest within-species divergence was 0.3% in *Z. m. maderaspatanus.* Therefore, whenever the dataset was reduced to one individual per taxon (for computational reasons), the four samples of *Z. stenocricotus* were maintained. Saturation plots showed saturation of 3rd codon transitions in the three genes (not shown). Base frequencies of the three genes and of each gene codon position were homogeneous across taxa ($P \geq 0.98$ in all χ^2 tests). Both incongruence tests supported congruence of the phylogenetic signal of the three genes (ILD: $P > 0.72$; Bayes factor in favour of linked topology, $2\log B_{12} = 158.12 >> 10$, value considered to provide very strong evidence against the alternative model – Table 2.2).

Phylogenetic inference

The substitution models selected for each gene by MRMODELTEST were all parameter rich. The GTR + Γ + I was the best fit for the ATPase6, the GTR + Γ for the ND3, and the HKY + Γ + I for the cyt *b*. For the combined dataset, partition by three codon positions had the best log-likelihood (Table 2.2), and was the partition that fared better in comparisons with all others (Table 2.3). Partition by gene was the worst, both with linked and unlinked gene topologies (Table 2.2). The HKY + Γ + I model was the one that best described both positions 1 and 2, and GTR + Γ + I was the best fit for position 3. Partition of the combined dataset into two partitions (the two first positions together) had the second highest likelihood, and evidence against it as assessed by Bayes factors was not as strong as for the other cases (Table 2.3).

Table 2.1.
Sequence characteristics of the 3 mtDNA genes for all ingroup taxa

	N bases	Var	Info	%A	%C	%G	%T	R. rate	α	pinv	ts/tv
ATP6											
All	661	202	154	30.6	36.8	9.3	23.3	1.091	2.616	0.628	8.453
1st	220	40	31	28.5	38.5	17.0	15.9	0.434	0.012	0.000	31.867
2nd	220	3	1	14.5	29.8	8.9	46.8	0.014	equal	0.000	2.013
3rd	221	159	122	46.2	36.0	4.6	13.2	2.545	2.681	0.000	6.572
Cyt b											
All	290	69	51	26.1	32.5	17.4	24.0	0.727	0.976	0.625	7.070
1st	96	13	7	21.2	24.8	30.4	23.7	0.548	0.616	0.710	7.829
2nd	97	3	2	21.6	21.8	17.5	39.0	0.066	0.681	0.817	-
3rd	97	53	42	35.0	47.7	3.9	13.4	2.382	0.745	0.000	6.562
ND3											
All	351	104	78	31.0	34.0	10.2	24.8	1.055	3.634	0.656	10.447
1st	117	17	14	27.7	31.1	19.5	21.7	0.400	0.238	0.516	20.815
2nd	117	11	6	17.1	24.5	12.5	45.9	0.215	3.203	0.842	21.369
3rd	117	75	58	41.6	44.0	4.8	9.6	2.384	0.849	0.000	8.114

Var: variable sites - Info: informative sites - R. rate: relative rate - α: shape parameter of the gamma distribution - pinv: proportion of invariable sites - ts/tv: transition to transversion ratio. All parameters estimated in PAUP* with maximum likelihood from a neighbour-joining topology built with LogDet corrected distances.

Table 2.2.

Estimated model likelihood (harmonic mean) of the models tested.

Model	partitions	DNA model	parameters	log-likelihood
[ATP6-ND3-Cytb]	1	GTRIΓ	6	-7103.85
ATP6,ND3,Cytb [1]	3	GTRIΓ (1) GTRΓ (2) HKYIΓ (3)	14	-7164.47
ATP6,ND3,Cytb [2]	3	GTRIΓ (1) GTRΓ (2) HKYIΓ (3)	17	-7243.53
[pos1-pos2],pos3	2	GTRIΓ	11	-6798.93
pos1,pos2,pos3	3	HKYIΓ (1,2) GTRIΓ (3)	15	-6678.08

Square brackets bound elements of a partition. Commas separate partitions. Pos i = codon position i.
[1]: the 3 genes share the same topology. [2]: each gene can have an independent topology.

Table 2.3.

Summary of Bayes factor tests for the models of table 2.2. Entries are twice the log of the Bayes factors in the comparison between models M_1 and M_2 ($2\log B_{12}$). Positive values indicate support for the row model (M_2); negative values indicate support for the column model (M_1).

Model	ATP6,ND3,Cytb [1]	ATP6,ND3,Cytb [2]	[pos1-pos2],pos3	pos1,pos2,pos3
[ATP6-ND3-Cytb]	121.24	279.36	-609.84	-851.54
ATP6,ND3,Cytb [1]	0	158.12	-731.08	-972.78
ATP6,ND3,Cytb [2]		0	-889.20	-1130.90
[pos1-pos2],pos3			0	-241.70

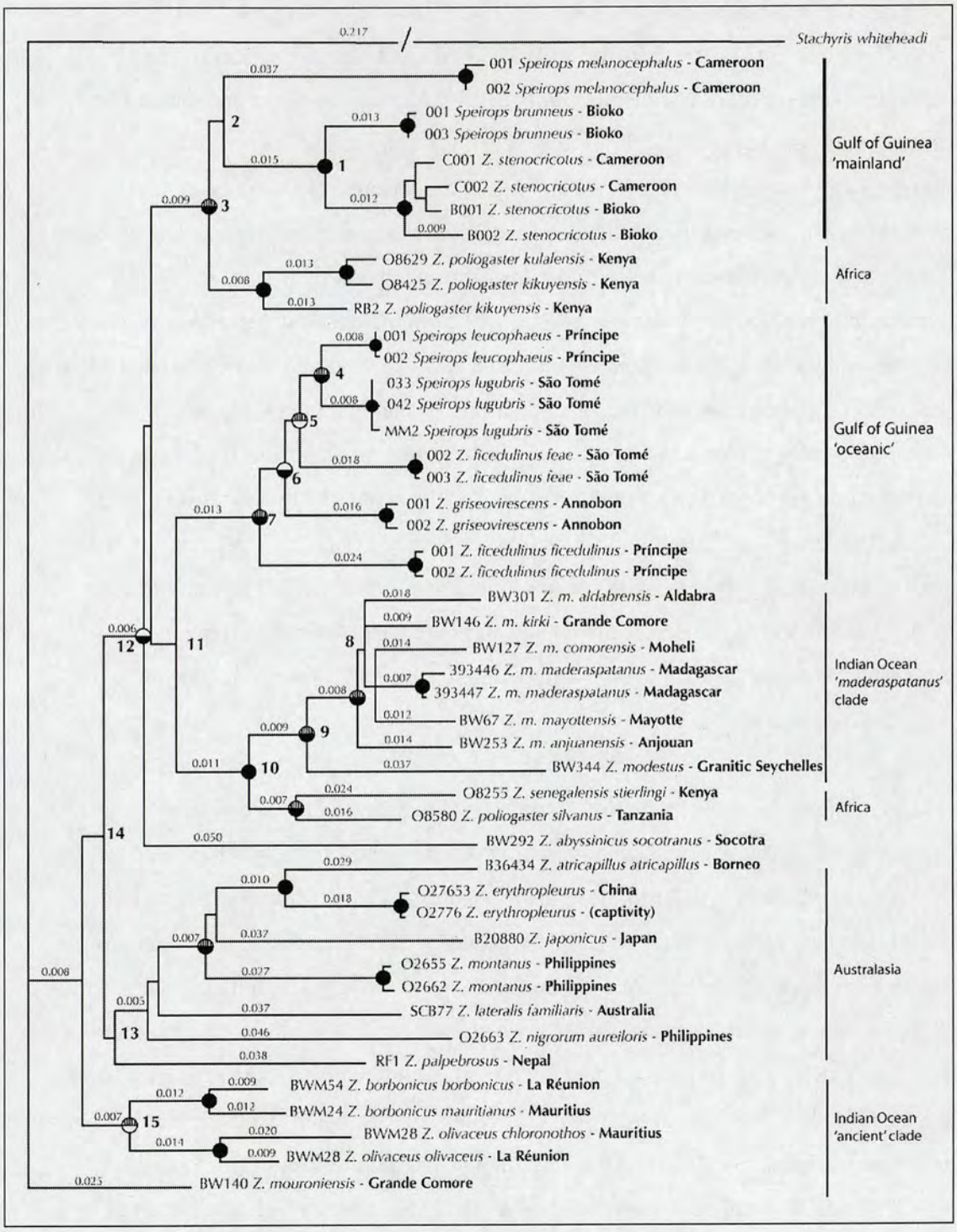
Square brackets bound elements of a partition. Commas separate partitions. Pos i = codon position i.
[1]: the 3 genes share the same topology. [2]: each gene can have an independent topology.

Parsimony analysis recovered 8 most parsimonious trees that differed only in the placement of two terminal nodes, which were not supported by bootstrap (< 70%). The ML tree had the same topology as the Bayesian consensus tree. When support for nodes is taken into account (bootstrap ≥ 70%, posterior probability > 0.95) all trees had very similar topologies (Fig. 2.5). The only difference occurred in the handling of the Asian species. In MP, *Z. nigerrimus* was excluded from the remaining Zosteropidae ingroup with a 76% bootstrap support, with the other Asian species except for *Z. palpebrosus* forming a clade with an 82% bootstrap support. In

the ML tree, all the Asian species form a clade that includes the Australian *Z. lateralis*, but it received no support. Overall, there was little discrepancy between bootstrap and posterior probability support (Fig. 2.5). The clade of the Mascarene taxa (*Z. borbonicus* spp and *Z. olivaceus* spp) was well supported by bootstrap (85%) but not by posterior probabilities (0.82). Within the Gulf of Guinea oceanic island species, the placement of *Z. griseovirescens* is not well supported by bootstrap (64%) but it is by posterior probability (0.98), whereas the opposite happens for the placement of *Z. ficedulinus feae* (84% and 0.81 respectively).

The same relationships were recovered for the species included in the study of the Indian Ocean white-eyes radiation (Warren *et al.* 2006), with the few differences being due to lower resolution of the dataset used here (1302 bp instead of 1514 bp), and will therefore not be discussed. In both cases *Z. mouroniensis* from the Grande Comore forms a basal polytomy.

Fig. 2.5. (next page). Maximum-likelihood tree obtained from the combined data set using the successive approximations approach. Bayesian inference gave the same topology. Maximum-parsimony recovered the same topology except for the position of the Asian taxon *Z. nigrorum*, which came as a basal polytomy. ML branch lengths are indicated when ≥ 0.005. Nodal support is indicated by circles, with parsimony bootstrap on the top hemisphere (striped ≥ 75%, black ≥ 90 %), and Bayesian posterior probabilities on the lower hemisphere (black > 0.95). Numbered nodes refer to those for which divergence dates have been estimated (Table 2.5).


0.217
Stachyris whiteheadi
0.037
001 Speirops melanocephalus - Cameroon
002 Speirops melanocephalus - Cameroon
2
0.013
001 Speirops brunneus - Bioko
003 Speirops brunneus - Bioko
0.015
1
C001 Z. stenocricotus - Cameroon
C002 Z. stenocricotus - Cameroon
0.009
3
0.012
B001 Z. stenocricotus - Bioko
0.009
B002 Z. stenocricotus - Bioko
Gulf of Guinea 'mainland'
0.013
O8629 Z. poliogaster kulalensis - Kenya
0.008
O8425 Z. poliogaster kikuyensis - Kenya
0.013
RB2 Z. poliogaster kikuyensis - Kenya
Africa
0.008
001 Speirops leucophaeus - Príncipe
002 Speirops leucophaeus - Príncipe
4
033 Speirops lugubris - São Tomé
0.008
042 Speirops lugubris - São Tomé
5
MM2 Speirops lugubris - São Tomé
0.018
002 Z. ficedulinus feae - São Tomé
6
003 Z. ficedulinus feae - São Tomé
0.013
0.016
001 Z. griseovirescens - Annobon
7
002 Z. griseovirescens - Annobon
0.024
001 Z. ficedulinus ficedulinus - Príncipe
002 Z. ficedulinus ficedulinus - Príncipe
Gulf of Guinea 'oceanic'
0.018
BW301 Z. m. aldabrensis - Aldabra
0.009
BW146 Z. m. kirki - Grande Comore
0.006
12
11
8
0.014
BW127 Z. m. comorensis - Moheli
0.007
393446 Z. m. maderaspatanus - Madagascar
0.008
393447 Z. m. maderaspatanus - Madagascar
0.012
BW67 Z. m. mayottensis - Mayotte
0.009
9
0.014
BW253 Z. m. anjuanensis - Anjouan
0.011
10
0.037
BW344 Z. modestus - Granitic Seychelles
Indian Ocean 'maderaspatanus' clade
0.024
O8255 Z. senegalensis stierlingi - Kenya
0.007
0.016
O8580 Z. poliogaster silvanus - Tanzania
Africa
14
0.050
BW292 Z. abyssinicus socotranus - Socotra
0.029
B36434 Z. atricapillus atricapillus - Borneo
0.010
0.018
O27653 Z. erythropleurus - China
O2776 Z. erythropleurus - (captivity)
0.007
0.037
B20880 Z. japonicus - Japan
0.008
0.027
O2655 Z. montanus - Philippines
O2662 Z. montanus - Philippines
0.005
0.037
SCB77 Z. lateralis familiaris - Australia
13
0.046
O2663 Z. nigrorum aureiloris - Philippines
0.038
RF1 Z. palpebrosus - Nepal
Australasia
0.009
BWM54 Z. borbonicus borbonicus - La Réunion
0.012
0.012
BWM24 Z. borbonicus mauritianus - Mauritius
0.007
15
0.020
BWM28 Z. olivaceus chloronothos - Mauritius
0.014
0.009
BWM28 Z. olivaceus olivaceus - La Réunion
0.025
BW140 Z. mouroniensis - Grande Comore
Indian Ocean 'ancient' clade

The Gulf of Guinea white-eyes are not monophyletic, being grouped in two well-supported clades derived from distinct African ancestors (bootstrap support of 78 and 84%, posterior probabilities of 1 for both). One clade groups the taxa from Cameroon and Bioko, and the other groups all taxa from the three Gulf of Guinea oceanic islands. Because of the proximity of Bioko to the mainland and the fact that it was connected to it several times in the past, the Cameroon-Bioko clade will be designated the 'mainland clade', as opposed to the 'oceanic clade'. This result supports the hypothesis that archipelago radiation, rather then multiple colonizations from the mainland, was the major process of lineage diversification in the Gulf. Under these relationships, the four species that form the endemic genus *Speirops* are not monophyletic, and *S. melanocephalus* and *S. lugubris* are not closely related as has sometimes been assumed (see Introduction). In the mainland clade, *S. melanocephalus* makes a basal polytomy, and *Z. stenocricotus* from Bioko and the mainland do not form distinct clades. In the oceanic clade, the two subspecies of *Z. ficedulinus* are paraphyletic, with *Z. griseovirescens* in between. These relationships are not clear as *Z. griseovirescens* makes a polytomy either with *Z. f. ficedulinus* or *Z. f. feae*. Shimodaira-Hasegawa (SH) and Bayes factors tests rejected the two classical taxonomy views: the monophyly of *Speirops* and a sister relationship between *S. melanocephalus* and *S. lugubris* (Table 2.4).

The geographic mode of speciation (sympatry x allopatry) of the species co-occurring on the same island cannot be fully determined from the phylogenetic tree. *S. brunneus* (Bioko) is sister to the *Z. stenocricotus* clade that has populations both in Bioko and Cameroon, so that either of the modes would be possible. *Z. ficedulinus ficedulinus* from Príncipe is not sister to *S. leucophaeus* with which it co-occurs, and therefore the sympatry-allopatry problem is not an issue. Nevertheless, Shimodaira-Hasegawa (SH) tests could not reject the hypothesis of sympatric speciation for these species (Table 2.4). *Z. ficedulinus feae* from São Tomé is sister to a clade that comprises a species of *Speirops* from São Tomé and another from Príncipe, which could also fit with either of the speciation modes. Speciation could have occurred in sympatry followed by a colonisation to Príncipe, or in allopatry if the species from São Tomé dispersed first to Príncipe where it diverged and then back-colonised São

Tomé. Again, SH tests could not reject the sympatric speciation hypothesis. Bayes factors were against both sympatric speciation hypotheses, but the values obtained were the smallest from all Bayes factors tests performed in this study (Table 2.4). The comparison of the model of sympatric speciation of *Z. f. feae* and *S. lugubris* gave a $2\log B_{12}$ value of 6.24. This is close to the valeu below which evidence for the alternative model is positive (but not strong); between 0 and 2, both models would be considered equivalent (Kass & Raftery 1995). As these classes of support only constitute guidelines, the small value obtained suggests that the data are compatible with either model.

Table 2.4.
Shimoidara-Hasegawa (SH) and Bayes factors (BF) tests of alternative phylogenetic hypotheses

	SH			BF	
Constraint	-lnL[1]	Δ-lnL	P	-lnL[1]	$2\log B_{12}$
Unconstrained tree	7000.82	(best)		6678.08	
Monophyly *Speirops*	7116.45	115.63	0.001	6745.52	134.88
S. melanocephalus and *S. lugubris* as sisters	7957.61	56.79	0.001	6708.92	61.68
Z. f. ficedulinus and *S. leucophaeus* as sisters	7012.29	11.47	0.382	6695.21	34.26
Z. f. feae and *S. lugubris* as sisters	7006.31	5.89	0.619	6681.20	6.24

[1]:Log-likelihood of the ML tree. [2]:Harmonic mean of the log-likelihood of the trees sampled with BI.
SH test: significant values ($P<0.05$) indicate a rejection of the proposed model.
BF: $2\log B_{12} > 6$ *and* $2\log B_{12} > 10$ indicates strong *and* very strong evidence against proposed model.

None of the methods of ancestral state reconstruction could determine the entry point to the Gulf of Guinea oceanic archipelago. Under an ordered parsimony analysis, Príncipe, São Tomé, and mainland Africa could all have been the area of the MRCA to the clade (Fig. 2.6). Colonisation of Annobón, and a second colonisation of Príncipe occurred from São Tomé (Fig. 2.6). In an unordered parsimony reconstruction (i.e., every dispersal event equally likely) every area above a node could have been the ancestral area (not shown), so that any island could have been the entry point. The ancestral areas method reflects this uncertainty by assigning equal probabilities of Príncipe and São Tomé being the ancestral area of the oceanic clade (Table 2.5). Bayesian inference using MRBAYES (searching only on the trees that contain the node of interest) infers the African mainland as the area of the MRCA of the oceanic clade, and São Tomé as the area of the next node up the tree (Table 2.6). BAYESMULTISTATE, the method that best takes into account the phylogenetic uncertainty, also estimated very similar posterior probabilities for Príncipe and São Tomé being the area of the ancestor to the oceanic clade (Fig. 2.7). These results suggest that Príncipe and São Tomé were colonised at the same time by the same ancestor from Africa. A contemporary colonisation of both islands could have followed a stepping-stone route (e.g. Príncipe → São Tomé), or derive from two independent colonisations from mainland. Tests of different colonisation scenarios (Fig. 2.3) performed with BAYESMULTISTATE favour the model of a single colonisation (Models 1-4, Tables 2.7 and 2.8) followed by dispersal events between neighbouring islands. The worst model (after the unordered model) was the one that considered all island taxa to have derived from mainland colonisations (model 7, Tables 2.7 and 2.8). When inter-island dispersal events are added to this model, its likelihood increases significantly (model 8, Tables 2.7 and 2.8), which suggests again that dispersal between islands was crucial in the diversification of this group. Overall, the different lines of evidence support a contemporary colonisation of the islands of São Tomé and Príncipe by the same ancestor, followed by inter-island dispersal events. São Tomé, the central island, appears to have been the source of colonisers both for Annobón and Príncipe.

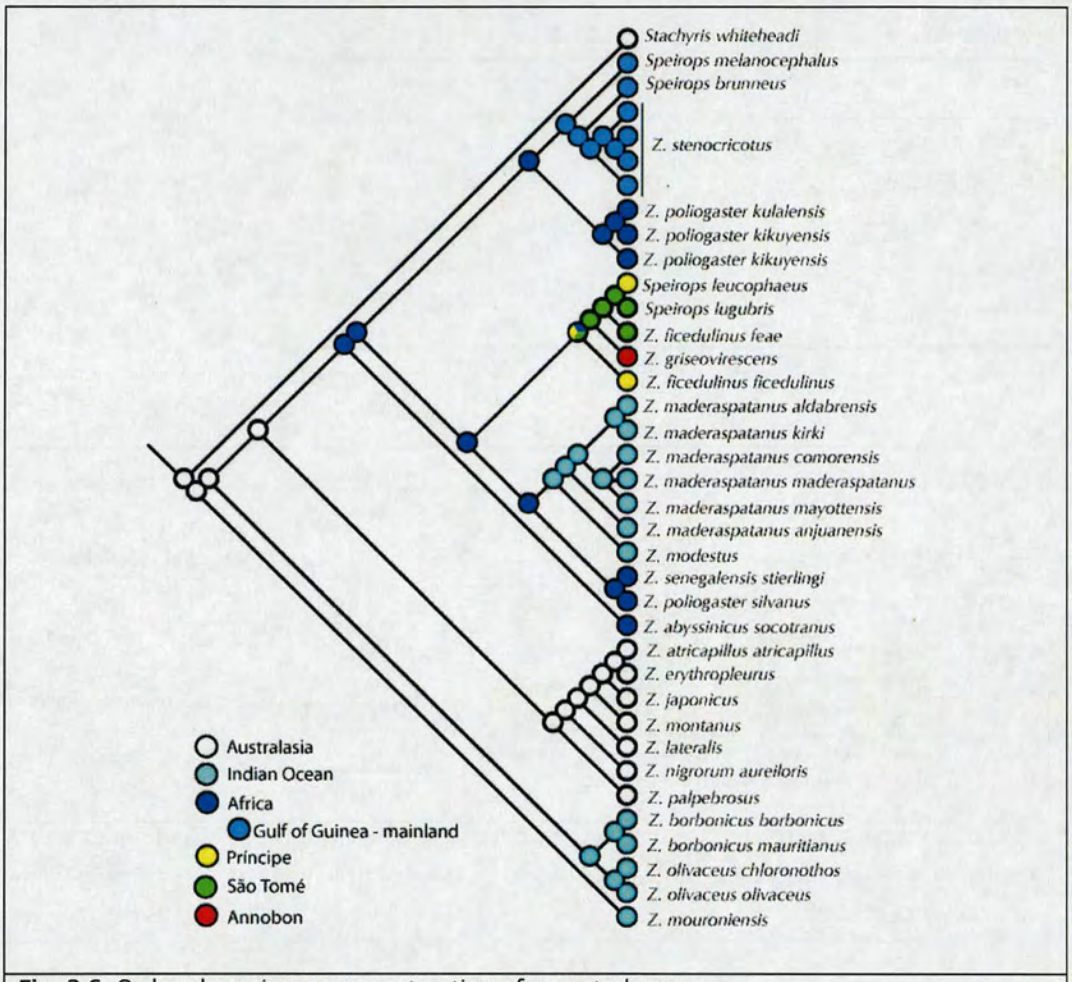


Fig. 2.6. Ordered parsimony reconstruction of ancestral areas.

Table 2.5.

Application of the ancestral areas method (Bremer 1992) to the Gulf of Guinea oceanic islands clade, (the node not supported by Bayesian inference was collapsed for the analysis; Fig. 2.5).

Area	Gains	Losses	Gains/ Losses	Ancestral area probability
Príncipe	2	3	0.66	1
São Tomé	3	3	0.66	1
Annobón	1	3	0.33	0.5

Table 2.6.

Bayesian posterior probabilities for the location of the MRCA of the Gulf of Guinea oceanic islands clade (node 7, Fig 2.5) and of the next node up the tree (node 6). Probabilities estimated in MRBAYES; highest values in bold.

Area	Oceanic clade	Node 6
Africa mainland	**0.706**	0.025
Cameroon-Bioko	0.007	0.002
Príncipe	0.147	0.173
São Tomé	0.121	**0.766**
Annobón	0.019	0.033

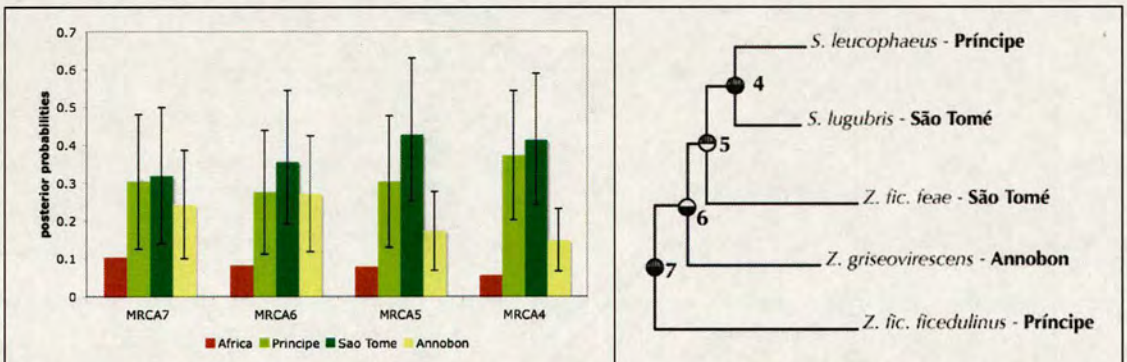


Fig. 2.7. Ancestral state probabilities, with standard deviations, for the MRCA of nodes 4 to 7 (depicted right) as estimated with BAYESMULTISTATE. Standard deviations for 'Africa' probabilities not shown for clarity.

Table 2.7.

Estimated model likelihood (harmonic mean) of the colonisation models tested. See Fig. 2.3 for a schematic representation of the different models.

Model	Name	Parameters	log-likelihood
1	Stepping-stone, entry Príncipe	3	-22.25
2	2 with archipelago dispersal	5	-22.01
3	Stepping-stone, entry São Tomé	3	-22.16
4	3 with archipelago dispersal	5	-22.03
5	*Speirops* back-colonisation 1	4	-22.84
6	*Speirops* back-colonisation 2	6	-24.21
7	Multiple colonisations from mainland	3	-30.09
8	7 with archipelago dispersal	7	-22.82
9	One-step dispersal	10	-25.60
10	Unordered	12	-30.34

Table 2.8.

Summary of Bayes factor tests for the models of table 2.7. Entries are twice the log of the Bayes factors in the comparison between models M_1 and M_2 ($2\log B_{12}$). Positive values indicate support for the row model (M_2); negative values indicate support for the column model (M_1). $0 < |2\log B_{12}| < 2$: models are equivalent (shaded cells); $2 < |2\log B_{12}| < 6$: positive support; $6 < |2\log B_{12}| < 10$: strong support; $|2\log B_{12}| > 10$: very strong support. Model numbers as in Table 2.7 and Fig. 2.3.

Model	2	3	4	5	6	7	8	9	10
1	-0.48	-0.18	-0.43	1.19	3.93	15.68	1.15	6.73	16.18
2		0.29	0.05	1.67	4.41	16.15	1.62	7.20	16.66
3			-0.25	1.38	4.12	15.86	1.33	6.91	16.37
4				1.62	4.36	16.11	1.58	7.16	16.61
5					2.74	14.48	-0.04	5.53	14.99
6						11.74	-2.78	2.79	12.25
7							-14.53	-8.95	0.50
8								5.58	15.03
9									9.46

Estimation of divergence and colonisation times

Warren *et al.* (2006) obtained an average pairwise substitution rate of 4.66% Ma^{-1} for their dataset consisting of ATPase6, ATPase8, ND3 and cyt *b*. This rate was estimated from the node that separates the lowland Grande Comore white-eye *Z. madesraspatanus kirki* from the other white-eyes in the *madesraspatanus* clade using the maximum estimate of the volcanic origin of the island (0.5 Ma) as the time of the earliest possible colonization. This calibration could not have been done with the three-gene dataset used in this study as the Anjouan white-eye *Z. m. anjuanensis* position is not resolved and forms a polytomy with *Z. m. kirki* (Fig. 2.5). In such a situation the maximum age for the node would be 11.5 Ma, the maximum age of Anjouan, of little use for calibration purposes. The 4.66% rate was corrected for the data set used in this study, which did not include the ATPase8, by comparing the genetic distances (from node to tip on the ML clock-enforced tree) of the nodes in common and obtaining the average difference ratio. This gave an average pairwise substitution rate of 4.95% for the ATPase6, ND3 and cyt *b* combined sequence

(giving a 2.475% rate along each lineage, the value used for the estimation of divergence times). Divergence time estimates obtained by Warren *et al.* (2006) for nodes in common were always within the 95% confidence interval of the Bayesian estimates of this study (Table 2.9). The time range for different colonisation events (Table 2.10) was determined with the MACA/MRCA model of colonisation time estimates (Fig. 2.4), and by taking into consideration the inferred ancestral area(s) of the nodes of interest. For example, the colonisation of São Tomé could have happened anywhere between the MACA to the oceanic clade (node 11) and the MRCA of *Z. griseovirescens* and the species above (node 6), rather than between the MACA and MRCA of the clade that contains the most ancient species present in São Tomé, *Z. ficedulinus feae* (nodes 6 and 5).

Table 2.9.

Bayesian divergence time estimates for the most recent common ancestor (MRCA) of different clades, with lower and upper 95% confidence intervals. For comparison, estimates obtained by Warren *et al.* (2006) in their study of the diversification of Indian Ocean Zosteropidae are presented (asterisks indicate nodes with fewer species in Warren's *et al.* study). Warren *et al.* estimates are based on a dataset with 212 extra bp, and were obtained by directly dividing the genetic distance from the node of interest to the tip (branch length on the clock-enforced tree) by the substitution rate per lineage. Node numbers and clade names correspond to those presented in Fig. 2.5. Africa* consists of the clade of all African species except those of the basal Indian Ocean species (*Z. mouroniensis* and Mascarene group). Dates are in million years before present.

Node	Clade	Time for the MRCA	Warren *et al.*
1	*S. brunneus* & *Z. stenocricotus*	0.462 – **0.632** – 0.816	
2	Gulf of Guinea 'mainland' (GGM)	0.960 – **1.206** – 1.446	
3	GGM - Africa	1.022 – **1.254** – 1.470	
4	'Oceanic' *Speirops*	0.224 – **0.362** – 0.510	
5	4 + *Z. ficedulinus feae*	0.405 – **0.576** – 0.740	
6	5 + *Z. griseovirescens*	0.520 – **0.792** – 0.886	
7	Gulf of Guinea 'oceanic' (GGO)	0.696 – **0.900** – 1.116	
8	Calibration	0.482 – **0.616** – 0.760	**0.500**
9	Indian Ocean *maderaspatanus* (IOM)	0.792 – **0.984** – 1.186	**1.020**
10	IOM-Africa	0.982 – **1.204** – 1.396	**1.200**
11	GGO-IO-Africa	1.342 – **1.606** – 1.894	
12	Africa*	1.606 – **1.872** – 2.178	
13	Australasia	1.480 – **1.836** – 2.240	**1.620**
14	Africa*-Australasia	1752 – **2.006** – 2.286	**1.800***
15	Indian Ocean Mascarenes	0.846 – **1.102** – 1.380	**1.220**
16	Zosteropidae	1754 – **2.006** – 2.286	**1.840***

Table 2.10.

Estimates of the time range for different colonisation events, using the MACA/MRCA model (Fig. 2.4). Mean MRCA and MACA estimates are presented. Africa* consists of the clade of all African species except those of the basal Indian Ocean species (*Z. mouroniensis* and Mascarene group). Dates are in million years before present.

Colonization Event	Time range	MRCA	MACA
Gulf of Guinea oceanic archipelago (= Príncipe 1)	0.696 – 1.894	0.900	1.606
São Tomé	0.520 – 1.894	0.792	1.606
Annobón	0.520 – 0.886	0.900	na
Príncipe 2	0.224 – 0.740	0.362	0.576
Gulf of Guinea - Bioko	0.462 – 1.446	0.632	1.206
Gulf of Guinea – Bioko/Cameroon	0.960 – 1.470	1.206	1.254
Indian Ocean *maderaspatanus*	0.792 – 1.396	0.984	1.204
Indian Ocean Mascarenes	0.846 – 2.286	1.102	2.006
Africa*	1.606 – 2.286	1.872	2.006

Phylogenetic analyses - microsatellite data

Of the 14 microsatellites tested for cross-amplification in 12 individuals from five Gulf of Guinea taxa, only six were genotyped for the entire data set used for phylogenetic inference (ZL12, ZL18, ZL22, ZL44 ASE19, ASE64). This data set comprised 77 individuals, representing all eight taxa (nine populations) of the Gulf of Guinea and the outgroup, *Z. capensis*. Of the remaining eight microsatellites, two did not amplify properly, three were monomorphic and three had only two alleles. ZL12, ZL22 and ASE19 amplified as di-nucleotide repeats, whereas ZL18, ZL44 and ASE64 amplified as tetra-nucleotide repeats.

Of the six microsatellites used, only ZL12 and ASE64 showed significant levels of variation (Table 2.11 and 2.12). ZL22 had five alleles with three making up 90% of samples (respective frequencies of 60, 20 and 10%). ZL44 had five alleles, but a single one (220) had a frequency of 93%. ASE19 had only three alleles, with one (185) only present twice in the same taxon, and another (187) present at a frequency of 84%. ZL18 had three alleles but was essentially monomorphic for allele 134, with only two instances of allele 138 found in *S. lugubris*, and one instance of allele 142 found in *S. leucophaeus*.

Table 2.11.

Total number of alleles sampled and Nei's (1987) estimators of heterozygosity for the 6 polymorphic microsatellite loci. H_O: observed heterozygosity across all samples; H_T: expected heterozygosity (gene diversity) across all samples; H_S: within-taxon/population gene diversity.

Locus	N Alleles	H_O	H_T	H_S
ZL12	13	0.708	0.85	0.639
ZL18	3	0.03	0.03	0.03
ZL22	5	0.391	0.613	0.314
ZL44	5	0.23	0.291	0.199
ASE19	3	0.03	0.222	0.054
ASE64	12	0.657	0.879	0.722

To take into account the differences in allelic diversity, phenograms of genetic distances were built for 3 different data sets: i) all 6 loci; ii) 5 loci – ASE19 excluded; iii) 3 loci – ZL12, ZL22, ASE64; iv) 2 loci – ZL12, ASE64. Both $(\delta\mu)^2$ and D_C genetic distance estimates had large standard errors (Table 2.13 and 2.14 for the six loci data set). This problem was more extreme for $(\delta\mu)^2$, where the standard error was on average 68% of the estimated value compared to 32% for D_C.

Independently of the data set used, neighbour-joining trees constructed with D_C distances recovered the same 'mainland' and 'oceanic' clades as the mtDNA sequence data (Figs. 2.5 & 2.8). The relationships within the 'oceanic' clade were nevertheless different. In the mtDNA phylogeny, *Speirops lugubris* is located at the tip of the tree, sister to *S. leucophaeus*, and *Zosterops f. ficedulinus* is at the base of the 'oceanic' clade (Fig. 2.5). In the microsatellite NJ tree, these two species change positions (Fig. 2.8). All $(\delta\mu)^2$ trees also inferred a *Z. f. ficedulinus-S. leucophaeus* sister relationship, but additionally they put *Speirops lugubris* together with the mainland species (Fig. 2.8)

Table 2.12. (next page)

Characteristics of six polymorphic microsatellite loci among the Gulf of Guinea Zosteropidae and *Zosterops capensis* (outgroup in phylogenetic inference). *n*: number of individuals genotyped. H_O and H_E observed and expected heterozygosity. Allele size range shown.

		Smel	Sbru	Sleu	Slug	ZsteC	ZsteB	ZficP	ZficS	Zgri	Zcap
ZL12	*n*	*3*	*10*	*10*	*10*	*5*	*5*	*4*	*10*	*10*	*9*
N alleles		3	1	4	5	6	4	2	7	3	4
H_O		1.00	0.00	0.60	0.80	0.60	0.80	1.00	1.00	0.50	0.78
H_E		0.61	0.00	0.55	0.67	0.80	0.66	0.50	0.79	0.64	0.72
Min		98	98	106	96	92	96	106	102	106	98
Max		106	98	112	110	112	112	110	118	110	108
Mode		98	98	106	96	100	98	106	106	106	98
ZL18	*n*	*3*	*10*	*10*	*10*	*5*	*5*	*4*	*10*	*10*	*10*
N alleles		1	1	2	2	1	1	1	1	1	1
H_O		0.00	0.00	0.10	0.20	0.00	0.00	0.00	0.00	0.00	0.00
H_E		0.00	0.00	0.10	0.18	0.00	0.00	0.00	0.00	0.00	0.00
Min		134	134	134	134	134	134	134	134	134	134
Max		134	134	142	138	134	134	134	134	134	134
Mode		134	134	134	134	134	134	134	134	134	134
ZL22	*n*	*3*	*9*	*10*	*10*	*3*	*5*	*4*	*10*	*10*	*9*
N alleles		2	3	2	3	1	2	1	3	1	4
H_O		1.00	0.56	0.20	0.70	0.00	0.60	0.00	0.30	0.00	0.56
H_E		0.50	0.48	0.10	0.54	0.00	0.42	0.00	0.27	0.00	0.56
Min		154	154	150	152	154	154	154	154	154	152
Max		158	158	154	156	154	156	154	158	154	158
Mode		154	156	154	154	154	154	154	154	154	152
ZL44	*n*	*3*	*10*	*10*	*10*	*5*	*5*	*4*	*10*	*10*	*10*
N alleles		2	1	1	2	3	2	1	1	1	2
H_O		1.00	0.00	0.00	0.10	0.60	0.20	0.00	0.00	0.00	0.40
H_E		0.50	0.00	0.00	0.10	0.48	0.18	0.00	0.00	0.00	0.50
Min		220	220	220	214	220	220	220	220	220	216
Max		224	220	220	220	224	224	220	220	220	220
Mode		220	220	220	220	220	220	220	220	220	220
ASE19	*n*	*1*	*10*	*10*	*10*	*5*	*5*	*4*	*10*	*10*	*9*
N alleles		1	2	1	2	1	1	1	2	1	1
H_O		0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00
H_E		0.00	0.00	0.00	0.18	0.00	0.00	0.00	0.18	0.00	0.00
Min		187	187	189	185	187	187	187	187	187	187
Max		187	187	189	187	187	187	187	189	187	187
Mode		187	187	189	187	187	187	187	187	187	187
ASE64	*n*	*3*	*10*	*10*	*10*	*5*	*5*	*4*	*10*	*10*	*10*
N alleles		4	5	3	8	5	3	3	5	5	5
H_O		0.67	0.80	0.50	0.90	0.60	0.20	0.50	1.00	0.80	0.60
H_E		0.67	0.74	0.64	0.83	0.76	0.54	0.41	0.74	0.70	0.57
Min		396	384	380	384	376	384	380	384	384	384
Max		412	416	392	408	396	408	396	400	400	404
Mode		404	412	384	392	388	388	380	388	396	384

Smel: *Speirops melanocephalus* – **Sbru**: *S. brunneus* - **Sleu**: *S. leucophaeus* – **Slug**: *S. lugubris*

ZsteC: *Zosterops stenocricotus*, Cameroon – **ZsteB**:*Z. stenocricotus*, Bioko – **ZficP**: *Z. ficedulinus ficedulinus* (Principe) – **ZficS**: *Z. f. feae* (São Tomé) - **Zgri**: *Z. griseovirescens* – **Zcap**: *Z. capensis*

Table 2.13.

Genetic distances among Gulf of Guinea Zosteropidae and an outgroup based on six microsatellite loci. Above diagonal: $(\delta\mu)^2$ distances (Goldstein *et al.* 1995). Below diagonal: standard error. See Table 2.12 for species names legend.

	Smel	Sbru	Sleu	Slug	ZsteC	ZsteB	ZficP	ZficS	Zgri	Zcap
Smel		0.738	6.646	1.109	1.268	1.355	6.896	4.456	3.079	3.582
Sbru	0.433		7.348	1.215	1.241	1.112	7.468	5.775	4.527	2.945
Sleu	3.976	4.232		2.781	2.769	2.825	0.183	0.498	1.148	1.488
Slug	0.819	0.814	1.447		0.079	0.064	2.805	1.724	1.135	0.725
ZsteC	0.972	0.680	1.496	0.034		0.044	2.811	1.881	1.305	0.682
ZsteB	1.144	0.631	1.590	0.029	0.015		2.834	2.004	1.492	0.562
ZficP	4.457	4.391	0.144	1.655	1.641	1.688		0.512	1.228	1.368
ZficS	2.519	4.120	0.311	1.267	1.446	1.654	0.463		0.197	1.450
Zgri	1.668	3.545	0.893	0.972	1.137	1.314	1.151	0.147		1.485
Zcap	2.817	1.667	1.020	0.595	0.479	0.368	0.999	1.137	0.895	

Table 2.14.

Genetic distances among Gulf of Guinea Zosteropidae and an outgroup based on six microsatellite loci. Above diagonal: D_C distances (Cavalli-Sforza & Edwards 1967). Below diagonal: standard error. See Table 2.12 for species names legend.

	Smel	Sbru	Sleu	Slug	ZsteC	ZsteB	ZficP	ZficS	Zgri	Zcap
Smel		0.378	0.659	0.535	0.307	0.368	0.457	0.469	0.447	0.468
Sbru	0.113		0.597	0.435	0.417	0.298	0.429	0.409	0.394	0.406
Sleu	0.124	0.157		0.557	0.578	0.576	0.406	0.407	0.411	0.664
Slug	0.111	0.138	0.117		0.488	0.391	0.449	0.329	0.341	0.437
ZsteC	0.126	0.134	0.154	0.113		0.279	0.354	0.407	0.329	0.488
ZsteB	0.127	0.115	0.164	0.114	0.097		0.405	0.384	0.405	0.394
ZficP	0.148	0.183	0.138	0.120	0.154	0.164		0.313	0.207	0.543
ZficS	0.142	0.158	0.113	0.058	0.126	0.137	0.126		0.215	0.424
Zgri	0.143	0.169	0.141	0.065	0.145	0.164	0.129	0.070		0.463
Zcap	0.145	0.123	0.121	0.077	0.149	0.117	0.172	0.113	0.143	

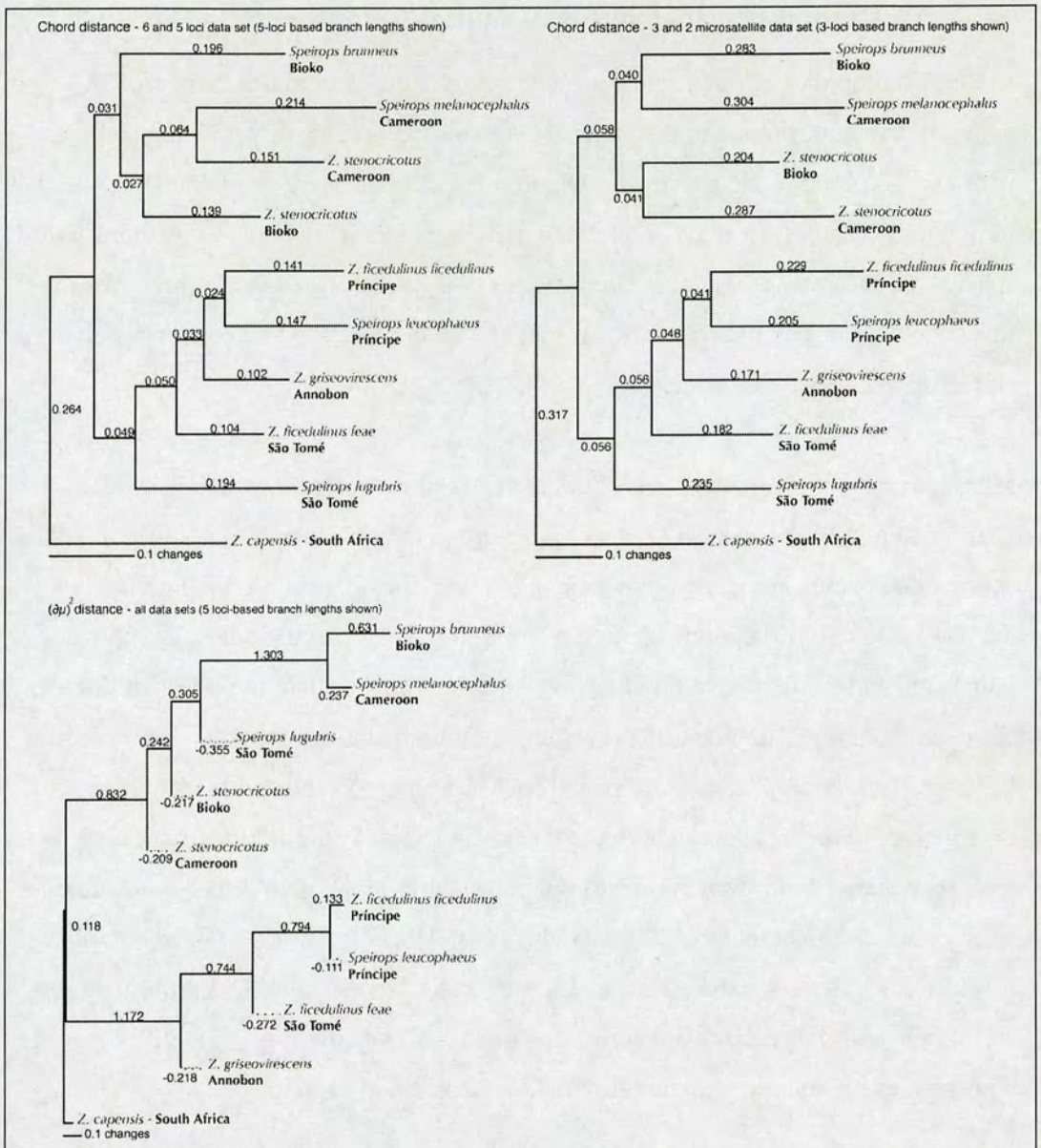


Fig. 2.8. Neighbour-joining topologies obtained for the different microsatellite loci data sets (see text) and estimated from D_C (Cavalli-Sforza & Edwards 1967) and $(\delta\mu)^2$ (Goldstein *et al.* 1995) distances.

Morphological analyses of Gulf of Guinea white-eyes

In the Zosteropidae, altough females tend to be smaller than males there are generally no siginificant sex differences (Moreau 1957). This pattern was confirmed in *Speirops lugubris*, the species for which more samples were available (105 sexed individuals). Therefore, male and female measurements were pooled together for all species. Even if any of the non-sexed species does present sexual size dimorphism, this should not affect the conclusions since the main objective was to compare the size differences between species.

Measurements of sympatric species never overlapped, except for the bill width of the Cameroon and Bioko pairs (Fig. 2.9). *Zosterops griseovirescens* (Annobón), the only species that occurs alone, had dimensions between the oceanic *Zosterops* and *Speirops*; the only exceptions were in bill length (larger than the *Speirops* spp from São Tomé and Príncipe), and in the tarsus that was smaller than the tarsus of the *Speirops* from São Tomé, but that overlapped with the tarsus from the *Speirops* from Príncipe. The *Speirops/Zosterops* pair from Mt. Cameroon showed the smallest divergences in the measurements except regarding tarsus length. In particular, *S. melanocephalus* from Cameroon was not particularly large, with dimensions often close or smaller than those of *Z. griseovirescens*. The populations of *Z. stenocricotus* from Bioko were larger than those of Cameroon for tarsus and bill measurements, and smaller for tail (despite not separating on the phylogeny: Fig. 2.5). Differences in body size (wing, mass) were larger in the São Tomé and Príncipe pairs.

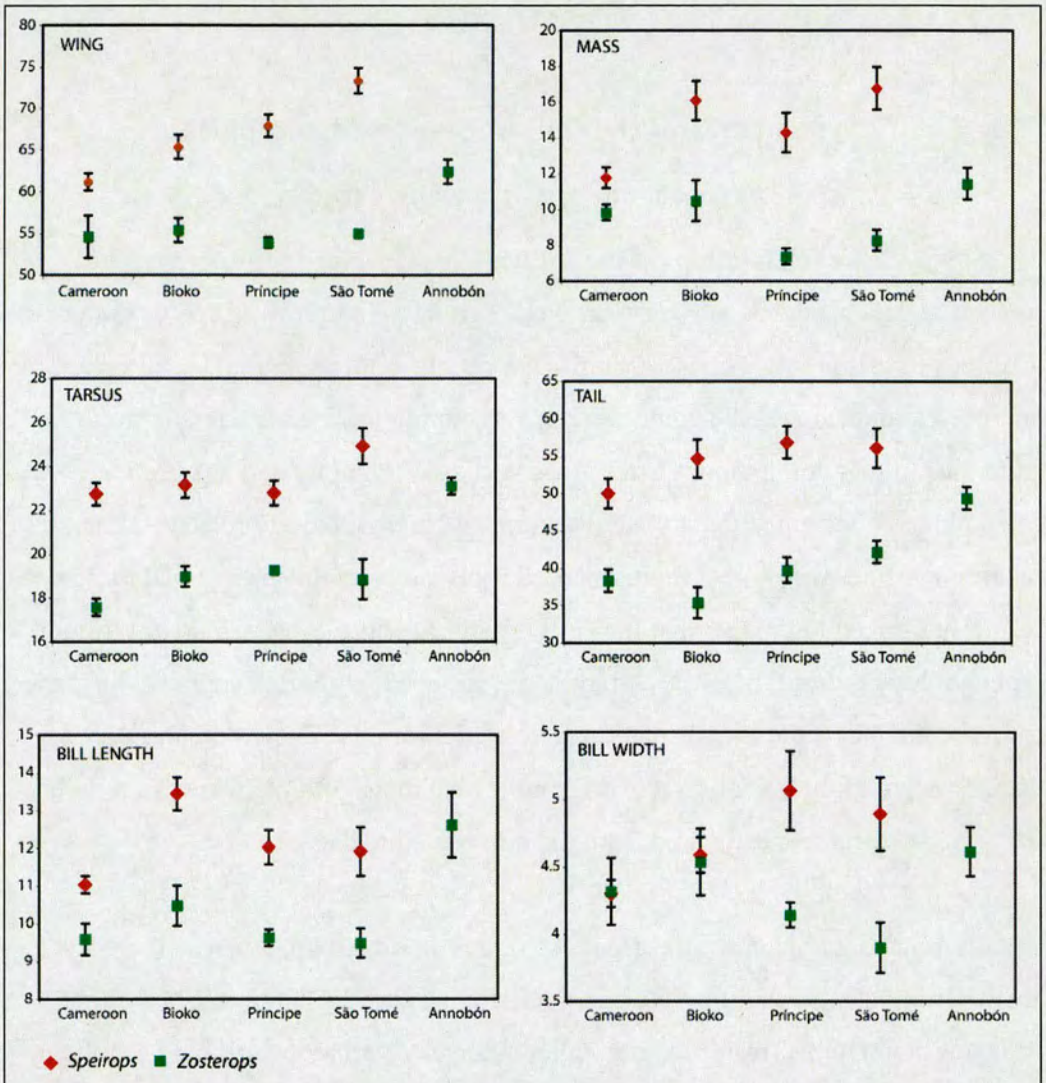


Fig. 2.9. Morphological measurements with standard deviations for the Gulf of Guinea Zosteropidae (mass in g, all other in mm).

DISCUSSION

Systematics of the Gulf of Guinea Zosteropidae

Molecular phylogenies: a different story

The phylogenetic relationships of the Gulf of Guinea white-eyes inferred with molecular data provide a very distinct picture from the currently accepted systematic affinities of this group. The most significant results, both supported by mitochondrial and nuclear data are: 1) the genus *Speirops*, grouping the four aberrant white-eyes from the Gulf, is not monophyletic; 2) the species *S. lugubris* and *S. melanocephalus*, often treated as conspecific, are valid species and are not closely related; 3) the 'aberrant' white-eyes do not represent earlier invasions to the islands. All these results are linked to the fact that the Gulf of Guinea white-eyes also do not form a monophyletic group. The mtDNA phylogeny grouped, with high support, the taxa from the three oceanic islands in a separate clade from the Cameroon and Bioko taxa. This grouping was also recovered in the neighbour-joining phenogram built from the D_C distances estimated from the microsatellite data.

Results 1 and 3 go against the current consensus on the group, whereas there has always been debate regarding the conspecificity of the São Tomé and Mt. Cameroon *Speirops* populations (result 2), although not the close relationships between them – except perhaps in Jones & Tye (2006). Using the mtDNA phylogeny, the alternative hypotheses for 1 and 2 (monophyly of *Speirops* and sister species relationship of *S. lugubris* and *S. melanocephalus*, respectively) were rejected both by Shimoidaira-Hasegawa tests and the comparison of Bayes factors.

Both Moreau (1957) and Amadon (1953) had hinted at the possibility that the four *Speirops* species could be included in *Zosterops* "without great violence to the facts" (Amadon 1953, p. 430). Admitting the impossibility of determining conclusively the relationships of these species with *Zosterops* they kept *Speirops* as a practical way of distinguishing the four aberrant Gulf of Guinea white-eyes, but in doing so they tacitly agreed a close relationship between these four species. This study shows both

that the *Speirops* species are within *Zosterops* but also that they do not constitute a monophyletic group.

Additionally, *Z. ficedulinus* was paraphyletic in all phylogenies (including microsatellite-based ones). In the maximum parsimony mtDNA phylogeny *Z. griseovirescens* (Annobón) forms a basal polytomy with *Z. f. ficedulinus* from Príncipe, whereas in Bayesian inference it forms a polytomy with *Z. f. feae* from São Tomé. Nei's uncorrected pairwise sequence divergence (Nei 1987) between the two subspecies was 3.7% (4.5% corrected), a value within the range of mitochondrial genetic distances between congeneric species documented in birds (Shields & Helm-Bychowski 1988; Johns & Avise 1998). For comparison, genetic divergence between the very distinct *S. lugubris* and *S. leucophaeus* was only 1.5% (corrected and uncorrected distances). The uncertainty regarding the relationship of these 3 taxa could reflect a true polytomy indicating a rapid radiation with retention of ancestral polymorphisms, as is suggested by the wide overlap of the divergence time estimates, or it could just be an artefact of insufficient sequence data (DeSalle *et al.* 1994; Hoelzer & Melnick 1994; Funk & Omland 2003). The clarification of this situation is important from a conservation perspective: whereas the *Z. ficedulinus* population from São Tomé is common (pers. obs. *contra* current assessments summarised in: Jones & Tye 2006) the population from Príncipe is currently very rare (pers. obs.), and appears to have been so for c. 100 years (Jones & Tye 2006, and references therein).

Mitochondrial versus microsatellite data

The results just presented were supported both by the mtDNA and the microsatellite data. The only discrepancy comes from the phenograms based on $(\delta\mu)^2$ distances which put *S. lugubris*, from São Tomé, within the 'mainland' clade. This most likely reflects a difficulty of the $(\delta\mu)^2$ distance to deal with the dataset rather than a real result, as $(\delta\mu)^2$ distances are known to perform poorly in topology reconstruction (Takezaki and Nei 1996). This is supported by the fact that several branch lengths of the $(\delta\mu)^2$ trees were negative, which could either simply indicate that they are very small or, more seriously, that the estimated distances were not additive and could not

therefore be properly handled by the NJ algorithm. For these reasons, inferences based on the D_C distances were considered more reliable and are the ones discussed.

Microsatellite- and mtDNA-based inferences differed mainly in the within-clade relationships. Differences between mitochondrial and nuclear data can be highly informative (Nicholls 2001). For example, the proximity between two species inferred by mtDNA but not nDNA can indicate past hybridization events. Because mtDNA is not linked to the nuclear genome, which contains the loci that determine individual fitness, it can penetrate the genome of a foreign population with very low levels of hybridisation (Takahata & Slatkin 1984). This explains the introgression of mtDNA across species borders, with reported cases ranging from insects (Ting *et al.* 2000; Shaw 2002) to amphibians (Weisrock *et al.* 2005), birds (Degnan 1993), and mammals (Goodman *et al.* 1999). Nevertheless, in the present study I feel it is unwise to make many conclusions out of the differences observed, as there is ample evidence to suggest that the microsatellite data were not very reliable. Only six microsatellites could be used, and the estimated distances had large standard errors. Of most concern is the fact that the information was mainly restricted to two loci. This makes the probability of congruence between the gene and species trees very low. It has been suggested that, because the effective population size of the nDNA is four times larger than that of the mtDNA, to have an equivalent amount of confidence to the mtDNA in recovering the true tree requires the use of more than 16 nuclear loci (Moore 1995). Furthermore, one of the variable loci (ASE64) was a tetra-nucleotide repeat, likely to be less suitable for phylogenetic inference than di-nucleotide repeats (Primmer & Ellegren 1998). Therefore the following detailed discussion of the evolutionary history of the Zosteropidae of the Gulf of Guinea will be based mainly on the mtDNA results. In this sense, one should keep in mind that we might be discussing only a part of the story.

Evolutionary history of the Gulf of Guinea Zosteropidae

1. Origins and colonisation routes of the Gulf of Guinea Zosteropidae

The Gulf of Guinea white-eyes consist of two clades, likely derived from different mainland ancestors. The 'oceanic' clade groups all oceanic island taxa, whereas the 'mainland' clade groups the Bioko and Mt. Cameroon ones. The mainland clade is sister to *Z. poliogaster* populations from the populations of the Aberdares and Mt. Kenya (two neighbouring mountains in East Africa). Nevertheless, as the systematics of mainland white-eyes is still very much a matter of speculation and taxon sampling on the continent was not extensive, the Cameroon-Bioko taxa could be more closely related to other lineages. For example, no samples of *Z. senegalensis* from the immediate vicinity of the range of *Z. stenocricotus* (Upper Guinea region and south of Gabon) were obtained. It should be noted that *Z. stenocricotus* is still considered by many authors a race of *Z. senegalensis*, and that those who accept the split do not deny their close affinities – although as Moreau (1957) concluded and this study amply corroborated, morphological similarity in white-eyes is not relevant from a phylogenetic perspective. The nearest mainland relative of the 'oceanic clade' could not be determined as it made a polytomy with all other African white-eye clades, excluding the species from the Mascarenes and *Z. mouroniensis* from Grande Comore – the 'ancient Indian Ocean white-eyes' *sensu* Warren *et al.* (2006). This polytomy could reflect lack of resolution due to insufficient sequence data. The dataset used for the study of the diversification of the Indian Ocean white-eyes (Warren *et al.* 2006) managed to resolve some nodes that the data set of this study could not, by the inclusion of only 212 extra bp (in a total of 1514 bp). Alternatively, the non-identification of the closest mainland relative could happen because it went extinct or because it was not sampled, as supported by the long branch that leads to the oceanic clade.

1.1. The temporal framework

The substitution rate used to infer divergence times was based on the rate obtained in the study of Indian Ocean white-eye diversification (Warren *et al.* 2006). This rate was based on a single calibration point and might therefore be incorrect.

Nevertheless, divergence dates are useful from a relative perspective. For example, the overlap of the estimates for several colonisation events suggests that they could have happened very close in time. This conclusion is independent of the true substitution rate since the overlap range is proportional to the substitution rate used (cf. Drummond & Rambaut 2003).

The substitution rate of 4.95% used in this study is faster than the typical avian mtDNA rate of 2% (García-Moreno 2004; Lovette 2004), but similar to the mtDNA rate found in *Hemignathus* honeycreepers from the island of Hawaii (Tarr & Fleischer 1993). The estimate might err on being too slow because it used the geological time estimate of the emergence of one island as the earliest possible time that island was colonised. On the other hand, if polymorphisms evolved in the ancestral population before the colonisation event, genetic distances (and therefore substitution rates) would be overestimated (Edwards 1997). The average divergence time estimates obtained for the Gulf of Guinea taxa with the 4.95% rate all fall within the Pleistocene, which began 1.8 Ma ago. Even if the true rate were twice as slow, the origin of *S. melanocephalus*, the oldest species in the group, would only be pushed back as far as the Pliocene/Pleistocene border. Therefore white-eye diversification in the Gulf of Guinea can be considered a recent radiation (Fjeldså 1994; Fjeldså & Lovett 1997). All speciation events are much more recent than the ages of the islands where they occur, with estimated sub-aerial ages varying between 5 Ma for Annobón, and 30 Ma for Príncipe (Lee *et al.* 1994).

The climatic cycles of the Pleistocene (Bartlein & Prentice 1989) led to cycles of expansion and contraction of the rainforests of central Africa (Fredoux 1994; Jahns 1996; Maley 1996; Dupont *et al.* 2000). Glacial periods were characterised by an increase in aridity and a decrease in temperature. Aridity led to the fragmentation of the lowland rainforest whereas the lower temperatures allowed the expansion of montane forests to low altitudes (up to 1000 m below current levels). Lowland rainforest biota would be isolated, whereas the connectedness between montane biota increased (Maley 1996). During the warmer and wetter interglacials, the lowland forest cover would expand again over the equatorial belt of Africa, whereas the

montane forest would contract to higher altitudes (the present situation). In the most extreme arid periods, lowland forest would be restricted to a few refuges surrounding the mountains of Upper Guinea, the Albertine Rift, and Cameroon and Gabon (Moreau 1966; Endler 1982; Mayr & O'Hara 1986; Prigogine 1987; Lovett 1993; Fjeldså *et al.* 2005). The presence of relict angiosperms on São Tomé that are absent from the mainland (paleoendemics) suggests that the oceanic islands could have acted as stable refuges before the other Cameroon mountains (Figueiredo 1994; Plana *et al.* 2004). Nevertheless, intermediary situations between full isolation and total connectedness of lowland forests must have been the predominant pattern (Maley 1996).

The cycles of contractions and expansion of the rainforest, although more intense in the last 0.9 Ma, have been felt for the last 2.8 Ma (deMenocal 1995). Even taking into consideration errors on the estimate of the substitution rate, the evolutionary history of the African white-eyes was therefore contemporary with and probably influenced by these eco-climatic cycles. Throughout these cycles, the montane regions are thought to have been very active diversification centres, where most recent radiations would have occurred (Fjeldså 1994; Fjeldså & Lovett 1997; Roy 1997; Roy *et al.* 1997; Fjeldså *et al.* 2005). Most species in this study that occur in the African forest belt are associated with montane regions, and their young age supports the montane speciation model. The white-breasted white-eye *Z. abyssinicus*, a lowland bird, forms a basal politomy in the African clade (which excludes the 'ancient Indian Ocean species'). Nevertheless, data on all lowland white-eyes are needed to determine their ancestor-descendant relationship with the montane species.

Another important consequence of the Plio-Pleistocene climatic changes was the change in sea level, which would have linked Bioko to mainland several times (Chapter 1). For example, during the last 500,000 years Bioko was part of the mainland for a total of about 250,000 years distributed in periods ranging from 30,000 to 70,000 years, and it was part of the mainland as recently as 10,000-12,000 years ago (Rohling *et al.* 1998).

1.2 The mainland clade

The inclusion of the Bioko taxa in the mainland clade is in accordance with other studies showing a close relationship between the fauna and flora of Bioko and Cameroon in general, and the montane forests of Bioko and Mt. Cameroon in particular, with many endemics exclusive to these two areas (Moreau 1966; Jones 1994; Smith *et al.* 2000; Blom & Schipper 2004; Graham *et al.* 2005). This reflects both the many periods in the past when Bioko was connected to mainland, and the proximity of the two areas. *Speirops brunneus* is currently restricted to a single mountain of Bioko despite other mountains on the island being available (Pérez del Val 1996), suggesting that it evolved a high degree of sedentariness and that the origin of the species might be linked to vicariance rather than to dispersal events. On the other hand, the 32 km that currently separate Bioko from Cameroon are likely to have allowed dispersal events between the two areas after isolation (e.g., Smith *et al.* 2005). Dispersal was suggested by the *Z. stenocricotus* sequence data: Bioko has both a very distinct haplotype (as divergent from other *Z. stenocricotus* haplotypes as the sister species *S. lugubris* and *S. leucophaeus* are divergent between them) and a haplotype very close to haplotypes from Cameroon. At the same time, four morphological measurements were significantly different between populations suggesting that differentiation is taking place, and similar haplotypes could either reflect instances of past hybridization events or retention of ancestral polymorphisms (Edwards 1997). Any further understanding of the links between Bioko and Cameroon avifaunas during periods of isolation will require a comparative phylogeographic approach, and the use of both nuclear and mitochondrial data.

The systematic position of the aberrant species from Mt. Cameroon, *S. melanocephalus*, could not be resolved with the mtDNA data but instead formed a basal polytomy, indicating that it is an old species in the clade. It was not sister to *S. brunneus* from Bioko, from which it diverged by 4.6% (uncorrected pairwise distance). This latter was sister to *Z. stenocricotus* from which it diverged by 2.2% (uncorrected pairwise distance). Stresemann (1948) had formerly suggested that *S. brunneus* was a *Zosterops*, but subsequent authors found little support to back his arguments (Amadon 1953; Moreau 1957). Microsatellite inferences inferred a sister

species relationship between the two *Speirops* from the mainland clade. This could indicate that hybridization between *S. brunneus* and *Z. stenocricotus* had led to introgression of mtDNA. The very restricted distribution of *S. brunneus* could be invoked as facilitating hybridization between this species and the much more common and widespread *Z. stenocricotus*, where their ranges overlap. The very distinct appearances of both species would suggest otherwise. Considering the low reliability of the microsatellite data used in this study, the hybridization scenario remains nothing more than a speculation until more nuclear data can be gathered.

1.3. The oceanic clade

The monophyly of the populations from the three oceanic islands supports the hypothesis that archipelago radiation, rather than multiple colonisations from mainland, played a major role in the diversification of this group: all taxa of the three oceanic islands are derived from the same founder. The importance of inter-island dispersal for the diversification of the oceanic taxa was highlighted in tests of alternative colonisation scenarios (Fig 2.3, p. 48; Table 2.8, p. 63), which consistently favoured models with dispersal between neighbouring islands, and very strongly rejected the model relying exclusively on independent colonisations from the mainland (model 7). When inter-island dispersal events were added to this model it became highly supported (model 8) – this suggests that the increase in likelihood of the model is explained by the inter-island dispersal events. São Tomé, the island in the middle of the chain, was inferred as the likely source of colonisers both for Príncipe and Annobón. The predominant SSW winds favour dispersal from São Tomé to Príncipe, but Príncipe could also have been a source of colonisers to São Tomé (discussed below), just as São Tomé was for Annobón. Dispersal between islands could be favoured because the moderate distances between them (c. 150 km) can make them visible from each other: on clear days, Príncipe can be seen from São Tomé from altitudes around 700 m (pers. obs.), and the islands' cloud cover can be seen even from sea-level (Jones & Tye 2006). Visibility of the islands would have been greater at periods of low sea level when both Príncipe and Annobón were c. 10 times their current sizes (Fig. 1.2). The fact that the white-eyes, in general, and the species from the Gulf in particular, are gregarious and move in flocks increases the

chances of successful colonisations. This characteristic is probably what explains their success in colonising islands worldwide (Moreau 1957; Clegg *et al.* 2002a).

The ancestral coloniser of the oceanic islands probably reached São Tomé and Príncipe at about the same time, an inference supported by all ancestral state reconstruction methods. This could have been achieved by island-hopping (with the entry point being either on Príncipe or on São Tomé), or by two independent colonisations from the same mainland stock. The island-hopping model was the more likely, but the double colonisation scenario could not be rejected. The very wide 95% confidence interval for the colonisation time of the Gulf of Guinea oceanic archipelago (from 0.7 to 1.9 Ma^{-1}) could indicate an old colonisation followed by a long period before radiation took place. It seems more likely to be an artefact of incorrect taxon sampling, and in particular of having missed the true MACA of the oceanic clade.

After the initial colonization of Príncipe and São Tomé, Annobón was colonised from São Tomé. These first colonisations led to the current three *Zosterops* species from the oceanic islands, whose polytomic relationship could indicate a rapid radiation (see 'Systematics'). Finally, the two oceanic island *Speirops* represented the most recent speciation events, in contrast to the generally accepted hypothesis that they represented the earliest invasions (Amadon 1953; Moreau 1957; Jones & Tye 2006). They diverge only by 1.5%, a value associated with intraspecific variation in birds (Shields & Helm-Bychowski 1988; Johns & Avise 1998).

2. Speciation mode of sympatric species

The mtDNA phylogeny did not provide a final answer to the mode of speciation of the sympatric pairs of white-eye species. Speciation in allopatry was supported for the pairs *Z. stenocricotus/S. melanocephalus* and *Z. f. ficedulinus/S. leucophaeus*, but both modes could explain the origins of the pairs *Z. stenocricotus/S. brunneus* and *Z. ficedulinus feae/S. lugubris*. Shimoidaira-Hasegawa tests could not reject the sympatric model of speciation for these two pairs, whereas Bayes factor comparisons favoured the allopatric model (albeit with not particularly high support values).

S. brunneus (Bioko) is sister to the *Z. stenocricotus* clade that has populations both in Bioko and on the mainland surrounding the Gulf of Guinea. The population of *Z. stenocricotus* is widespread over its mainland range and *S. brunneus* is restricted to altitudes above 1900 m on a single mountain on Bioko (Pico Basilé, 3010m). This pattern seems to favour a double colonisation model. One could envisage the following hypothesis: a montane ancestor would reach Bioko during a glacial period, when Bioko was part of the mainland and the montane forests of the neighbouring mountains of Pico Basilé and Mt. Cameroon had extended their lower limits. The populations from Bioko and Cameroon would separate and be restricted to high altitudes during an interglacial, and would diverge sufficiently until a second wave of colonisers again reached the island (overseas dispersal) or peninsula (during a glaciation).

Z. ficedulinus feae from São Tomé is sister to a clade that includes both a species from São Tomé and another from Príncipe. Speciation could have occurred in sympatry followed by a colonisation to Príncipe, or in allopatry if the species from São Tomé dispersed first to Príncipe where it diverged and then back-colonised São Tomé. Microsatellite data only add more 'noise' as they also support sympatric speciation for the *Z. f. ficedulinus*/*S. leucophaeus* pair.

The allopatric speciation models require more dispersal events, which should nevertheless be more likely than the conditions needed for sympatric speciation on small islands (Grant 2001; Newton 2003; Coyne & Orr 2004). Ultimately, however, no final conclusion can be made.

3. Causes of phenotypic evolution

Overall, the results of this study confirm Moreau's conclusion, after his extensive revision of the systematics of the African Zosteropidae, that morphology and plumage colour in this group convey very little phylogenetic information (Moreau 1957). The fact that *S. melanocephalus* and *S. lugubris* are not closely related despite sharing similarities not only in colour but also in song also supports Moreau's assertion that song is also not a reliable phylogenetic character in this group. A similar result was provided by a mtDNA phylogeny of the Micronesian white-eyes, where the aberrant (in colour and size) *Rukia oleagina* was found to be well within the *Zosterops* (Slikas *et al.* 2000). The other Zosteropidae genera were also defined on the basis of atypical size and colour patterns (Mees 1957, 1961, 1969; Sibley & Monroe Jr 1990; Fry 2000), and further molecular studies might conclude that rather than representing older or distinct lineages, they are all part of the genus *Zosterops*. This would probably make *Zosterops* the most speciose passerine genus (c. 100 spp), depending on the results of the molecular revision of the sunbird family Nectariniidae (c. 130 spp in tribe Nectariini: Cheke & Mann 2001).

3.1. Clarification of the Gulf of Guinea 'aberrancy' problem

As the 'aberrant' Gulf of Guinea white-eyes do not share the same history, a logical line of enquiry would be to determine if the 'aberrant' characters evolved by convergence or have different causes. In particular, one would like to set up hypotheses that could explain the origin of the aberrant characters, and why, for the entire Afrotropical region (inc. Madagascar), these are restricted to the Gulf. This is nevertheless a flawed question. The concept of 'aberrancy' is meaningless from a phylogenetic perspective, as it only says that a given species or trait is different from the norm. Because there are many ways to be different, 'aberrancy' does not necessarily reflect synapomorphies, i.e., derived characters shared by descent, the basis of phylogenetic reconstruction. In the four Gulf of Guinea *Speirops*, many of the so-called 'aberrant' characters are not shared between them. Moreau was aware of this: " The several *Speirops* differ from each other and from "normal" *Zosterops* in so many characters and to such varying degrees (...)", but " (...) in view of the

uncertain phylogeny of these Gulf of Guinea birds, no violence is done, and practical convenience is served, by keeping all four abnormal birds in *Speirops*." (Moreau 1957, p. 391 & 415). Therefore, the aberrancy of the Gulf of Guinea white-eyes was interpreted as reflecting shared history. The only aberrant characters that are shared by all *Speirops* are: increase in size and loss of the green and yellow colour. These characters are not exclusive to the Gulf of Guinea species, being found among many insular and montane *Zosterops* species (Moreau 1957). In this respect, and as discussed below, the dimensions of the Bioko and Mt. Cameroon *Speirops* are not aberrant after all (Moreau 1957).

3.2 Phenotypic evolution and the Gulf of Guinea setting

The confrontation of the patterns of morphological diversity in the Gulf with the evolutionary history of the group allows the setting up of hypotheses for the processes behind phenotypic evolution in the region. Two important factors should be considered: i) all Gulf of Guinea taxa are island taxa; ii) 8 out of the 9 Gulf of Guinea populations occur as pairs of sympatric or parapatric species. Phenotypic evolution of passerine birds on islands follows a general trend of increase in bill and body size (Grant 1968; Blondel 2000; Clegg & Owens 2002; Scott *et al.* 2003) which is also found on the Gulf of Guinea islands (Amadon 1953). Additionally, many island birds tend to lose bright colours and marked plumage patterns (Blondel 2000; Grant 2001), a pattern that is again found extensively on the Gulf of Guinea (Amadon 1953). Increase in bill size is probably associated with the adoption of a generalist strategy, which is thought to be selected for in resource-limited island environments, and made possible due to the reduction in predation and interspecific competition (Grant 1965a; Keast 1970; Lack 1976; Alatalo *et al.* 1986; Scott *et al.* 2003). Increase in body size could be a consequence or prerequisite of increase in bill size (Grant 1965a), or be a consequence of the higher levels of intraspecific competition in depauperate island environments (MacArthur *et al.* 1972; Robinson-Wolrath & Owens 2003). The loss of bright colours and marked plumage patterns has been hypothesised to reflect a reduction in sexual selection that would result from a decrease of aggressiveness in crowded environments and/or the increase in relatedness between island individuals (MacArthur *et al.* 1972; Griffith 2000).

Finally, the co-existence of closely related species is expected to affect patterns of phenotypic evolution either directly or indirectly. In the former case, resource competition between similar species will favour phenotypic divergence (character displacement) in order to reduce competition pressure, leading to specialization of each species to a different niche dimension (Brown & Wilson 1956). In the latter case, co-existing species might represent size assortment during community build-up: extinction would eliminate one or both species from similar-phenotype pairs (Case & Sidell 1983). The occurrence of sympatric white-eyes is almost exclusively restricted to islands, and it is in these situations that the most aberrant white-eyes, classified in different genera, are found (Moreau 1957; Mees 1969; Lack 1971). This indicates that a combination of the island environment and inter-specific interactions must play an important role in phenotypic evolution in this otherwise homogeneous group.

3.3. The oceanic Gulf of Guinea white-eyes

The white-eyes from the oceanic islands of the Gulf of Guinea are all derived from the same *Zosterops*-like ancestor. The three *Zosterops* taxa are closely related. The population of Annobón occurs alone, whereas the populations of São Tomé and Príncipe co-occur with a closely related *Speirops*. On Annobón, birds are exceptionally large for lowland *Zosterops*, whereas on São Tomé and Príncipe the *Zosterops* populations are the smallest and the *Speirops* are the largest in Africa (Moreau 1957; Hall & Moreau 1970). The size of *Z. griseovirescens* from Annobón is intermediate in size to the large and small populations of São Tomé and Príncipe, but its bill length is larger. This pattern, together with the common ancestry of all species, supports a role for inter- and intraspecific competition in driving phenotypic evolution of the oceanic species. On Annobón, where interspecific competition is very low (only one other forest passerine, a paradise flycatcher, is present), intraspecific competition is very high, and there is no adult predation, a generalist strategy can evolve, which is reflected in the increase in bill and body size. On São Tomé and Príncipe where two closely related populations occur in full sympatry, interspecific competition is hypothesised to have led to evolution through character displacement (rather than size assortment) of two specialized species, as is reflected

by the very large size differences. Ecological studies are needed to demonstrate that the sympatric species do indeed not overlap on their use of resources. Larger birds can feed on larger prey, whereas smaller birds have access to thinner branches (Lack 1971; Diamond 1973). The *Zosterops* of São Tomé and Príncipe tend to be confined to the highest levels of the canopies, while the *Speirops* explore all heights (pers. obs.). This could indicate a specialization of small sized birds in exploring resources within the lighter branches. The smaller size also allows different feeding behaviours, with the tiny *Zosterops* able to explore the underside of leaves by hanging upside-down from the central vein (pers. obs.). An alternative (but not necessarily exclusive) hypothesis is that the larger *Speirops* exclude *Zosterops* from the lower heights.

3.4. The mainland Gulf of Guinea white-eyes

The *Zosterops/Speirops* species pairs of the mainland clade are less divergent in size than the oceanic island pairs. In both Bioko and Mt. Cameroon the populations do not currently occur in full sympatry but in parapatry. The larger species are restricted to high altitudes and the smaller to lower altitudes, with their ranges overlapping in their distribution limits. This altitudinal segregation would restrict the opportunity for character displacement to the areas of overlap, leading to smaller size differences. In Bioko, where the zone of range overlap is apparently larger than on Mt. Cameroon (pers. obs.; Stresemann 1948; Pérez del Val 1996), the size difference is bigger. An alternative and more likely hypothesis is that each population has evolved different adaptations to the altitudinal belt they inhabit. Moreau (1957) has shown that morphometric characters of African Zosteropidae vary significantly and independently with altitude and temperature. According to the correlations found, the Bioko and Mt. Cameroon *Speirops* have the dimensions expected for the altitudinal range they inhabit. The areas of overlap are therefore the altitudinal limits of each species. In this case, interspecific interactions are more likely to lead to variable dynamics of competitive exclusion of the less adapted species at any given time rather than to character displacement (Lack 1971; Diamond 1973). Field studies of the morphological variation along the altitudinal distribution of each species could determine if divergence is higher in the areas of overlap (in this study morphological

measurements came, for each species pair, from areas of allopatry). An interesting pattern is the size differences between *Z. stenocricotus* populations of Bioko and Mt. Cameroon. The population from Bioko is significantly larger than the one from Cameroon for the bill and tarsus, but has a smaller tail. This suggests that differentiation is taking place, with the increase in bill size, and possibly tarsus length, reflecting the evolution of a generalist strategy on an insular environment and/or adaptation to altitude (on Bioko the species is restricted to heights above 800 m, whereas it can be found down to sea-level on Mt. Cameroon: Pérez del Val 1996; Fry 2000).

Conclusions

The history of the Gulf of Guinea white-eyes has two components: mainland and oceanic. Molecular data supported the 'archipelago radiation model' (Chapter 1) for explaining the diversification of the white-eyes on the three oceanic islands. This is the first case described for a bird group in the Gulf of Guinea. The two more atypical species (*Speirops* spp) represent the most recent speciation events rather than old colonisations as previously believed. Their level of genetic divergence is typical of that found at the intraspecific level in birds, indicating a very fast rate of phenotypic evolution. Reasons for this can only be speculated on, but it is likely that character displacement played a major role. Interestingly, the sympatric speciation mode cannot be excluded for the origin of one of the most aberrant species, and it is under this mode that fast phenotypic diversification is expected (Via 2001). Additionally, island environments offer two favourable conditions for sympatric speciation: high intraspecific competition and low interspecific competition (Orr & Smith 1998; Dieckmann & Doebeli 1999). In the mainland clade, the opposite pattern is observed with the typical *Zosterops* being derived from atypical *Speirops*. Body size evolution on Bioko and Mt. Cameroon is in agreement with the expected patterns of adaptation to different altitudes and temperatures. Causes for the 'atypical' colours of the four *Speirops* (from white to brown) cannot be determined. Plumage colour is a trait known for its fast rate of evolution (Price & Birch 1996; Omland & Lanyon 2000; Warren *et al.* 2005). The differences observed are all melanin-based, a trait which

has been shown to be under the control of a single locus in several wild bird populations (Mundy 2005). From a systematic perspective, the genus *Speirops* is therefore invalid: it is not monophyletic and its members are well within the *Zosterops* genus. There is no such thing as a set of shared 'aberrant' characters between the four *Speirops* but only differences from the typical *Zosterops* template, which evolved for different reasons (character displacement and fast evolution on islands, and adaptation to montane regions). The two subspecies of *Zosterops ficedulinus* are paraphyletic, and further genetic and field studies are warranted to assess their specific status.

This study provided a much-needed clarification of the evolution of the Gulf of Guinea white-eyes, but was unable to provide a complete picture because many deep nodes could not be resolved. This can only be improved by a combination of extra sequence and taxon sampling. Nuclear and mitochondrial data should be used in order to identify instances of hybridization, which could have played an important role in the evolution of this group with high dispersal abilities. Considering the great confusion surrounding the affinities of the mostly homogenous mainland populations, the ideal approach would consist of building a phylogeny with a phylogeographic sampling strategy (cf., Barraclough & Nee 2001). As the Zosteropidae are one of the most speciose passerine families, and their radiation in Africa is very recent and associated with the climatic changes of the Pleistocene, such an endeavour promises to be very rewarding.

3

ON THE ORIGIN OF THE ENIGMATIC SÃO TOMÉ GROSBEAK *Neospiza concolor*

The São Tomé grosbeak *Neospiza concolor* is a monotypic passerine genus endemic to the island of São Tomé. It is one of the least known birds in the world: after the collection of three individuals in 1888 and 1890, it disappeared for 101 years, and has only been sighted about five times since its rediscovery in 1991. The aura of mystery that surrounds this species has been enhanced by its peculiar morphology that has made its affinities uncertain, having been considered both a weaver (Ploceidae) and a finch (Fringillidae). Three individuals were mist-netted in 2003 and 2005, and blood samples taken for genetic analyses. Mitochondrial and nuclear sequences showed that *Neospiza concolor* is an Old World finch (Fringillidae: Fringillinae) sister to the Príncipe seedeater *Serinus rufobrunneus*, another Gulf of Guinea endemic with populations in São Tomé, Príncipe and Boné de Jóquei, a small islet offshore from Príncipe. The genus *Neospiza* is therefore synonymous with *Serinus*, making this species the largest *Serinus* in the world, 50 % heavier than the next largest *Serinus*. Data from 14 microsatellite loci inferred a closer relationship between *Neospiza* and the São Tomé population of *S. rufobrunneus*, than between all three allopatric *S. rufobrunneus* populations (a total of 217 *S. rufobrunneus* were genotyped). This pattern is consistent with either: i) *Neospiza* being a resource polymorphism of *S. rufobrunneus*; ii) hybridization between the two taxa; or iii) of a sympatric origin for *Neospiza concolor*. Most evidence favours the hypothesis of sympatric speciation, with allele sharing reflecting incomplete lineage sorting after the speciation event. Considering that sympatric speciation in birds is unlikely and not supported empirically, the microsatellite-inferred relationships need to be confirmed with sequence data. Both sequence and microsatellite data revealed unexpectedly high levels of genetic diversity in the three *Neospiza* individuals sampled, which could indicate that this species is more common than currently thought.

INTRODUCTION

The São Tomé grosbeak *Neospiza concolor* is one of the most enigmatic bird species in the world. The reasons for this are twofold: it is one of the rarest (or least observed) birds in the world, and it has a strange phenotype that justified the creation of a genus of its own, whose phylogenetic affinities are uncertain. In the extreme, the aura of mystery that it elicits has even led to the theory that it was a species created through hybridization by the Portuguese naturalist who discovered it, Francisco Newton – a republican (and a Freemason according to this theory) – as part of a grand subversive scheme to destabilise the monarchy (Guedes & Peiriço 1997)!

The São Tomé grosbeak was discovered in 1888 by Newton who collected a male in the forests of southeastern São Tomé, near São João dos Angolares. In 1890, Newton collected another two males in the southwestern forests, in the area of the rivers São Miguel and Quija. The type specimen is kept in the bird collections of the Natural History Museum at Tring, while the other two were lost in the 1978 fire that destroyed the National Museum of Natural History (Lisbon). Newton had already considered the species to be extremely rare (Naurois 1988) but probably never suspected that it would take 101 years for it to be rediscovered, in August 1991, in the river Xufe-xufe, close to the river São Miguel (Sargeant *et al.* 1992). Apart from two unconfirmed records in 1997 (Jones & Tye 2006), several individuals were confidently observed again (and photographed) only in January 2002, in the São Miguel area, feeding on fruits of the sugar plum *Uapaca guineensis* and the endemic *Dicranolepis thomensis* (Dallimer *et al.* 2003). In December 2002, together with the guides Pedro Leitão and Lúcio Primo, I observed one individual perching at two metres height in a flowering *Dicranolepis thomensis* in the southeast forests near the Umbumgu River, at an altitude of 300 m. On 27 January 2003, one female was captured in the same place, and on 12 February 2005, two males were captured in the same area as the Dallimer *et al.* (2003) records. Birds were ringed, measured, photographed, and blood and feather samples were obtained before releasing them. Genetic material could therefore be obtained to address the origin of this species (and to sex the three individuals as described in the methods).

Based on the shape of the head and on the massive parrot-like bill (with the upper mandible extending over the lower), the São Tomé grosbeak was originally classified as a weaver (Ploceidae) under the genus *Amblyospiza* (Bocage 1888), which currently includes a single species, the grosbeak weaver *A. albifrons*. In 1903 this placement was questioned and other traits, including the lack of the 10th primary feather, were used to consider this species to be a true finch (Fringillidae) related to canaries and seedeaters (*Serinus* sensu lato), but whose phenotypic distinctiveness warranted it to be placed in its own monotypic genus *Neospiza* (Salvadori 1903). Several authors, including Bocage, accepted this classification (Bocage 1904; Amadon 1953; Bannerman 1953). Shelley (1905) was the first to suggest that *Neospiza* was related to the Príncipe seedeater *Serinus rufobrunneus* and put them in the same genus (a change that was not adopted subsequently). The Príncipe seedeater is endemic to the islands of São Tomé and Príncipe and to Boné de Jóquei, a small islet 3 km offshore from Príncipe. Nevertheless, in a review of the family Ploceidae, Moreau could not find sufficient evidence to ascertain the phylogenetic affinities of *Neospiza*, and did not place it in either of the two putative families (Moreau 1960). Later Moreau would put *Neospiza* next to *Amblyospiza* (Moreau 1962). This was in part due to Mayr's communication to Moreau that *Neospiza* still had a 10th primary, albeit reduced (Moreau 1960). Amadon, who never directly inspected a grosbeak specimen, used detailed notes by James Chapin on the type specimen to reiterate its placement within the Fringillidae rather than the Ploceidae (Amadon 1965). Chapin described the 10th primary as being vestigial and having moved to the upper part of the wing, a pattern typical of many fringillids. Chapin further considered that *Neospiza* and *S. rufobrunneus* formed a group with the thick-billed seedeater *S. burtoni*, which has its largest-billed race on Mt. Cameroon. He proposed that *Neospiza* and *S. rufobrunneus* were derived from two independent colonisations of the same mainland stock, with *Neospiza* representing the first colonisation event. Amadon (1953) had also suggested that *S. rufobrunneus* and *S. burtoni* were the most likely closest relatives. This assessment was fully supported by Naurois who had studied *S. rufobrunneus* in the field and the two *Neospiza* specimens held at the Lisbon Museum (Naurois 1975b, 1988). Naurois added that the plumage colour of *Neospiza* was almost the same as the colour of the Príncipe and Boné populations of

S. rufobrunneus, suggesting that the colour difference between the population from São Tomé and *Neospiza* could be a result of character displacement (Naurois 1988). Having held more than 200 live *S. rufobrunneus* and three *Neospiza*, I was also struck by the similarities in plumage (Fig. 3.1), especially the fact that both species have exactly the same patch of light feathers between the chin and the breast (collar). Despite the 'danger' of using colour as a phylogenetic trait (see Chapter 2), it is difficult not to see *Neospiza* as a strange *S. rufobrunneus*.

In this study molecular data were used to determine the phylogenetic position of *Neospiza concolor*, to establish its relationship with *S. rufobrunneus*, and to identify the closest mainland relative of *S. rufobrunneus* (and *Neospiza* if both taxa were related). Mitochondrial and nuclear sequence data were used together with data from 14 microsatellite loci. Morphological measurements are presented for descriptive purposes due to the scarcity of data on *Neospiza*; measurements from a female *Neospiza concolor* are presented for the first time.

Fig. 3.1. The enigmatic São Tomé grosbeak *Neospiza concolor* (top: male; middle row: close-ups of the bill, with a female in the first photo) and the Príncipe seedeater *Serinus rufobrunneus* (bottom left), which is also present on São Tomé, have been considered to derive from two independent colonisations from the same mainland lineage. The thick-billed seedeater *S. burtoni* with a population on Mount Cameroon (pictured, bottom right) has been hypothesised as the nearest mainland relative.

METHODS

Taxon and character sampling strategy

A three-step hierarchical strategy was followed to ascertain the phylogenetic relationships of *Neospiza concolor*, and at the same time the affinities of *Serinus rufobrunneus*. First, the position of *Neospiza* in a large-scale mitochondrial phylogeny of the family Fringillidae (Yuri & Mindell 2002), the most speciose avian family, was determined. This was to confirm that it is an Old World finch (subfamily Fringillinae) in accordance with the most favoured taxonomic hypothesis. This phylogeny included among outgroup taxa representatives of two families to which *Neospiza* could also be related on morphological grounds (Ploceidae, where it was placed initially, and Passeridae). This phylogeny confirmed the close affinities of *Neospiza* with *Serinus*. Therefore, in a second step, the position of *Neospiza* within the genus *Serinus* was determined, together with the phylogenetic relationships of *Serinus rufobrunneus*, using mitochondrial and nuclear sequences. This phylogenetic inference supported a sister relationship of *Neospiza* with *Serinus rufobrunneus*, but the relationship between *Neospiza* and the three allopatric populations of *S. rufobrunneus* was polytomic. This polytomy was investigated in a third stage using a population genetics, microsatellite based, approach.

Laboratory procedures

Samples obtained for this study consisted either of frozen tissue or blood collected non-destructively from the brachial vein of mist-netted wild individuals (stored in absolute ethanol). Total genomic DNA was extracted using a 'DNeasy Tissue Extraction Kit' (Qiagen), following the protocol for DNA extraction from animal tissues. Adaptation of this protocol for blood samples stored in absolute ethanol is described on Chapter 2.

Position of *Neospiza* within the family Fringillidae

A total of 2924 bp of mitochondrial sequences were used to locate the position of *Neospiza*, *S. rufobrunneus*, and *S. burtoni* in a large-scale phylogeny of finches (Yuri & Mindell 2002). Sequences included: 874 bp of the small subunit ribosomal RNA

(12S), 675 bp of the NADH subunit 2 (ND2), the full ATP synthases 8 and 6 (ATP8, ATP6; 842 bp), six transfer RNAs (tryptophane, alanine, aspargine, cysteine, tyrosine, and lysine), and 85 bp from the 3'-end portion of cytochrome *c* oxidase subunit 3. Homologous sequences from 24 representatives of all subfamilies and 6 outgroup taxa were downloaded from GENBANK, where the new sequences obtained in this study were also deposited. Primers and PCR conditions are summarised in Appendix 3.1. DNA purification and sequencing procedures were the same as those described in Chapter 2.

Position of *Neospiza* within *Serinus* and affinities of *S. rufobrunneus*

To ascertain the relationships of *Neospiza* within the genus *Serinus*, a *Serinus* phylogeny was inferred from an 894 bp sequence of cytochrome *b* (cyt *b*). Taxon sampling included 23 *Serinus* species (19 from Africa, two from Europe, two from Asia) often represented by two individuals. Samples for the three allopatric populations of *S. rufobrunneus* (São Tomé, Príncipe, Boné) were included. The outgroup taxa were the common chaffinch *Fringilla coelebs* and the lark-like bunting *Emberiza impetuani*. The utility of the fifth intron of the β-fibrinogen gene (FIB5), an independent marker from the nucleus, was assessed in a subsample of 14 *Serinus* species. Primers and PCR conditions are summarised in Appendix 3.1. DNA purification and sequencing procedures were the same as those described in Chapter 2. Sequences have been deposited in GENBANK.

Relationships between *Neospiza* and the three *S. rufobrunneus* populations

The close relationships between *Neospiza* and *S. rufobrunneus* were further explored by taking advantage of a large microsatellite dataset developed for a population genetics study of *S. rufobrunneus* (Chapter 6). The dataset comprised 14 autosomal microsatellite loci identified for *S. rufobrunneus* and tested in 9 other *Serinus* species (Melo & Hansson 2006; Chapter 6) and on *Neospiza* (this study). Sampling included the three *Neospiza* individuals, 217 *S. rufobrunneus* (113 from São Tomé, 37 from Príncipe, 67 from Boné; Table 6.1), and between one and 13 samples of each of the nine other *Serinus* species (Appendix 6.2), totalling 250 individuals overall.

Phylogenetic analyses

Sequence data

Molecular phylogenies were estimated using model-based approaches (maximum likelihood, ML, and Bayesian inferences, BI), as implemented in PHYML version 2.4 (Guindon & Gascuel 2003) and MRBAYES 3.1 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). Thresholds for node support were set at 70 % for bootstrap values (Hillis & Bull 1993) and at 0.95 for posterior probabilities (Huelsenbeck & Ronquist 2001).

Likelihood models were estimated with MRMODELTEST 2.0 (Nylander 2004) and the best-fit models were selected using the Akaike Information Criterion (Akaike 1973). Clade supports for the ML analyses were assessed by non-parametric bootstrapping (Felsenstein 1985; 500 replicates). The compatibility of the phylogenetic signal of the different markers used for the large-scale phylogeny of the Fringillidae has been demonstrated previously (Yuri & Mindell 2002), and therefore the analysis was performed on the concatenated dataset.

For the Bayesian analyses, the most adequate partition of the sequence data was evaluated using the Bayes factors approach described in Chapter 2: biologically relevant partitions were defined (e.g., genes or codon positions) and used for phylogenetic inference. Each Bayesian search consisted of four incrementally heated Metropolis-coupled MCMC chains run for two million generations with trees sampled every 100 generations. The first 200,000 generations (2000 trees) were discarded ('burn-in' period) and the posterior probabilities were estimated for the remaining sampled generations. Two independent Bayesian runs initiated from random starting trees were performed for each data partition, and the log-likelihood values and posterior probabilities were checked to ascertain that the chains had reached stationarity. The effective sample size of the log-likelihoods was estimated in TRACER version 1.2 (Rambaut & Drummond 2003) to confirm further that the parameter space was properly explored. The harmonic means of the likelihood values of the phylogenies sampled after convergence of a MRBAYES run were used as approximations of the marginal likelihood of the respective partition model (M_1) to

be used in model comparison (M_1 vs. M_2) using Bayes factors (B_{12}). A value greater than 10 for $2\ln B_{12}$ was considered as a strong evidence against the alternative model, M_2 (Kass & Raftery 1995; Nylander *et al.* 2004). The concatenated analyses were performed freeing the different parameters (base frequencies, rate matrix, shape parameter, proportion of invariable sites) to vary between the partitions (genes and codon positions). For the best data partition, a new analysis was conducted with three independent runs of five million generations each, trees sampled every 500 generations, and the first 25% discarded as burn-in. Posterior probabilities were estimated from the 50% majority-rule consensus tree of the 22,500 sampled trees.

Microsatellite data

The program MICROSAT version 1.5 (Minch *et al.* 1995) was used to estimate pairwise chord distances (D_C; Cavalli-Sforza & Edwards 1967) between individuals, and the resulting distance matrices were used to construct neighbour-joining (NJ) phenograms in PAUP* version 4.0b10 (Swofford 2003). The chord distance has been shown to be one the best measures for topology reconstruction based on microsatellite data (Takezaki & Nei 1996). To explore whether the unbalanced dataset could have any effect on topology reconstruction, distances were estimated both for the full dataset (250 individuals, 13 taxa) and for a smaller, more balanced dataset. This consisted of 15 individuals from each of the three *S. rufobrunneus* populations, 13 *S. flaviventris* and the three *Neospiza.* The sub-sampling of *S. rufobrunneus* was done as follows: for Boné and Príncipe the first 15 individuals with no missing data were chosen (as there was only a single sampling site on Boné, and two on Príncipe with only four individuals sampled in the first site); for São Tomé, the first and last individuals with no missing data were chosen from the six more intensively sampled sites (Table 6.1), and the first individual with no missing data was chosen from another three sites.

Morphometrics of *Neospiza*

For descriptive purposes, measurements of the three *Neospiza concolor* individuals are presented and compared with measurements of *S. rufobrunneus* and *S. burtoni*, which are the three species that were thought to be related. The following

measurements were obtained: mass to the nearest 0.5 g with a Pesola spring balance; wing and tail length to the nearest 0.5 mm with a standard wing ruler; bill length, width and depth (height), and tarsus length to the nearest 0.1 mm with a digital calliper. These measurements were taken as follows: wing length (flattened), from the carpal joint to the tip of the longest primary; tail length, from the uropygial gland to tip of the central rectrix; tarsus length, from the tibiotarsus joint to the distal end of the tarsometatarsus when the foot is held to the leg; upper mandible length, from when the culmen enters the feathers of the head to tip; bill width and depth at the anterior end of nares. Because none of the species presents sexual dimorphism, sex was identified following Griffiths's *et al.* (1998) molecular protocol described in Chapter 2.

RESULTS

Molecular data characteristics

Several lines of evidence support a mitochondrial origin for the different mitochondrial fragments sampled: chromatograms were of good quality; all protein-coding sequences translated appropriately; sequence characteristics (Table 3.1) and phylogenetic reconstruction agreed with previously published results using the same markers (Arnaiz-Villena *et al.* 1999; Yuri & Mindell 2002; Ryan *et al.* 2004). Base composition variation among taxa did not show deviation from homogeneity (Fringillidae concatenated dataset: $\chi^2 = 44.07$, $P \approx 1$; cyt *b*: $\chi^2 = 9.41$, $P = 1$; FIB5: $\chi^2 = 1.99$, $P = 1$). Microsatellite characteristics are presented in Chapter 6. For the Fringillidae dataset, the best data partition, as determined by the Bayes factor approach, was by gene and by codon position (20 partitions), followed by partition by codon position (4 partitions; $2\ln B_{12} = 314$). All other partition schemes were much worse ($2\ln B_{12} > 1788$). For the *Serinus* dataset, partition of the cyt *b* by codon position was strongly favoured over non-partition (3 partitions; $2\ln B_{12} = 910$).

Table 3.1.

Sequence characteristics, based on all ingroup taxa, for the 'Fringillidae' and '*Serinus*' datasets. All markers from mitochondrial origin except FIB5.

	Sites	Var	Info	%A	%C	%G	%T	Model
FRINGILLIDAE								
12S	874	174	73	31.0	26.7	21.2	21.1	GTR+Γ+I
ND2								
All	675	337	282	30.2	35.6	11.0	23.2	GTR+Γ+I
1st	225	163	62	37.3	29.1	16.6	17.0	GTR+Γ
2nd	225	38	16	17.3	35.7	10.7	36.3	GTR+Γ+I
3rd	225	220	204	36.1	41.9	5.7	16.3	GTR+Γ
CO2								
All	85	43	32	30.0	33.9	11.5	24.6	GTR+Γ
1st	28	9	7	27.9	19.1	23.3	29.7	K2P+Γ
2nd	28	7	3	23.1	37.0	9.0	30.9	HKY+I
3rd	29	27	22	38.8	45.2	2.5	13.5	HKY+Γ
ATP8								
All	168	82	69	30.0	37.2	7.2	25.5	GTR+Γ+I
1st	56	22	14	37.4	32.0	4.6	26.0	GTR+Γ
2nd	56	14	12	15.5	41.3	11.0	32.2	GTR+I
3rd	56	46	43	37.2	38.4	6.0	18.4	GTR+Γ
ATP6								
All	684	300	237	30.0	35.7	10.1	24.2	GTR+Γ+I
1st	228	75	53	29.4	38.0	16.8	15.9	GTR+Γ+I
2nd	228	13	7	15.4	28.5	9.0	47.1	GTR+I
3rd	228	212	177	45.2	40.6	4.5	9.7	GTR+Γ+I
tRNAs	438	105	75	32.2	26.5	17.1	24.3	GTR+Γ+I
Concatenated	2924	1037	788	30.7	31.7	14.5	23.1	GTR+Γ+I
SERINUS								
Cyt *b*								
All	894	235	190	28.1	34.5	13.2	24.1	GTR+Γ+I
1st	298	33	22	24.1	29.0	24.8	22.0	SYM+Γ+I
2nd	298	8	4	19.5	27.0	13.2	40.3	HKY+I
3rd	298	194	164	40.7	48.1	1.5	9.8	GTR+Γ
FIB5 (Nuclear)	571	26	11	30.1	17.5	20.6	31.9	GTR+Γ+I

Var: variable sites - Info: parsimony informative sites. Models of sequence evolution: GTR, general time reversible (Rodríguez *et al.* 1990); HKY, Hasegawa *et al.* (1985); K2P, Kimura 2 parameter (Kimura 1980); SYM, symmetrical model (Zharkikh 1994); Γ = variation of nucleotide substitution rate among sites is described by a gamma distribution; I = a proportion of the sites are considered to be invariant.

Phylogenetic inference

Position of *Neospiza* within the Fringillidae family

The Fringillidae phylogeny inferred from a subset of the samples used by Yuri & Mindell (2002) recovered the same relationships as in that study (Fig. 3.2). A higher number of nodes were supported by BI than by ML, a likely consequence of the ability of MRBAYES to allow different models of sequence evolution for each data partition, in contrast to ML implementation methods that can only consider a single model for the entire dataset. Both inference methods supported the monophyly of *Serinus* and recovered a sister species relationship between *Neospiza* and *S. rufobrunneus* (Fig. 6.2). This sister species relationship reflects close affinities but might not be real since only three *Serinus* species were sequenced. These results placed therefore *Neospiza* within the Old World finches, and confirmed its close affinities to *S. rufobrunneus*.

Position of *Neospiza* within *Serinus* and affinities of *S. rufobrunneus*

The *Serinus* phylogeny inferred with the cytochrome *b* sequences recovered a sister species relationship between *S. rufobrunneus* and *Neospiza*, with strong support from both inference methods (Fig. 3.3). The *S. rufobrunneus* haplotype from São Tomé nevertheless formed a polytomy within the clade. This is probably a result of its large divergence from the haplotypes from Boné and Príncipe (uncorrected: 2%, corrected: 2.3%), which is close to the 2.4 % (uncorrected) or 2.9% (corrected) net genetic distance between *Neospiza* haplotypes and all of the *S. rufobrunneus* haplotypes. The Príncipe and Boné haplotypes diverge by only 0.2% (corrected and uncorrected). The closest species to these taxa was the streaky seedeater *S. striolatus* – a montane species presently occurring in East Africa – rather than the previously hypothesised *S. burtoni* with a population on nearby Mount Cameroon. The position of *S. burtoni* could not be ascertained. The FIB5 sequences, with only 11 informative sites, inferred a much less resolved tree with supported nodes concordant with those recovered by the cyt *b* analyses (not shown). Combining the two sequences did not improve resolution over the cyt *b* tree (not shown). The genetic diversity of *Neospiza* as inferred from only three haplotypes was large, with haplotypes differing by 53, 60, and 62 nucleotides out of a total of 894.

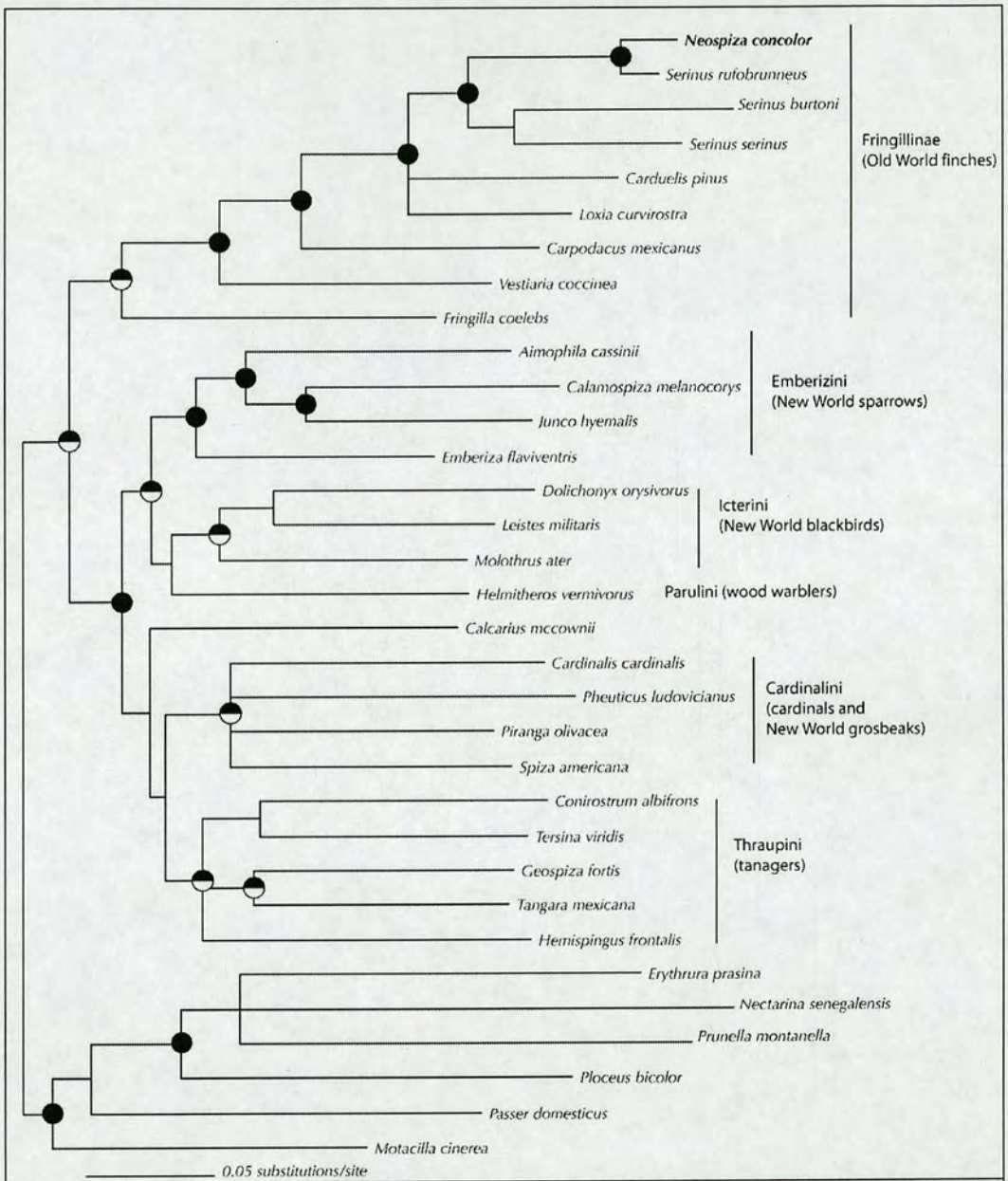


Fig. 3.2. Position of *Neospiza concolor*, *Serinus rufobrunneus*, and *S. burtoni* in the Fringillidae phylogeny of Yuri & Mindell (2002). 50% majority-rule consensus tree of 22,500 trees (sampled after stationarity) obtained by Bayesian inference from 2924 bp of mtDNA. A ML search gave the same topology. Nodal support is indicated by circles with Bayesian posterior probabilities in the top hemisphere (black: > 0.95), and ML bootstrap in the lower hemisphere (black: > 70%).

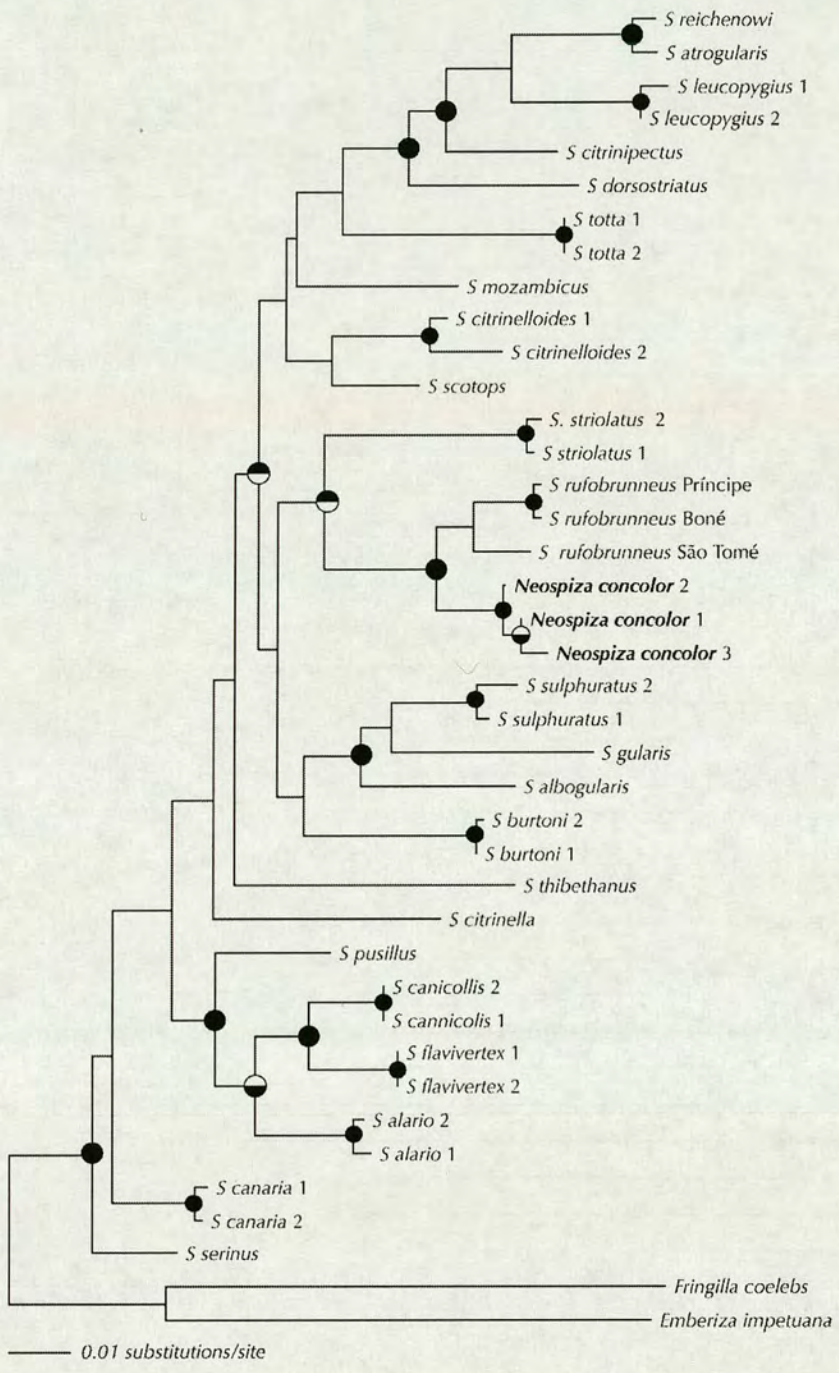


Fig. 3.3. Position of *Neospiza concolor* within the genus *Serinus*. 50% majority-rule consensus tree of 22,500 trees (sampled after stationarity) obtained by Bayesian inference from 894 bp of cyt *b*. ML inference gave a topology that only differed by recovering with support a sister relationship between *S. Serinus* and *S. canaria*. Nodal support is indicated by circles with Bayesian posterior probabilities in the top hemisphere (black: > 0.95), and ML bootstrap in the lower hemisphere (black: > 70%).

Relationships between *Neospiza* and *S. rufobrunneus* allopatric populations

Data from 14 microsatellite loci put the three *Neospiza* individuals well within the São Tomé population of *S. rufobrunneus*. This relationship is recovered by both datasets: 250 individuals-13 taxa (Fig. 3.4), and 61 individuals-5 taxa (Fig. 3.5). From the 52 alleles sampled in the three *Neospiza*, 13 were unique and, of the remaining 39, 70% were shared with São Tomé, 44% with Príncipe and 30% with Boné (percentages do not add up to 100 because the three *S. rufobrunneus* also share alleles). If this result reflects their true phylogenetic history, the three allopatric populations of *S. rufobrunneus* are paraphyletic. The validity of this result from a phylogenetic perspective is supported by the fact that the microsatellite data recovered exactly the same *Serinus* clades as those recovered by sequence data from this study and previously published phylogenies (Arnaiz-Villena *et al.* 1999; Ryan *et al.* 2004), with no detectable effect of the unequal sample sizes between taxa. For example, despite only a single sample from *S. sulphuratus* being available it was put together with *S. flaviventris*, and these taxa were in turn grouped with *S. albogularis* – as with the mitochondrial sequences. Mitochondrial data grouped *S. burtoni* with this group but without node support (this study; Ryan *et al.* 2004). Microsatellites recovered the same relationship lending support to this grouping. Microsatellite data also separated the three allopatric populations of *S. rufobrunneus*, a result confirmed by population genetic analyses discussed in detail in Chapter 6.

The genetic diversity sampled in the three *Neospiza* individuals was surprisingly high for a species thought to be extremely rare (Table 3.2). Out of 40 sampled genotypes, only 10 were homozygous. For most loci, observed and expected heterozygosity of *Neospiza* was greater than the values for the Príncipe and Boné populations of *S. rufobrunneus* and similar to the values of the São Tomé population. Locus Lswμ14 was monomorphic, but this is a locus for which evidence of null alleles is strong (Chapter 6). Diversity estimates from only three individuals are likely to be far from the population values but, if anything, diversity estimators such as allelic richness are expected to increase with increasing sample sizes.

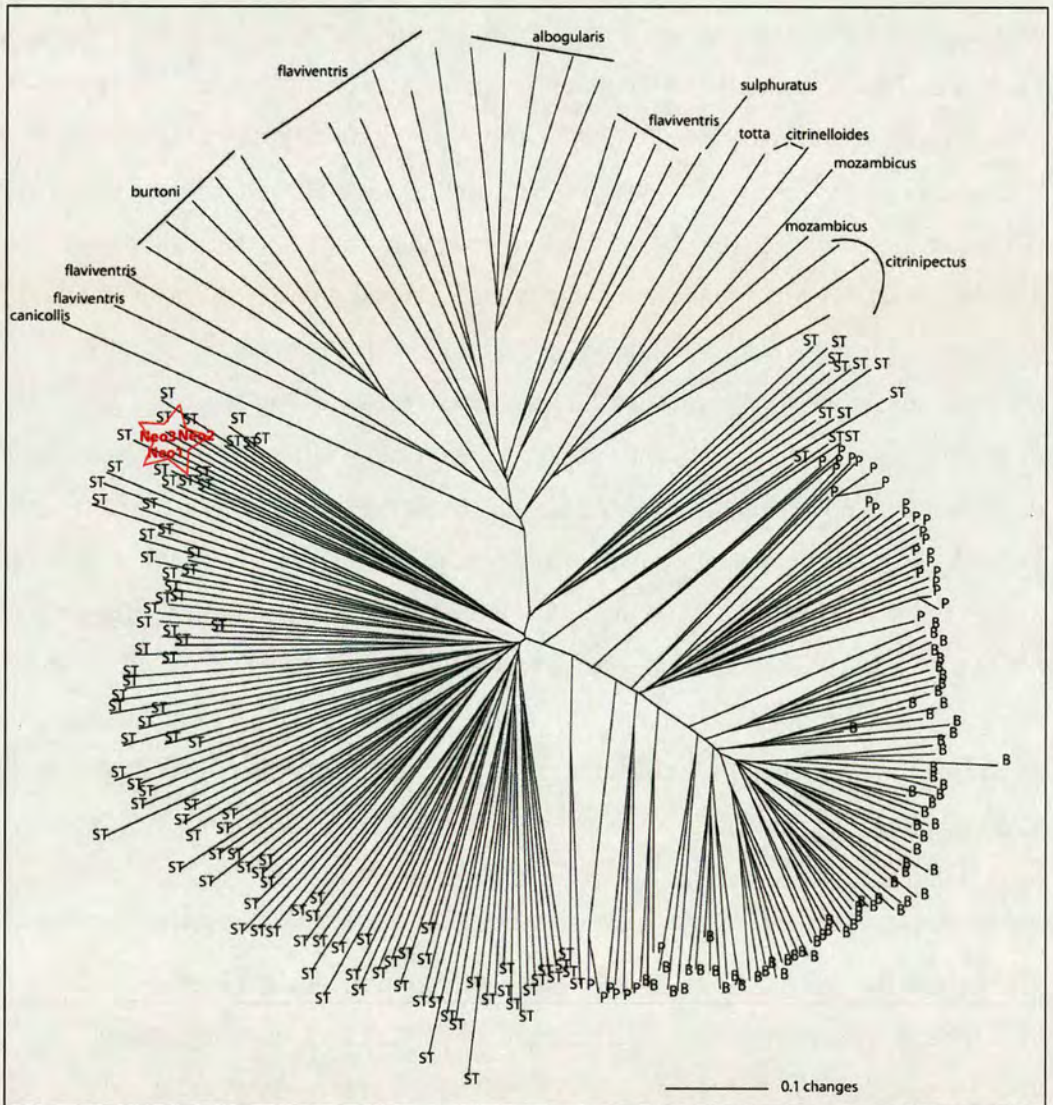


Fig. 3.4. Neighbour-joining tree built from pairwise chord distances (Cavalli-Sforza & Edwards 1967) between 250 individuals from 13 taxa estimated from 14 microsatellite loci. The position of the three *Neospiza* (Neo) individuals is highlighted by a star. Individuals from each of the three allopatric populations of *S. rufobrunneus* are labelled with the initials of each island: ST, São Tomé; P, Príncipe; B, Boné.

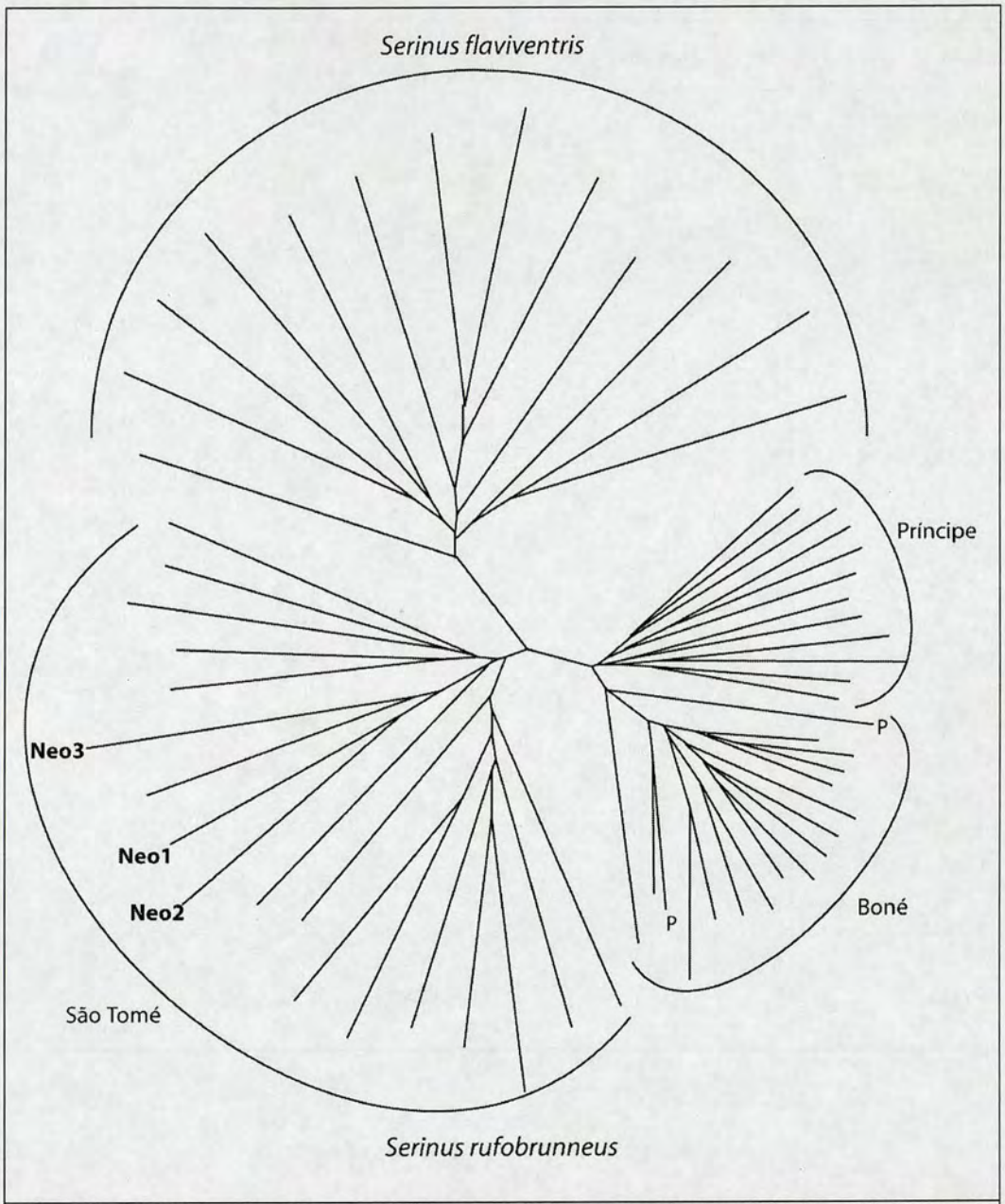


Fig. 3.5. Neighbour-joining tree built from pairwise chord distances (Cavalli-Sforza & Edwards 1967) estimated from 14 microsatellite loci between 61 individuals from 5 taxa: three allopatric populations of *Serinus rufobrunneus*, *S. flaviventris* and *Neospiza concolor* (Neo). Two individuals from Príncipe labelled with 'P' fell into the Boné group (islet 3 km offshore Príncipe).

Table 3.2. Comparison of microsatellite diversity between *Neospiza concolor* (only 3 sampled individuals) and the three *Serinus rufobrunneus* populations (São Tomé, n = 117; Príncipe, n = 37; Boné, n = 67). Complete details on the microsatellites can be found on Chapter 6.

Locus		*Neospiza*	*Serinus rufobrunneus*		
			São Tomé	Príncipe	Boné
Ase42	H_o	0.67	0.81	0.41	0.14
	H_E	0.44	0.74	0.33	0.18
Ase43	H_o	1.00	0.69	0.81	0.58
	H_E	0.63	0.83	0.74	0.70
Ase48	H_o	0.33	0.79	0.73	0.70
	H_E	0.61	0.88	0.69	0.73
Cuμ04	H_o	0.67	0.72	0.11	0.00
	H_E	0.61	0.71	0.16	0.00
Cuμ28	H_o	0.67	0.86	0.69	0.83
	H_E	0.61	0.88	0.80	0.72
Gf08	H_o	0.33	0.17	0.03	0.00
	H_E	0.28	0.23	0.03	0.00
LOX1	H_o	1.00	0.96	0.92	0.27
	H_E	0.75	0.95	0.89	0.44
LOX3	H_o	1.00	0.99	0.84	0.84
	H_E	0.83	0.98	0.93	0.82
LOX7	H_o	1.00	0.98	0.97	0.66
	H_E	0.83	0.98	0.97	0.92
LOX8	H_o	1.00	0.92	0.58	0.42
	H_E	0.72	0.97	0.91	0.62
Lswμ14	H_o	0.00	0.33	0.22	0.16
	H_E	0.00	0.52	0.31	0.26
Lswμ18	H_o	1.00	0.51	0.73	0.02
	H_E	0.78	0.69	0.85	0.02
Pdoμ4	H_o	1.00	0.77	0.92	0.76
	H_E	0.83	0.82	0.95	0.82
WBSW7	H_o	0.67	0.57	0.00	0.00
	H_E	0.44	0.55	0.00	0.00

H_O: observed heterozygosity; H_E: expected heterozygosity.

Morphometrics of *Neospiza*

Neospiza concolor is a very large '*Serinus*' with a massive bill even when compared with *S. burtoni*, the largest *Serinus* (after '*Neospiza*') with a very large bill. *Neospiza* males have a body weight that is double that of the males of its sister species, *S. rufobrunneus*, and a bill depth (height) that is almost double (93% larger). Wing length, bill length and width are 33, 44 and 48% greater in *Neospiza*, with the smallest difference being found for tarsus length (18%). The sexual size dimorphism

of *Neospiza* was very large for four out of the seven traits, with the two males having a body weight 24% greater than the female, a tarsus length 14% greater, and a bill length and depth 12 and 17% greater. The single measured female had unusually short tarsi compared with its body weight: they were only 3% greater than the tarsi of *S. rufobrunneus* females, whereas its body weight was 59% greater.

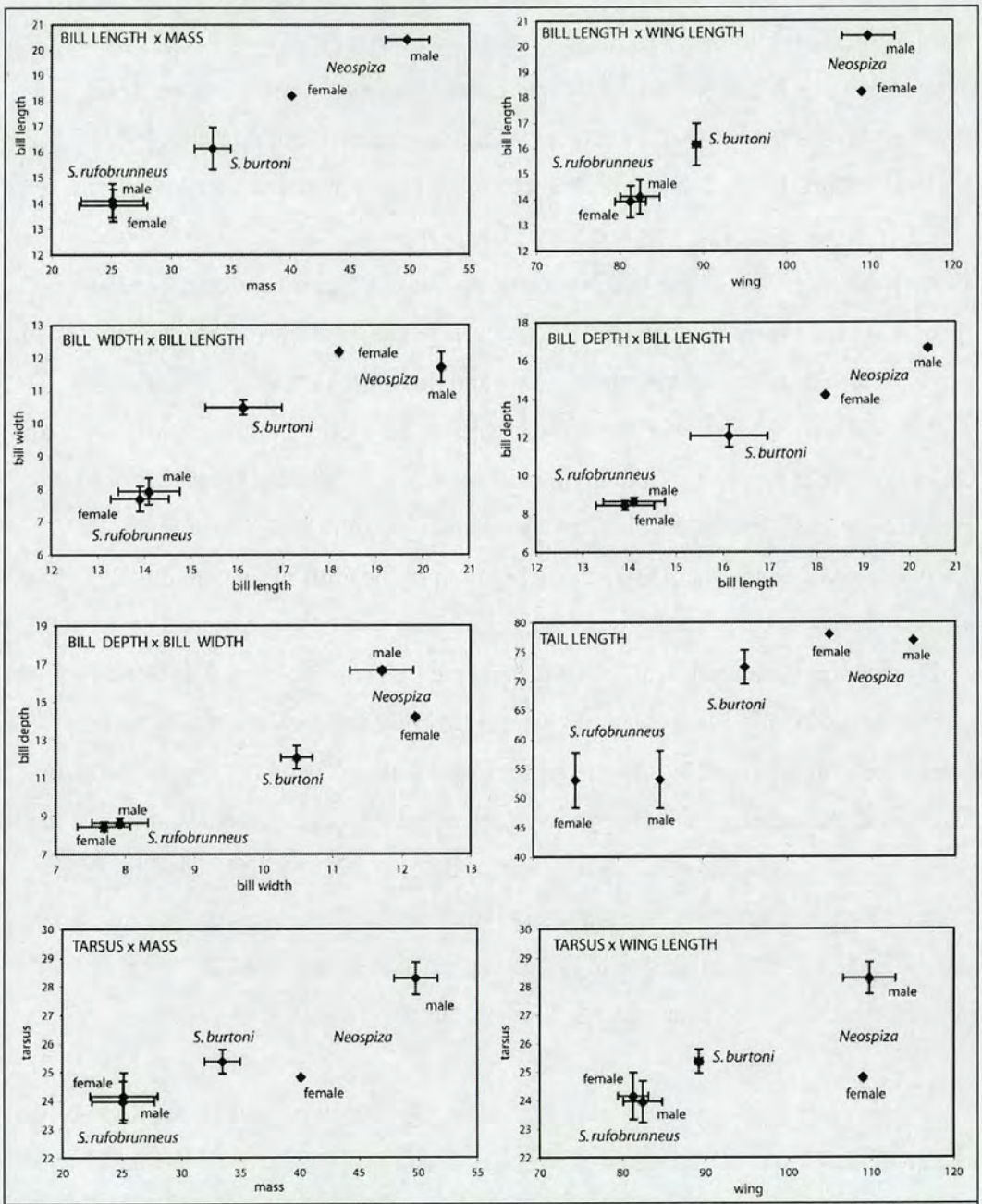


Fig. 3.6. Morphometrics of *Neospiza concolor* in comparison with *Serinus rufobrunneus* and *S. burtoni*. Mass in g, all other in mm. Sample sizes - *Neospiza*: 2 males, 1 female; *S. burtoni*: 5 individuals; *S. rufobrunneus*: 130 males, 72 females.

DISCUSSION

Neospiza concolor: a giant *Serinus*

This study confirmed that the São Tomé grosbeak *Neospiza concolor* is an Old World finch (Fringillidae: Frigillinae) rather than a weaver (Ploceidae), related to the genus *Serinus* – the most widely accepted taxonomic position (Salvadori 1903; Bocage 1904; Shelley 1905; Amadon 1953; Bannerman 1953; Amadon 1965; Naurois 1975b, 1988). Furthermore, it confirmed the hypothesis that *Neospiza* is closely related to the Príncipe seedeater *S. rufobrunneus* (Shelley 1905; Amadon 1965; Naurois 1988). These two co-occurring Gulf of Guinea endemics are sister species, which places *Neospiza* well within the genus *Serinus*, making the monotypic genus *Neospiza* synonymous with *Serinus* and therefore invalid. Hence, the São Tomé grosbeak is a giant *Serinus* – the largest in the world – joining other São Tomé endemics like the giant sunbird *Dreptes thomensis* and the giant weaver *Ploceus grandis* – the two largest species of their families – and the São Tomé speirops *Speirops lugubris* and the São Tomé population of the Gulf of Guinea thrush *Turdus olivaceofuscus* – the largest representatives of their families in Africa. The trend for body size increase among Gulf of Guinea island birds has been well documented (Amadon 1953), and fits the general pattern for body size evolution of birds on islands – small birds get bigger and larger ones smaller (Clegg & Owens 2002). In São Tomé, the endemic dwarf olive ibis *Bostrychia bocagei*, the smallest ibis in the world, provides good support for body size reduction of large birds. Despite following a common trend, the impressive magnitude of the *Neospiza* gigantism should be highlighted: males weight twice as much as males of its sister species *S. rufobrunneus* and 50% more than *S. burtoni*, the next largest *Serinus*.

Contrary to earlier suggestions (Amadon 1953, 1965; Naurois 1988), the thick-billed seedeater *S. burtoni* occurring on Mount Cameroon is not the closest living relative of the Gulf of Guinea island finches. Instead, the Gulf of Guinea taxa are sister to the streaky seedeater *S. striolatus*, a highland species currently restricted to East Africa, from Eritrea to Eastern Zaire and Northern Tanzania (Fry & Keith 2004). This

disjunct distribution pattern is common for many African montane forest birds (Mayr & O'Hara 1986; Prigogine 1987). It is explained by the environmental effects of the Plio-Pleistocene glacial cycles (Maley 1996). During glaciations the montane forest cover would extend to lower altitudes, increasing the connection between mountain ranges; during interglacials (the current situation) the montane forest retracted to the higher altitudes that are found only in western and eastern mountain ranges, thus creating disjunct distributions among montane species. A similar pattern has been reported for several endemic Gulf of Guinea groups: ferns and angiosperms (Figueiredo 1994, 1998; Plana *et al.* 2004), snails (Gascoigne 1994), frogs (Measey *et al.* 2007), and skinks (Jesus *et al.* 2005b). In this case, it may be that the common ancestor of *S. striolatus*, *S. rufobrunneus* and *Neospiza* went extinct from mainland West Africa. Alternatively, *S. burtoni* could represent this lineage on the mainland: its phylogenetic position could not be ascertained, but it was placed near *S. striolatus* both here and in Ryan *et al.*'s (2004) study.

The *Serinus* phylogeny included 19 of the 33 or 37 recognised African species (Clement *et al.* 1993; Fry & Keith 2004; respectively) and therefore there is the possibility that the true nearest relative was not sampled. Nevertheless, all species with populations on the mainland surrounding the Gulf of Guinea have been sampled, with the exception of *S. capistratus* considered to be part of a superspecies that includes *S. citrinelloides* (Fry *et al.* 2004), which was sampled. Additionally, if the superspecies groupings of Fry *et al.* (2004) are treated as single taxa, sampling in this study covered 19 out of 28 taxa, with most excluded taxa being very restricted East African endemics. The relationship of the Gulf of Guinea taxa with an East African taxon therefore constitutes a solid result, but a complete phylogeny of African *Serinus* using additional sequence data will be needed to construct an accurate phylogenetic hypothesis for the diversification of the group, as the present one lacks resolution in many nodes.

Neospiza concolor: speciation in sympatry?

The relationship of *Neospiza* with the three allopatric populations of *S. rufobrunneus* could not be determined from the mitochondrial data because the haplotype from the São Tomé seedeater population could not be placed within the clade. This polytomy is unlikely to be a 'hard polytomy' reflecting simultaneous speciation events (Hoelzer & Melnick 1994) because the branch lengths separating the São Tomé lineage from the others are large. On the contrary, the distinctiveness of the São Tomé haplotype might be in the origin of the polytomy: it is almost equidistant from the other *S. rufobrunneus* haplotypes and the *Neospiza* haplotypes, with its divergence from the Boné and Príncipe haplotypes (2.3%) being at the lower bound of species divergence in passerine birds (Johns & Avise 1998; Lovette & Bermingham 1999; Ryan & Bloomer 1999; Markland & Lovette 2005). A phylogeny where the *S. rufobrunneus* haplotypes are constrained to be monophyletic gained more support than a phylogeny where the São Tomé haplotype is constrained to be sister to the *Neospiza* haplotypes, as determined by the Bayes factor approach ($2\ln B_{12} = 8$). Under the guidelines of Kass & Raftery (1995), this Bayes factor value provides already strong evidence favouring the monophyletic model, but relative to most model comparisons this is a small value. For example, partition of the Fringillidae dataset by gene and codon position rather than partition by gene produces the same phylogeny, but the more complex partition receives a much higher support ($2\ln B_{12} = 1779$). Therefore, and as Bayes factor interpretation is subjective, a value of 8 does not provide sufficient confidence to resolve the polytomy. The clarification of the relationships between the *S. rufobrunneus* populations and *Neospiza* using mitochondrial data will require both more sequence data and a phylogeographic sampling. The former is needed to resolve the polytomy; the latter to increase haplotype sampling and estimate population differentiation: rare dispersal events between islands are likely (Chapter 6), and could lead to fast introgression of mitochondrial haplotypes across islands (Barton & Jones 1983; Takahata & Slatkin 1984). In other words, the distinct São Tomé haplotype might not be exclusive to São Tomé.

The microsatellite data placed the three *Neospiza* individuals well within the São Tomé seedeater population, making *S. rufobrunneus* a paraphyletic species. The validity of this result from a phylogenetic perspective was supported by the fact that these data recovered exactly the same *Serinus* relationships as the mitochondrial data (see Results). Paraphyly of *S. rufobrunneus* is consistent with a recent speciation event that led to the origin of *Neospiza* from the São Tomé population of *S. rufobrunneus* (Avise & Ball 1990; Funk & Omland 2003). Because *Neospiza* is currently restricted to São Tomé this pattern is in agreement with speciation having happened in sympatry. Alternatively, this pattern could indicate either that *Neospiza* and *S. rufobrunneus* are resource polymorphisms of the same species, or that *Neospiza* hybridises or has hybridised with *S. rufobrunneus*. These possibilities are discussed further below. Nevertheless, this pattern does not seem to conform to the hypothesis that *Neospiza* and *S. rufobrunneus* represent independent colonisations, at different times, of the same lineage (Amadon 1965; Naurois 1988). Incomplete lineage sorting could explain the sharing of alleles between the two species derived from different colonisations, but sharing should be higher between the most recent arrival and an allopatric population (Príncipe, Boné or a mainland species), which was not the case.

Neospiza as a case of sympatric speciation

The discrepancy between the paraphyletic pattern recovered by microsatellite data and the large genetic differentiation between *Neospiza* and the three *S. rufobrunneus* haplotypes inferred from mitochondrial sequences is consistent with the different times for lineage sorting of each marker. Lineage sorting, the process by which diverging populations fix different alleles by stochastic loss of common alleles and appearance of new ones by mutation, is faster for mitochondrial DNA due to its smaller effective population size, generally a quarter of that of nuclear loci (Funk & Omland 2003). The high variability of microsatellite loci in relation to mitochondrial DNA, and in particular the very high variability of many of the loci used (Chapter 6), will further increase the time needed for lineage sorting of the microsatellite dataset (Hedrick 1999). Despite the large sharing of alleles between *Neospiza* and the São Tomé population of *S. rufobrunneus*, 13 out of the 52 sampled *Neospiza* alleles were

unique to *Neospiza* indicating that differentiation of the nuclear genome is also taking place.

Sympatric speciation requires the evolution of a polymorphism, of assortative mating, and of a linkage between the two in the face of gene flow (Kirkpatrick & Ravigné 2002). These conditions are thought to be very difficult to meet in birds (Newton 2003), a group characterised by high levels of gene flow, which are expected to break any associations between polymorphic and assortative mating traits. In birds, the only case known so far of sympatric speciation has been described in the brood parasite indigobirds *Vidua* spp. (Sorenson *et al.* 2003). This is nevertheless a case with unique characteristics: assortative mating evolves before any polymorphism and by cultural, rather than genetic, evolution. Male songs and female preferences (for song and nests) result from imprinting on the hosts, and therefore reproductive isolation can evolve simultaneously with the colonization of a new host (Payne *et al.* 2000). The best proof for sympatric speciation – under the orthodox route – would be to find sister species confined to small areas such as oceanic islands (Coyne & Orr 2004). No such case has yet been found in birds (Coyne & Price 2000).

The data presented here are the strongest to date supporting sympatric speciation in birds. Coyne & Orr (2004, p. 142) listed four criteria that, if met, provide strong support for sympatric speciation: i) the species must be largely or completely sympatric; ii) the species must have substantial reproductive isolation; iii) the sympatric taxa must be sister groups, i.e., genetic similarity should not derive from hybridisation; iv) the existence of a past allopatric phase is very unlikely. The first criterion is met since *Neospiza*, endemic to São Tomé, is fully sympatric with its sister species, *S. rufobrunneus*. Nevertheless, only the microsatellite data supports a sister relationship between *Neospiza* and the *S. rufobrunneus* population with which it co-occurs. The second criterion is almost certainly met (see below '*Neospiza* as a resource polymorphism), and the third is likely to be met (see below 'Hybridisation between *Neospiza* and *S. rufobrunneus*'). The relationships inferred from the microsatellite loci support the last criterion unambiguously: the sister species to

Neospiza has indeed allopatric populations but *Neospiza* is most closely related to the sympatric population.

Overall, support for the sympatric origin of *Neospiza* is very much dependent on the confirmation of a sister relationship with the *S. rufobrunneus* population from São Tomé. This confirmation must come from phylogenies inferred from sequence data and using multiple samples per population. If the more parsimonious pattern of reciprocal monophyly of *Neospiza* and *S. rufobrunneus* is recovered, speciation in sympatry will neither be supported nor rejected, as *Neospiza* could be derived from any of the *S. rufobrunneus* populations. Evidence from the microsatellite data would nevertheless still favour speciation in sympatry.

Neospiza as a resource polymorphism

Resource-based polymorphisms can evolve when interspecific competition is low and a diversity of resources is available (Smith & Skúlason 1996). The first condition is typical of island environments, whereas the second is likely to occur in the highly diverse rainforest habitat of São Tomé, and in particular in the initial stages of colonisation from the mainland. *Neospiza* could then represent a morph of *S. rufobrunneus* that, with its bill typical of species that crush or crack their food, may specialise on a very hard food resource (Naurois 1988). Bill polymorphisms are found in several bird species, with the Galápagos' medium ground finch *Geospiza fortis* (Grant & Grant 1995, 2002c) and the African black-bellied seed-cracker *Pyrenestes ostrinus* (Smith 1993, 1997) providing two classic examples. The evolution of polymorphisms in bill size and shape can constitute an important path for speciation in birds, as was the case for the radiation of Darwin's finches (Lack 1947; Grant 1986; Schluter 2000).

In the present case, however, molecular and morphological data suggest that *Neospiza* and *S. rufobrunneus* are distinct species. The mitochondrial genetic divergence between *Neospiza* and *S. rufobrunneus* (2.9%) indicates a long period of isolation. If the most widely reported rate of cyt *b* evolution in birds of 2% per million years (García-Moreno 2004; Lovette 2004) applies in this case, *Neospiza*

could have diverged from *S. rufobrunneus* from 780,000 to 1.5 million years ago (Bayesian estimate using a strict molecular clock, as implemented in BEAST version 1.3: Drummond & Rambaut 2003). Additional sampling of *S. rufobrunneus* (and *Neospiza* if possible) is required to confirm this level of genetic divergence. In any case, the presence of unique microsatellite alleles in *Neospiza* supports also a long time of independent evolution. The very large morphological differences also make it unlikely that *Neospiza* and *S. rufobrunneus* represent polymorphisms of the same species. For comparison, the difference in weight between *Neospiza* and *S. rufobrunneus* is twice the difference in weight between what is probably the most disparate polymorphism in birds: the "mega-billed" and "small-billed" forms of *P. ostrinus* (Smith & Girman 2000). Nevertheless, the fact that *Neospiza* and *S. rufobrunneus* do not constitute polymorphisms does not invalidate the possibility that the evolution of resource polymorphisms within the ancestral species triggered the speciation event – if it occurred in sympatry.

Hybridisation between *Neospiza* and *S. rufobrunneus*

Sharing of alleles between two distinct species could result from hybridisation rather than incomplete lineage sorting (Funk & Omland 2003). Introgression of genetic material across species borders is expected to be faster and easier to detect for mitochondrial DNA, due to its smaller effective population size and lack of recombination (Barton & Jones 1983; Takahata & Slatkin 1984). Here, nuclear alleles were shared but there was no sharing of mitochondrial haplotypes, a pattern consistent with incomplete lineage sorting rather than hybridisation. Unfortunately, however, the mitochondrial sampling of both *Neospiza* and *S. rufobrunneus* was insufficient in this study, so the hybridization hypothesis cannot at present be rejected due to a lack of data. Adequate sampling of mitochondrial DNA to detect hybridisation is particularly important in birds, a group where females are the heterogametic sex. According to Haldane's rule (Haldane 1922; Orr 1997), any inviability arising from hybridisation will be more strongly expressed in females, reducing therefore the levels of introgression of mitochondrial DNA (Tegelstrom & Gelter 1990; Brumfield *et al.* 2001; Bensch *et al.* 2002; Crochet *et al.* 2003).

Genetic variability of *Neospiza*

Neospiza concolor is considered one of the rarest birds in the world and accordingly is classified under the highest IUCN threat category, 'Critical' (BirdLife International 2000). The genetic data obtained in this study from only three individuals detected much more genetic variation than expected for a rare species, although high genetic diversity has been occasionally documented in rare species, and especially in plants (Gitzendanner & Soltis 2000; Ellis *et al.* 2003). It may indicate that the population only recently crashed from a much larger number, or that *Neospiza* is more common than currently believed. As *Neospiza* has been considered extremely rare since its discovery almost 120 years ago one would expect that stochastic loss of genetic diversity would be more noticeable by now. Therefore, there is some hope that *Neospiza* is faring better than predicted from the rarity of observations. It could be a bird adapted to the canopy level, only coming down in specific situations – perhaps when one of its understorey food plants such as *Dicranolepis thomensis* is fruiting. The clarification of the São Tomé grosbeak enigma must include the identification of its food resources and, in particular, such foods as require such a massive bill to deal with them. This could then help focus searches on the best places for *Neospiza* in order to estimate its population size and investigate its interactions with *S. rufobrunneus*.

4

PHYLOGENETIC AFFINITIES OF THE GULF OF GUINEA THRUSH *Turdus olivaceofuscus* (Turdinae), AN ENDEMIC SPECIES FROM SÃO TOMÉ AND PRÍNCIPE ISLANDS, AND ASSESSMENT OF THE TAXONOMIC STATUS OF THE TWO POPULATIONS

The Gulf of Guinea thrush *Turdus olivaceofuscus* is endemic to the islands of São Tomé and Príncipe, with the populations in each island described as distinct subspecies (*T. o. olivaceofuscus* and *T. o. xanthorhynchus*, respectively). Morphologically, this species differs strikingly from mainland thrushes, which has made the inference of phylogenetic relationships difficult. A phylogeny of 23 *Turdus* species based on 1424 nucleotides of mitochondrial DNA revealed that *T. olivaceofuscus* is sister to the African thrush *T. pelios*, which occurs on the nearby mainland and Bioko Island. The previously hypothesised close relationship with the Comoros thrush *T. bewsheri* was not supported. The relationship between both Gulf of Guinea populations could not be established: their monophyly was not recovered, with all methods inferring a paraphyletic relationship, albeit with no support. This pattern, if confirmed, is consistent with each island having been colonised independently from the mainland. The genetic divergence between the two *T. olivaceofuscus* subspecies was high (uncorrected: 6.8%; corrected: 10.4%) – well within the range documented for congeneric bird species, and greater than the divergence found between several recognised *Turdus* species pairs. The clear phenotypic differentiation between the subspecies was concordant with the genetic divergence. These results support the splitting of *T. olivaceofuscus* in two species: the São Tomé thrush *T. olivaceofuscus* and the Príncipe thrush *T. xanthorhynchus*. This splitting has important implications for conservation. The Príncipe thrush is very rare (discovered in 1899, rediscovered in 1928, and rediscovered again only in 1997) but is currently classified under a minor IUCN threat category ('near-threatened') because it is put together with the common São Tomé population. As a separate species it will qualify immediately to the category 'Vulnerable', and could even qualify for the highest threat level, 'Critically Endangered', pending estimates on population size and demographic trends.

INTRODUCTION

The high levels of bird endemism in the Gulf of Guinea are concentrated on São Tomé and Príncipe, the two largest oceanic islands, where up to 28 endemics have been recognised (Stattersfield *et al.* 1998; Jones & Tye 2006). Of these, only five species are shared between both islands, three of which classified as distinct subspecies in each island: the Príncipe white-eye *Zosterops ficedulinus*, the Príncipe seedeater *Serinus rufobrunneus*, and the Gulf of Guinea thrush *Turdus olivaceofuscus*. This pattern suggests that isolation between islands has been an important factor in promoting bird diversification in this system. A similar pattern has been described for the angiosperms where out of 176 Gulf of Guinea endemics only 16 are shared between islands (Figueiredo 1994). Genetic data have corroborated the importance of isolation by suggesting that the two subspecies of the white-eye might actually constitute two cryptic species (Chapter 2), and confirming that genetic and phenotypic structuring between seedeater populations are high (Chapter 6). The two subspecies of *T. olivaceofuscus* (São Tomé: *olivaceofuscus*, Príncipe: *xanthorhynchus*) conform to this pattern, as they are by far the most clearly phenotypically differentiated subspecies pair.

The taxonomic position of *T. olivaceofuscus* has been difficult to ascertain because, phenotypically, it differs strikingly from its most likely nearest mainland relatives (species from the *T. olivaceus* complex). The most obvious differences are in size, with the thrush from São Tomé being one of the largest in the genus *Turdus*, and in colour, with *T. olivaceofuscus* having barred underparts contrasting with the uniform or streaked underparts of its putative mainland relatives. Therefore, it has been proposed that the Gulf of Guinea thrush and the similar Comoro thrush *T. bewsheri* – endemic to three Indian Ocean islands of the Comoros archipelago – are not closely related to any of the extant African species, but that both are representatives of an old mainland lineage now extinct (Keith & Urban 1992). Alternatively the distinctiveness of *T. olivaceofuscus* and *T. bewsheri* could just be a consequence of the trend for the fast evolution of 'aberrant' traits on islands as was shown to be the

case with the Gulf of Guinea endemics *Speirops* (Chapter 2) and *Neospiza* (Chapter 3). A recent molecular phylogeny of the *Turdus olivaceus* species complex did not recover a close and basal relationship between *T. bewsheri* and *T. olivaceofuscus* as would be expected if both were derived from the same ancient mainland lineage (Bowie *et al.* 2005). Nevertheless, the position of *T. olivaceofuscus* could not be resolved: it tended to group with the African thrush *T. pelios*, the thrush that occurs on the Gulf of Guinea mainland, but such a relationship was never supported. Therefore, in the case of *T. olivaceofuscus* the hypothesis that it derives from an extinct mainland lineage remains valid.

Contrarily to the uncertainty regarding the affinities of *T. olivaceofuscus*, the classification of the São Tomé and Príncipe populations as two subspecies of the same species has been widely accepted at least since 1924 (Sclater 1924). This consensus is interesting considering the large morphological and colour differences between the two populations: the *olivaceofuscus* race from São Tomé has been considered more similar to the nominate race of *T. bewsheri* from the Indian Ocean Anjouan Island, on the other side of the African continent, than to the neighbouring *xanthorhynchus* race from Príncipe (Naurois 1984). The population from Príncipe was initially described as a distinct species: it is smaller than *olivaceofuscus*, the bill is shorter and it is yellow rather than black, the legs are pale instead of dark, and the breast and belly bars are larger and darker (Salvadori 1901). Ascertaining the specific status of closely related allopatric populations is necessarily a relatively arbitrary issue (Daugherty *et al.* 1990): in the time of Salvadori any small difference found in an island population was used as justification for the description of a new species, whereas the evolutionary synthesis led to the adoption by taxonomists of the concept of geographic variation within species (Mayr 1942). The genus *Turdus* is a particularly good example of a group with high levels of geographic variants, which had made the inference of phylogenetic relationships difficult (Keith & Urban 1992; Clement & Hathway 2000; Bowie *et al.* 2005). Even if the 'polytypic species' position better reflects the current understanding of the evolutionary process its adoption often led to the downgrading of populations from species to subspecific level based mostly on the taxonomist's 'feeling' rather than on evidence (Mayr &

Diamond 2001). Deciding if a given allopatric population should be given subspecific or specific status might seem a relatively trivial issue, especially in light of the widely accepted view that there is a continuum between intraspecific evolutionary change and speciation (Mayden 1997; de Queiroz 2005). In the present case this issue is not trivial for two main reasons, one fundamental and the other applied, as described next.

The attribution of subspecific status to the two island populations assumes that they are sister populations, with the corollary that one island was colonised from the other. Considering the levels of phenotypic differentiation, an equally likely hypothesis in the Gulf of Guinea situation is that each population derives from independent colonisations from the mainland. Even if the colonisers were from the same mainland lineage – the most parsimonious scenario – they could have colonised the islands at different points in time. In this case the two populations would be paraphyletic: the population derived from the more recent colonisation event would be sister to the mainland species that provided the colonisers, and the older island population would be basal to this group. Current taxonomic classification could therefore misrepresent the biogeography of the Gulf of Guinea *Turdus*.

From an applied perspective, a correct taxonomic assignment is fundamental in biodiversity conservation (Avise 1989; Crowe *et al.* 1994; Collar 1997; Dillon & Fjeldså 2005) as the species is the most commonly used unit for the delineation of conservation priorities (e.g., Stattersfield *et al.* 1998; Myers *et al.* 2000; Olson *et al.* 2000; Stattersfield & Capper 2000; Fishpool & Evans 2001; Inskipp & Gillet 2005; IUCN 2006) and is certainly the unit best understood by legislators. A correct taxonomic assignment is of particular relevance in the present case. Whereas the thrush population on São Tomé is common and present in most habitats with tree cover – from primary forest to gardens in the capital city, the population on Príncipe has been rare since its discovery. A single specimen was collected in 1899 by the Italian naturalist Leonardo Fea who already considered it to be uncommon (Naurois 1984). Subsequent explorations could not find it, until 1928 when José Correia, the Portuguese collector working for the American Museum of Natural History, obtained

four specimens. After Correia, the species was only rediscovered in 1997 (Sasha Lima *in litt*). Other records followed that showed its range to be confined to the southern third of the island where mature forest still occurs, from the highest peak (Pico do Príncipe: 948 m) down to the southern coast. This is the area of highest conservation value, recently proclaimed a Natural Park albeit no measures on the ground have yet been implemented. Despite the trickle of records coming from ornithologists visiting the most inaccessible areas of the island, the pattern that emerges is one of rarity. Thrushes on Príncipe are very tame – much more so than the birds from São Tomé – and may forage on the ground a few metres away from humans (pers. obs.; Baillie & Gascoigne 1999) but nevertheless sightings remain scarce. Because the thrush was always observed in areas of mature forest, its range might be less than 30 km^2. On average, I sighted a single thrush in every 4-5 days' stay in the southern forests, suggesting that the population might be smaller than 250 individuals, the IUCN threshold to consider a species 'Endangered', the second highest threat level (IUCN 2001). Jonathan Baillie considered it to be frequent to common at altitudes above 250 m, but still considered the population to be certainly below 1,000 birds (Baillie & Gascoigne 1999), a figure that would qualify it as 'Vulnerable' (IUCN 2001). Evidently, a proper census of this population is required. What stands out is the importance of gathering solid evidence regarding the taxonomic status of this rare population. As a subspecies of *T. olivaceofuscus*, the rarity of this population is hidden by the abundance of the São Tomé population, which is responsible for placing *T. olivaceofuscus* in the IUCN category 'near-threatened' (BirdLife International 2000). Taxonomic assessments in such situations are therefore anything but trivial, with incorrect assignments in similar situations being responsible for extinctions that otherwise could possibly have been averted (e.g., Daugherty *et al.* 1990; Hazevoet 1996).

In December 2004, one thrush was captured on Príncipe for the first time since Correia's visit in 1928. This individual was mist-netted at 130 m on the southern forests close to Ribeira Porco (1°32'N 7°22'E). It was ringed, measured and a blood sample obtained from the brachial vain before being released again. This sample offers therefore an opportunity to assess from a molecular perspective the

relationship between the São Tomé and Príncipe thrush populations and could potentially allow resolution of the relationship of these insular populations with the mainland species. In this study I used mitochondrial sequence data to address these two questions. The levels of genetic differentiation of the São Tomé and Príncipe populations were interpreted together with patterns of morphological and colour differentiation in order to determine the likelihood that the two populations constitute different species.

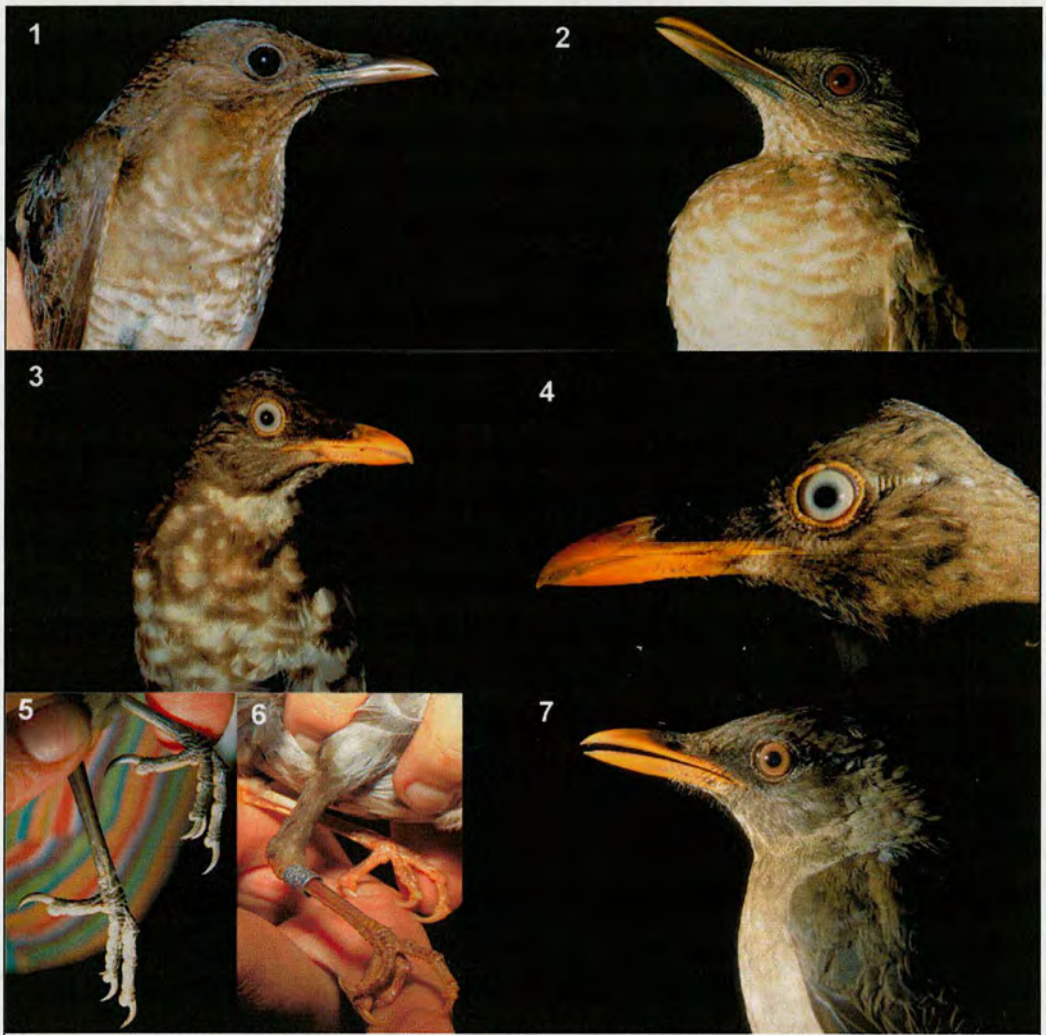


Fig. 4.1. The Gulf of Guinea thrush *T. olivaceofuscus* (1–4) and the African thrush *T. pelios* from nearby mainland (7). *T. olivaceofuscus* is currently classified into two very distinct subspecies: *olivaceofuscus* from São Tomé (1,2) and *xanthorhynchus* from Príncipe (3,4). The two subspecies differ in morphometry (see introduction and results) and in coloration. The bill of the nominate race is dark brown with variable extensions of light brown–dark yellow on its edges and tip, the iris is dark brown, and the feet are dark grey (5). This contrasts with the bright yellow bill, the almost white iris and the pale legs (6) of *xanthorhynchus*, which also has a yellow eye-ring. Additionally the patterns of the underparts differ, with *xanthorhynchus* presenting larger and more well defined bars (close to spots) and a clear white throat, and *olivaceofuscus* showing thin bars and a 'dirtier' throat. The barred underparts of the insular populations are not found on the mainland species most likely to be the closest relatives, which have either uniform (7) or streaked underparts. The 'eye-stripe' of naked skin present in this group is clearly visible on photo 4.

METHODS

Taxon and character sampling

In total 27 individuals from 22 taxa were analysed in this study (Appendix 4.1.). The entire *Turdus olivaceus* complex was rooted on the song thrush *T. philomelos* a broadly distributed Old World species (introduced to Australia and New Zealand), which forms one of the basal branches within *Turdus* (G. Voelker, pers. comm.). Other outgroup taxa included were: *T. hortulorum*, *T. ruficollis*, *T. iliacus*, *T. grayi*, *T. plebejus* and *T. infuscatus*. Phylogenetic inference was based on the complete sequences of the mitochondrial genes NADH dehydrogenase subunit 2 (ND2; 1041 bp) and subunit 3 (ND3; 351 bp), and 32 bp of flanking transfer RNA (glycine and arginine).

Laboratory procedures

DNA was extracted from frozen tissues or blood using either a Puregene DNA isolation kit (Gentra Systems), or using a phenol/chloroform/isoamylalcohol extraction protocol (Bowie *et al.* 2003). Samples conserved in ethanol were washed as described on Chapter 2 and for both methods the digestion step was extended to at least 12 hours. Primers and PCR conditions for the amplification and sequencing of the ND2 and ND3 (with flanking tRNAs) are presented in Appendix 4.2. PCR products were electrophoresed on 1.5% low-melting point agarose gels (FMC Bioproducts), stained with ethidium bromide and visualized under UV light. Amplicons of the appropriate length were cut out of the gel and purified using GELase™ (Epicentre Technologies). The purified products were cycle-sequenced using Big Dye terminator chemistry (Applied Biosystems, Inc.) precipitated with 3M ammonium acetate or 100% isopropanol, rinsed in ethanol, dried and re-suspended in formamide-EDTA solution, and run on an ABI 3100 automated DNA sequencer. Sequences were obtained from both strands of DNA for each individual, and some individuals were sequenced several times if any base ambiguity was encountered. Internal primers were used to facilitate double-stranded sequencing of the entire ND2 gene (Appendix 4.2). All sequences were checked using the program Sequencher 3.0

(Gene Codes Corp) and aligned to the chicken (*Gallus gallus*) mtDNA sequence (Desjardins & Morais 1990) to test for the presence of any insertions or deletions, as well as to check that no stop-codons were present. Sequences have been deposited in GENBANK.

Phylogenetic analyses

The compatibility of the phylogenetic signal of the different markers used for the large-scale phylogeny of the Turdinae was previously demonstrated (Bowie *et al.* 2005) and therefore all analyses were performed on the concatenated dataset. Parsimony analyses (MP) were conducted in PAUP* version 4.0b10 (Swofford 2003) using a heuristic search, implementing stepwise addition with 1000 random addition replicates, and tree bisection-reconnection (TBR) branch-swapping. Because transitions often accumulate at higher rates than transversions, MP analyses were conducted under empirical 3rd position transition:transversion weighting schemes (10x ts:tv), with 1st and 2nd positions being weighted the same as 3rd position transversions. Clade support was estimated using the non-parametric bootstrap (Felsenstein 1985) with 500 replicates, with each replicate containing one random addition replicate. Maximum likelihood analyses (ML) were conducted in PAUP* using a full heuristic search with 50 random addition replicates and TBR branch-swapping. Clade support was estimated using 100 bootstrap replicates. MODELTEST (Posada & Crandall 1998) was used to select the most appropriate model of nucleotide evolution.

MRBAYES version 3.1 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003) was used to conduct a Bayesian approach to phylogenetic inference (BI). The most adequate partition of the sequence data was evaluated using the Bayes factors approach described in Chapter 2: different biologically relevant partitions were defined (e.g., genes or codon positions) and used for phylogenetic inference. The best model of sequence evolution for each partition was estimated with MRMODELTEST version 2.0 (Nylander 2004). For each partition model, two independent runs of two million generations, each with four Metropolis-coupled MCMC chains (one cold and three heated chains) were run simultaneously to

optimize efforts to find peaks in tree-space. The first 200,000 generations (2000 trees) were discarded ('burn-in' period) and the posterior probabilities were estimated for the remaining sampled generations. Log-likelihood values and posterior probabilities were checked to ascertain that the chains had reached stationarity. The effective sample size of the log-likelihoods was estimated in TRACER version 1.2 (Rambaut & Drummond 2003) to confirm further that the parameter space was properly explored. The harmonic means of the likelihood values of the phylogenies sampled after convergence of a MRBAYES run were used as approximations of the marginal likelihood of the respective partition model (M_1) to be used in model comparison (M_1 vs. M_2) using Bayes factors (B_{12}). A value greater than 10 for $2\ln B_{12}$ was considered as a strong evidence against the alternative model, M_2 (Kass & Raftery 1995; Nylander *et al.* 2004). The analyses of the combined dataset were performed freeing the different parameters (base frequencies, rate matrix, shape parameter, proportion of invariable sites) to vary between the partitions (genes and codon positions). For the best data partition, a new analysis was conducted with three independent runs of five million generations each, trees sampled every 500 generations, and the first 25% discarded as burn-in. Posterior probabilities were estimated from the 50% majority-rule consensus tree of the 22,500 sampled trees.

Morphological and colour data

The level of phenotypic differentiation between the São Tomé and Príncipe thrush populations was assessed from two distinct datasets, one from individuals captured in the field and the other from museum specimens (Appendix 4.3). All individuals from the field dataset were captured with mist-nets, ringed, measured and a blood sample was obtained. Their sex was determined with a genetic-based protocol (Griffiths *et al.* 1998), as described in Chapter 2. Sex of the individuals from the museum dataset had been determined by gonad inspection at the time of collection. Measurements from the field dataset were obtained by myself, and measurements from the museum data set were obtained by Nigel Collar. The colour of the bill, leg and iris was recorded for all individuals.

The field dataset consisted of a single male captured on Príncipe and 24 individuals captured on São Tomé (13 males, 11 females). The following measurements were obtained: mass to the nearest 0.5 g with a Pesola spring balance; wing and tail length to the nearest 0.5 mm with a standard wing ruler; bill length, width and depth (height), tarsus length, longest toe length, and head + bill length to the nearest 0.1 mm with a digital calliper. These measurements were taken as follows: wing length (flattened), from the carpal joint to the tip of the longest primary; tail length, from the uropygial gland to tip of the central rectrix; tarsus length, from the tibiotarsus joint to the distal end of the tarsometatarsus when the foot is held to the leg; upper mandible length, from when the culmen enters the feathers of the head to tip; bill width and depth at the anterior end of nares; longest toe length (the middle anterior toe) from the base of the first phalange to the end of the third where it links with the claw; head + bill length from the back of the skull to the tip of the bill.

The museum dataset consisted of 10 individuals from São Tomé (six males, five females) and three from Príncipe (two males, one female). Tarsus, wing and bill length were measured to the nearest mm. Tarsus was measured as above. Wing length was measured without flattening. Bill length was measured from the base of the skull (equivalent to field measure) and from the tip of the nares.

The small datasets were inappropriate for significance testing. Therefore results are presented only for descriptive purposes, from both a univariate and multivariate perspective. A principal component analysis (PCA) was used for the multivariate description of morphological differences. PCA was conducted in PRIMER version 5 (Primer-E Ltd) on standardised data with the components extracted from a covariance matrix.

RESULTS

Sequence characteristics

The mitochondrial origin of the different sequences has been demonstrated previously (Bowie *et al.* 2003; Bowie *et al.* 2005): sequences were obtained from

chromatograms of good quality; all protein-coding sequences aligned without gaps or insertions and translated appropriately; sequence characteristics (Table 4.1) were similar to previously published results using the same markers (e.g., Voelker 1999; Markland & Lovette 2005; Moyle *et al.* in press), including the subfamily Turdinae, which includes the genus *Turdus* (Klicka *et al.* 2005). Variation in base composition among taxa did not show deviation from homogeneity (ND2: $\chi^2 = 12.69$, P = 1; ND3: $\chi^2 = 9.88$, P = 1; concatenated: $\chi^2 = 18.53$, P = 1). The best data partition, as determined by the Bayes factor approach, was by gene and by codon position (7 partitions), followed by partition by codon position (4 partitions; $2\ln B_{12} = 49$). Partition by gene or no partition were significantly worse ($2\ln B_{12} > 383$).

Table 4.1. Sequence characteristics based on all taxa, except the root *T. philomelos*

	Sites	Var	Info	%A	%C	%G	%T	Model
ND2								
All	1041	360	249	28.3	35.2	13.0	23.5	GTR+Γ
1st	347	74	46	32.8	30.0	18.6	18.6	GTR+Γ+I
2nd	347	34	18	16.0	33.5	11.3	39.2	GTR+Γ+I
3rd	347	252	185	36.1	42.2	9.0	12.7	GTR+Γ
ND3								
All	351	127	77	25.5	35.6	14.1	24.8	GTR+Γ+I
1st	117	23	12	21.4	35.5	22.8	20.3	SYM + Γ
2nd	117	13	8	17.5	27.1	12.2	43.2	HKY+Γ+I
3rd	117	91	57	37.6	44.2	7.4	10.7	HKY+Γ
tRNAs	32	3	3	46.7	16.5	22.3	14.5	HKY+I
Concatenated	1424	490	329	27.9	35.0	13.4	23.6	GTR+Γ

Var: variable sites - Info: parsimony informative sites. Models of sequence evolution: GTR, general time reversible (Rodríguez *et al.* 1990); SYM, symmetrical model (Zharkikh 1994); HKY, Hasegawa *et al.* (1985); Γ = variation of nucleotide substitution rate among sites is described by a gamma distribution; I = a proportion of the sites are considered to be invariant.

Phylogenetic inference

All inference methods recovered a sister species relationship between *T. olivaceofuscus* and *T. pelios* (Fig. 4.2). This was well supported by Bayesian inference (BI), was just supported by maximum likelihood (ML; bootstrap = 70%), but was not supported by maximum parsimony (MP; bootstrap = 54%). Additionally, all methods recovered a paraphyletic relationship between the two Gulf of Guinea populations, with the population from São Tomé (*T. o. olivaceofuscus*) occupying a basal position, and the population from Príncipe (*T. o. xanthorhynchus*) as sister to *T. pelios*. Nevertheless, this received no support and therefore the two populations make a polytomy with *T. pelios*. This polytomy was investigated with BI by performing two additional searches: one where the Gulf of Guinea populations were constrained to be monophyletic; another where the paraphyletic pattern was enforced. The Bayes factor in favour of paraphyly was of $2\ln B_{12} = 3.5$. According to general guidelines (Kass & Raftery 1995), this value constitutes positive evidence for the paraphyletic pattern. Nevertheless, Bayes factors favouring one model over another tend to be much larger – for example, in this study the Bayes factor favouring a partition by gene and codon position over a partition by gene only was of 766, even though both partitions recovered the same topology (see also: Irestedt *et al.* 2004). Therefore, based on empirical experience, a value of 3.5 does not constitute unambiguous evidence in favour of one model, and the relationship between both subspecies should still be considered uncertain. This study confirmed that *T. bewsheri* from the Comoros is sister to *T. libonyanus* and is not closely related to *T. olivaceofuscus* – there is as much divergence between them as between them and some European and Central American species (Table. 4.2).

Genetic divergence between the Gulf of Guinea populations was large (uncorrected: 6.8%; corrected: 10.4%), and greater than between several *Turdus* species (Table 4.2). Some of these species have only recently been recognised (*helleri, roehli, smithi*: Bowie *et al.* 2003, 2005), but divergence between the Gulf of Guinea populations was also greater than between the very distinct *T. bewsheri* and *T. libonyanus* (uncorrected: 5.1%; corrected: 7.2%).

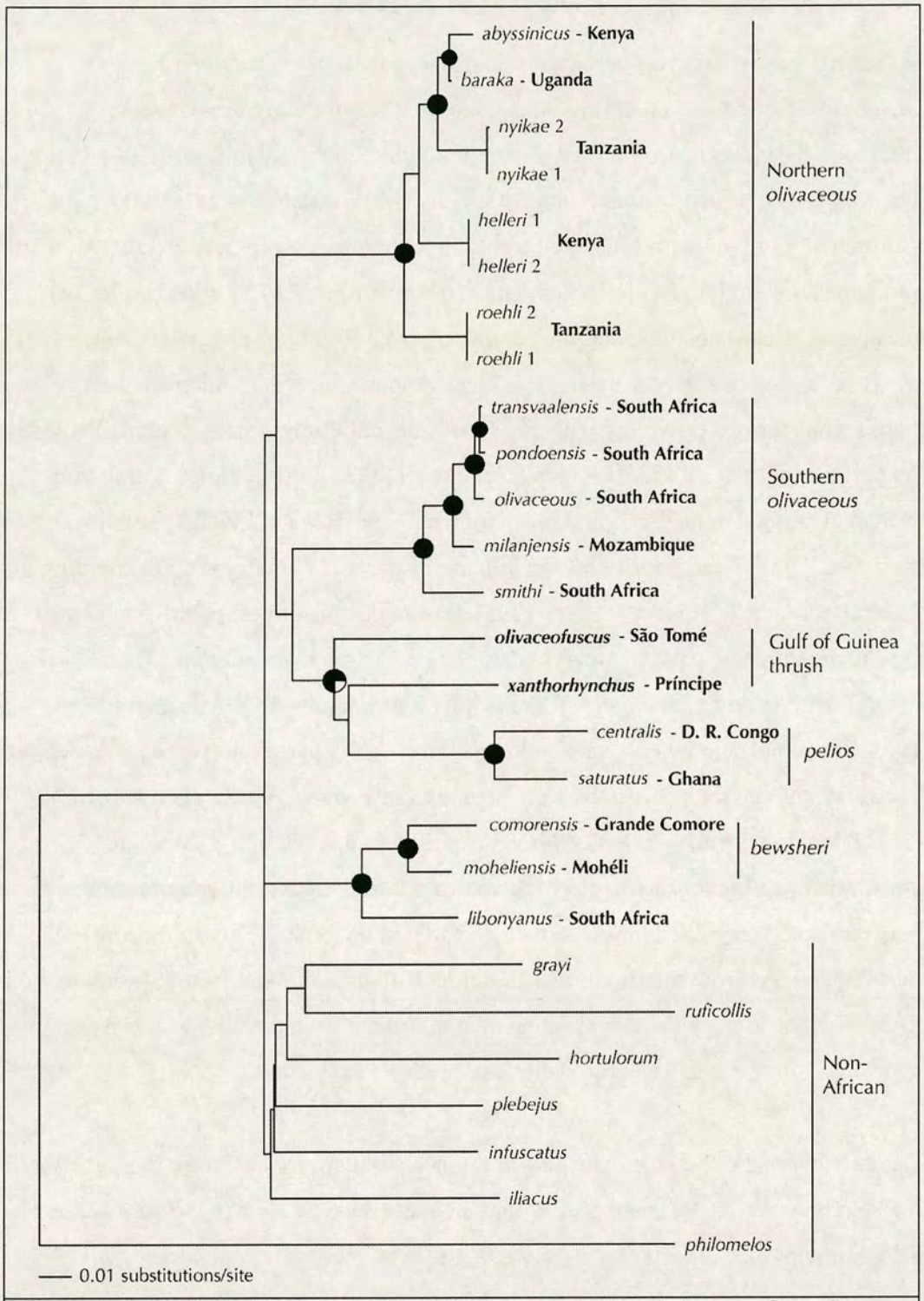


Fig. 4.2. Maximum-likelihood tree obtained from the combined dataset of *Turdus* spp. Bayesian inference recovered the same topology; maximum parsimony recovered the same topology for supported nodes. Nodal support in circles: ML bootstrap and MP bootstrap on bottom left and right hemispheres, Bayesian posterior probabilities on top hemisphere. Black: bootstrap support ≥ 70 %, posterior probability > 0.95.

Table 4.2.

Percentage pairwise sequence divergence between *Turdus* from the different clades inferred from ND2, ND3 and flanking tRNAs (Fig. 4.2). Clade names in bold. Below diagonal: uncorrected distances. Above diagonal: distances corrected with the GTR + Γ model, with α = 0.23, as inferred in MODELTEST (Posada & Crandall 1998). Values in bold = distances between recognized species that are smaller than distances between the Príncipe and São Tomé populations of *T. olivaceofuscus* (shaded).

	Northern *olivaceous*					**Southern *olivaceous***					**G. Guinea**		***pelios***		***bewsheri***		***lib***	**Non-African**						
	aby	*bar*	*nyi*	*hel*	*roe*	*tra*	*pon*	*oli*	*mil*	*smi*	ST	P	*cen*	*sat*	*com*	*moh*	*lib*	*gra*	*ruf*	*hor*	*ple*	*inf*	*ili*	*phi*
abyssinicus	-	0.8	3.2	**3.7**	**4.7**	17.1	16.6	17.2	16.6	17.2	15.9	14.9	15.7	14.6	15.6	13.9	13.9	18.1	24.6	18.8	17.9	16.5	16.8	38.0
baraka	0.8	-	2.4	**2.8**	**3.6**	15.2	14.9	15.3	14.7	15.7	14.6	14.4	15.3	14.3	15.2	13.7	13.5	17.1	23.6	18.7	16.3	14.9	15.1	37.3
nyikae	2.7	2.0	-	**4.4**	**5.3**	17.6	17.4	17.7	17.7	17.8	16.0	15.7	17.8	16.4	17.9	16.0	16.4	19.3	26.5	22.6	17.5	18.1	17.9	37.9
helleri	**3.1**	**2.5**	**3.5**	-	**4.7**	16.0	15.5	16.1	15.0	16.5	14.5	14.5	16.2	14.7	15.0	13.4	14.6	18.4	24.0	18.8	16.4	16.7	15.5	40.1
roehli	**3.8**	**3.0**	**4.1**	**3.8**	-	17.0	16.9	16.2	15.5	15.9	15.8	13.7	16.2	15.1	16.2	14.6	14.1	18.0	24.0	20.0	15.6	16.4	14.7	39.1
transvaalensis	9.3	8.7	9.4	8.9	9.3	-	0.3	0.5	1.7	**4.3**	13.3	14.4	21.8	20.6	19.4	19.9	18.4	18.5	24.3	22.2	18.8	17.1	19.2	37.2
pondoensis	9.2	8.6	9.3	8.7	9.3	0.3	-	0.7	1.7	**4.1**	12.9	13.8	21.3	20.7	19.1	19.1	18.2	18.4	23.7	22.1	18.2	16.3	18.7	36.3
olivaceous	9.4	8.8	9.5	9.0	9.1	0.5	0.6	-	1.8	**4.5**	12.7	13.7	21.0	20.6	18.4	18.6	17.2	18.3	24.0	21.6	17.7	16.0	18.5	36.4
milanjensis	9.2	8.6	9.4	8.6	8.9	1.5	1.5	1.6	-	**4.1**	12.5	13.1	19.6	19.9	17.9	18.8	17.9	18.2	23.3	20.8	17.5	15.9	18.4	40.3
smithi	9.3	8.8	9.3	9.0	9.0	**3.4**	**3.3**	**3.5**	**3.3**	-	12.4	12.5	20.0	19.9	18.5	20.0	17.4	20.0	21.7	20.4	16.9	16.3	17.5	37.1
olivaceofuscus (ST)	9.1	8.6	9.1	8.6	9.1	8.1	8.0	7.9	7.8	7.7	-	**10.4**	13.5	13.9	17.1	15.5	16.6	15.5	25.7	21.4	17.4	14.9	16.7	35.5
xanthorhynchus (P)	8.7	8.5	8.8	8.5	8.3	8.4	8.2	8.1	7.9	7.7	**6.8**	-	14.4	12.6	16.9	15.5	15.0	14.4	24.4	21.1	16.0	15.4	17.0	32.8
centralis	9.0	8.8	9.6	9.1	9.1	10.7	10.6	10.5	10.1	10.2	8.1	8.4	-	5.8	18.1	15.7	18.5	18.3	28.6	22.0	21.0	20.0	21.5	37.4
saturatus	8.7	8.6	9.3	8.7	8.9	10.5	10.5	10.4	10.3	10.2	8.4	7.8	4.4	-	18.0	17.4	16.1	17.9	26.5	22.6	18.8	19.8	21.0	33.7
comorensis	8.9	8.8	9.4	8.7	9.1	10.0	9.9	9.7	9.6	9.6	9.2	9.3	9.7	9.7	-	3.9	**8.5**	21.3	23.0	17.5	16.6	15.8	17.7	46.4
moheliensis	8.4	8.3	9.0	8.1	8.6	10.0	9.7	9.6	9.7	9.7	8.6	8.8	9.0	9.6	3.2	-	**7.2**	19.2	21.2	16.3	17.0	15.0	18.0	36.2
libonyanus	8.2	8.2	9.0	8.5	8.5	9.7	9.5	9.3	9.5	9.2	9.1	8.6	9.9	9.2	**5.7**	**5.1**	-	18.4	21.6	16.9	17.0	16.2	16.8	37.8
grayi	9.8	9.5	10.0	9.8	9.8	9.7	9.7	9.7	9.7	10.1	9.0	8.5	9.9	9.8	10.5	10.0	9.8	-	18.7	21.1	15.6	16.1	19.8	37.5
ruficollis	12.1	11.9	12.4	12.1	12.1	11.9	11.8	11.8	11.8	11.2	12.3	12.1	13.1	12.6	11.7	11.3	11.4	10.5	-	20.8	19.9	21.9	21.5	40.8
hortulorum	10.1	10.0	11.0	10.0	10.5	10.7	10.6	10.5	10.3	10.1	10.5	10.7	10.9	11.2	9.6	9.3	9.4	10.5	11.3	-	19.0	19.3	18.8	37.8
plebejus	9.7	9.2	9.4	9.2	8.9	9.5	9.4	9.3	9.3	9.0	9.6	9.1	10.6	10.0	9.3	9.4	9.4	8.9	10.9	10.0	-	15.6	16.1	32.2
infuscatus	9.2	8.7	9.6	9.2	9.2	9.4	9.1	9.0	9.0	8.9	8.7	8.8	10.3	10.3	8.8	8.5	9.0	9.1	11.5	10.0	8.9	-	16.6	36.0
iliacus	9.4	8.9	9.7	9.0	8.8	9.9	9.8	9.7	9.7	9.2	9.2	9.4	10.8	10.5	9.4	9.4	9.1	10.1	11.3	10.0	9.1	9.2	-	37.0
philomelos	15.1	14.9	15.0	15.4	15.2	14.7	14.6	14.6	15.2	14.6	14.5	14.0	14.8	14.2	16.1	14.6	14.8	14.8	15.9	14.9	13.9	14.6	14.7	-

Phenotypic data

Colour traits

The colour of the bill, iris and legs constituted synapomorphies: colour traits were shared among all individuals from the same population and different between populations. Differences were in agreement with the original description by Salvadori (1901; see 'Introduction').

Morphological traits

Albeit no statistical analyses were performed due to the very small sample sizes, both the field and the museum datasets depict a clear differentiation between Príncipe and São Tomé birds (Figs. 4.4 & 4.5). Tail length did not differ between populations and the single weight available for Príncipe birds (one male) was within the lower bounds of São Tomé males. Similarity in weight contrasts with a large difference in wing length that results in a smaller wing/mass ratio for the Príncipe male. The higher mass for females compared to males is a typical pattern of thrushes, even if otherwise the males are larger (Naurois 1984). The data suggest that longer bills are also thinner (smaller width and depth). The shape of the feet differed between São Tomé and Príncipe, with the Príncipe bird having a very long middle toe in relation to tarsus length (Fig. 4.4.a). The small wing/mass and tarsus/longest toe ratios of the Príncipe bird is in agreement with the field records of a bird whose movements are mostly restricted to the ground. For the field dataset, the first two principal components explained only 53% of the variation, with four components needed to explain 83% of the variation. The first two principal components explained 79% of the variation of the museum dataset.

Fig. 4.4.a. Legend next page. Measurements units: mass in g, all others in mm.

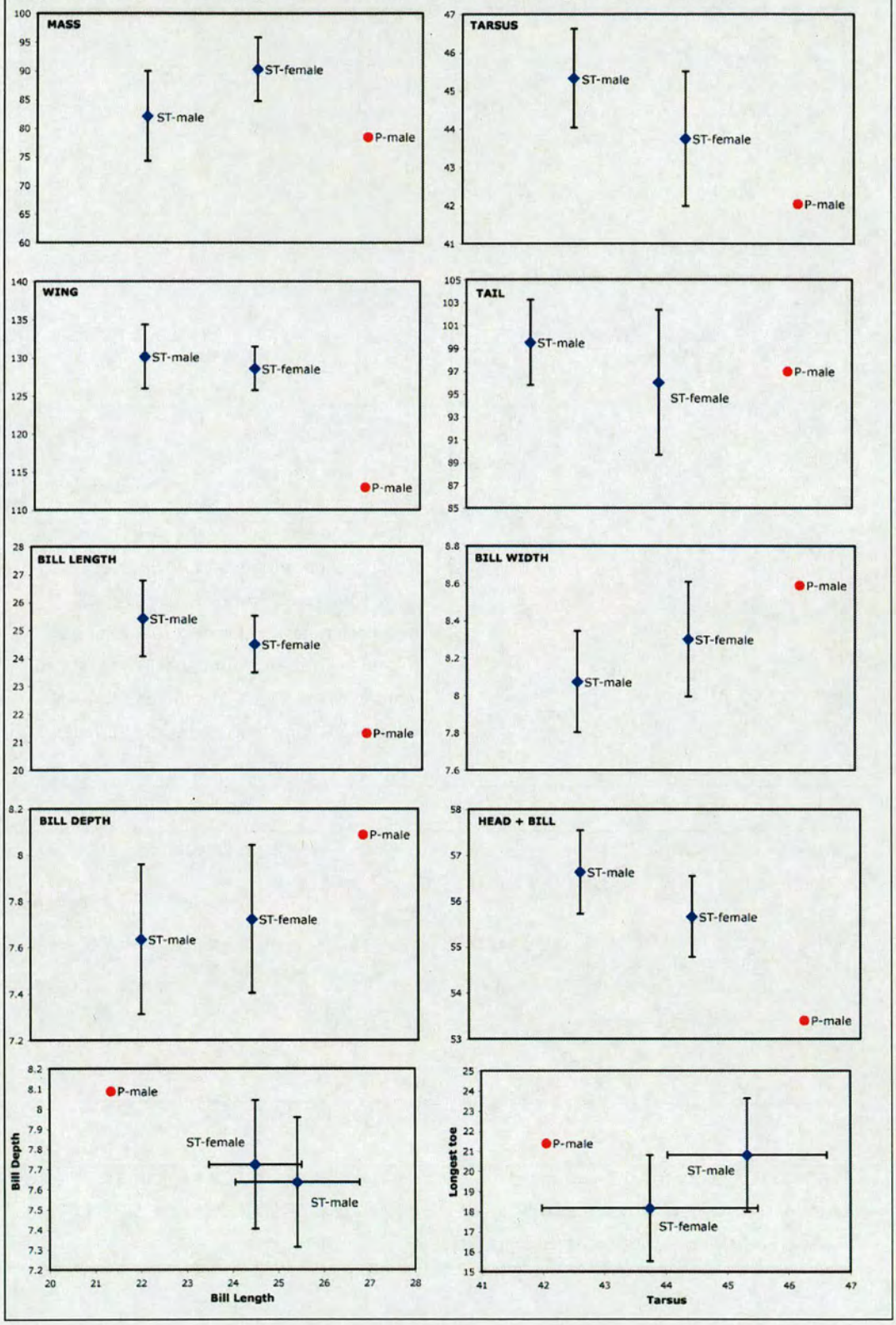


Fig. 4.4.b. Measurements units: mass in g, all others in mm.

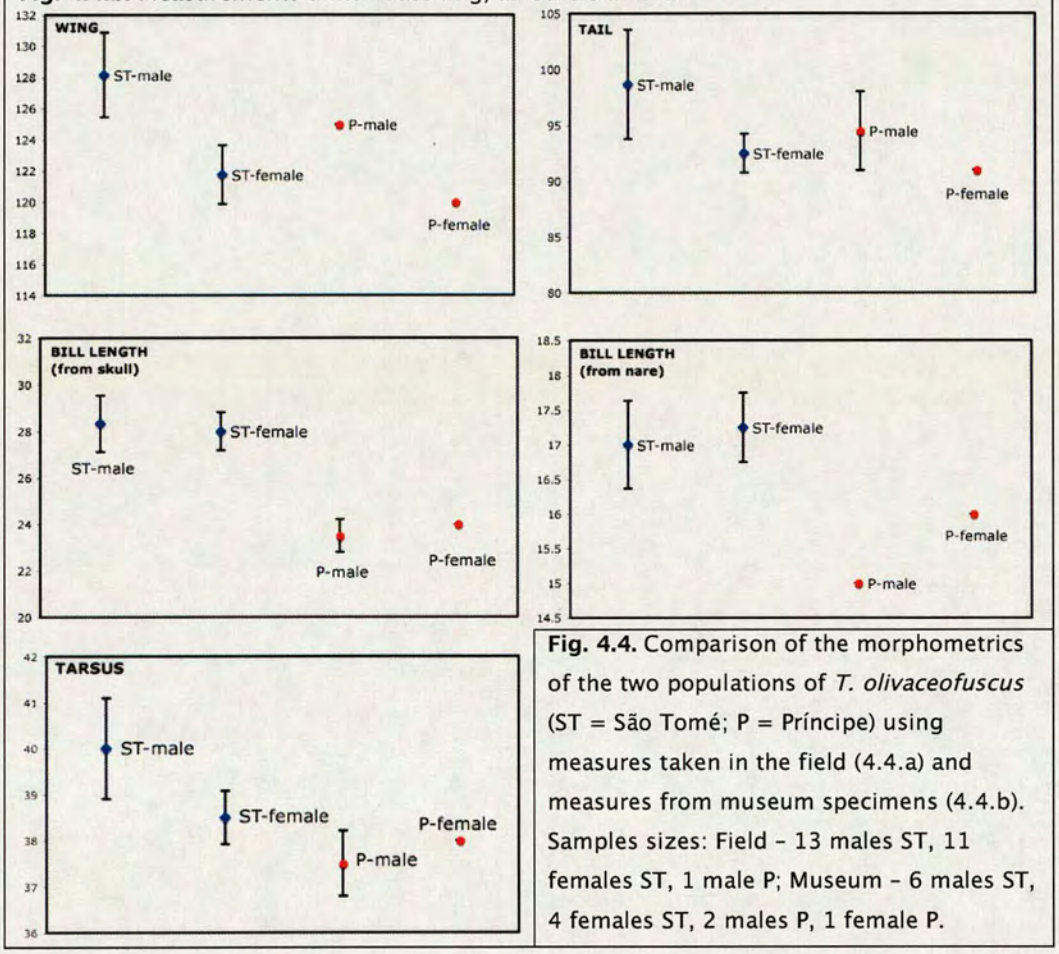


Fig. 4.4. Comparison of the morphometrics of the two populations of *T. olivaceofuscus* (ST = São Tomé; P = Príncipe) using measures taken in the field (4.4.a) and measures from museum specimens (4.4.b). Samples sizes: Field - 13 males ST, 11 females ST, 1 male P; Museum - 6 males ST, 4 females ST, 2 males P, 1 female P.

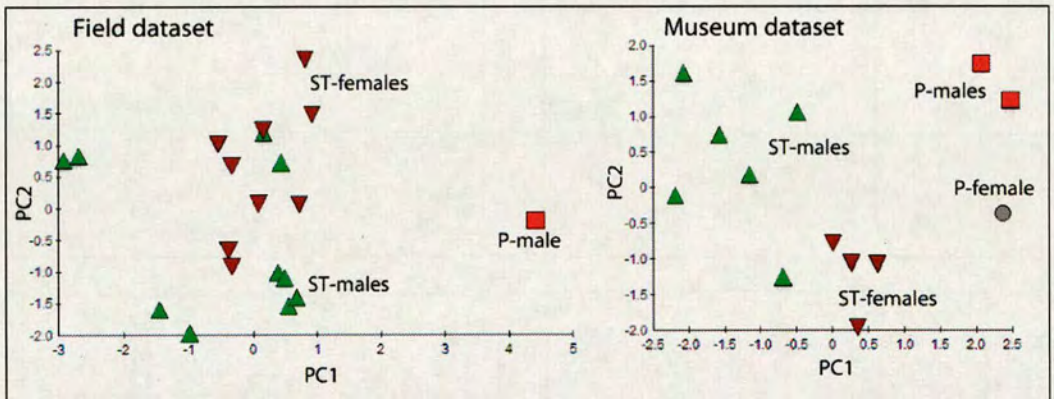


Fig. 4.5. PCA plots for the two morphometric datasets. PC1 and PC2 explain 53.2% of the variation of the field dataset and 79.1% of the museum dataset. P: Príncipe, ST: São Tomé. See methods for description of the datasets.

DISCUSSION

Phylogenetic affinities of the Gulf of Guinea *Turdus*

The Gulf of Guinea thrush *T. olivaceofuscus* is sister to the African thrush *Turdus pelios*, which occurs on nearby mainland, including Mt. Cameroon, and on Bioko Island (Borrow & Demey 2001). This relationship had been recovered in a previous study but had not received node support (Bowie *et al.* 2005). This result rejects the hypothesis based on phenotypic similarity that the Comoros thrush *T. bewsheri* and *T. olivaceofuscus* descended from the same mainland lineage now extinct (Keith & Urban 1992). For both species, a close mainland relative could be identified. The phenotypic differences between the insular and mainland species are not indicative of old age, but likely result from fast evolutionary change on islands. The same discordance between amount of phenotypic change and time since divergence has been found for all Gulf of Guinea birds included in this study (white-eyes, finches, kingfishers: Chapters 2, 3, 5).

Albeit with no support, all phylogenetic methods recovered a paraphyletic relationship between the two *T. olivaceofuscus* populations, with the Príncipe population *T. o. xanthorhynchus* being sister to the African thrush *T. pelios*, and the São Tomé population *T. o. olivaceofuscus* being basal to this group. If colour of the bare parts retained any phylogenetic signal in this case, then the similarities between the Príncipe population and *T. pelios* in these traits are concordant with such relationship. Paraphyly of *T. olivaceofuscus* would be consistent with independent colonisations of each island from mainland that occurred at different times. Confirmation of this pattern will require more extensive genetic sampling since taxon sampling was complete.

The taxonomic status of the thrush from Príncipe

The high level of genetic divergence between the São Tomé and Príncipe thrush populations together with the indication that they might be paraphyletic strongly supports the treatment of the two populations as different species. There is a large number of 'species concepts' that propose different criteria for the delimitation of species, but ultimately all agree that species constitute distinct evolutionary lineages (Mayden 1997; de Queiroz 1998, 2005) – it is mostly on the degree of distinction required for species delimitation that they differ. The 10.4% corrected divergence between the two populations (6.8% uncorrected) is indicative of a long time of isolation, strongly suggesting that the two populations constitute well-separated evolutionary lineages, i.e., valid species under the general 'lineage-based' species definition. Such a level of divergence would also qualify the two populations as distinct species under more restricted species concepts such as the 'evolutionary' and 'phylogenetic' species concepts (Wiley 1978; Cracraft 1989; respectively), had the genetic sampling on Príncipe not been restricted to a single individual. Increasing genetic sampling of Príncipe birds is essential to confirm that intra-population polymorphism does not change the divergence level significantly. This possibility seems nevertheless unlikely because the very small size of the Príncipe population is likely to be reflected in low genetic diversity, and the large phenotypic differences are consistent with a long period of isolation.

A valid point of contention concerns the utility of using genetic divergence data for inferring the taxonomic status of populations. Substitution rates vary among genes and taxa (Mindell & Thacker 1996) and even among demographic contexts – they can for example be faster on island populations (Johnson & Seger 2001). Therefore, there is no 'magic number' above which diverging populations can be considered distinct species. For example, whereas a 2% pairwise genetic divergence value of mitochondrial DNA in birds corresponds generally to a separation of about one million years (García-Moreno 2004; Lovette 2004), variation of substitution rates across taxa can make the 2% value correspond to a divergence time estimate of 200,000 years (in penguins: Lambert *et al.* 2002) and of more than two million years

(in albatrosses: Nunn & Stanley 1998). Nevertheless, genetic distances can provide valuable evidence of specific status when put into context and when concordant with differentiation in other traits (e.g., Helbig *et al.* 1996; Ryan *et al.* 1998; Ryan & Bloomer 1999). Such evidence is particularly valuable for assessing the taxonomic status of allopatric populations where direct evidence for reproductive isolation cannot be obtained (except through captive breeding experiments). The divergence value obtained in this study is well within the range of mitochondrial genetic distances between congeneric species documented in birds (Shields & Helm-Bychowski 1988; Johns & Avise 1998) with species pairs diverging sometimes as little as 2-2.5% (Lovette & Bermingham 1999; Ryan & Bloomer 1999), including island-mainland pairs (Markland & Lovette 2005). More significantly, many valid *Turdus* species showed smaller genetic distances (see 'Results'). Further support comes from the concordance between the high levels of genetic divergence and phenotypic differentiation. The two populations show levels of phenotypic differentiation in the same order of magnitude as other congeneric African thrushes, with the population from São Tomé being more distinct from the Príncipe population than it is to the more distantly related *T. bewsheri* from Anjouan. Additionally, there is some evidence that song, an important trait in species delimitation in birds (Catchpole & Slater 1995), differs between the two populations (S. Lima *in litt*; Baillie & Gascoigne 1999). I never heard the thrush sing during any of my visits to its range (always carried in November-December), a pattern that contrasts with the very vocal behaviour of the São Tomé birds during the same period.

In summary, despite the limitations of the dataset and the problems associated with determining the taxonomic status of closely related allopatric populations, it is recommended to split the Gulf of Guinea thrush *T. olivaceofuscus* into two distinct species because: i) genetic differentiation was well above levels documented for many congeneric species of birds; ii) genetic and phenotypic differentiation were similar or larger than between other *Turdus* species; iii) the monophyly of the two populations was not supported, suggesting that they might derive from independent mainland colonisations; iv) the subsuming of the Príncipe and São Tomé populations within the same species was purely subjective, i.e., not based in any research. The

rarity of the Príncipe birds together with the above arguments further justifies the attribution of species status to the Príncipe population. As a subspecies of *T. olivaceofuscus* the rarity of the Príncipe thrush is overlooked because of the abundance of the São Tomé population; as a separate species it would immediately qualify to move from a relatively minor threat level category ('near threatened') to one of the highest priority categories ('Endangered' or 'Vulnerable'), based on population size and restricted range alone (IUCN 2001).

Splitting of the Gulf of Guinea thrush *T. olivaceofuscus* will simply require returning to the original nomenclature for the two taxa: on São Tomé it remains *T. olivaceofuscus* Hartlaub 1852, and on Príncipe it returns to *T. xanthorhynchus* Salvadori 1901. The respective common names could indicate the island where each occur: São Tomé thrush and Príncipe thrush, respectively. It follows from this proposal that a research programme on the Príncipe thrush should be launched having as its main objectives: i) to consolidate the phylogenetic findings by increasing the genetic and phenotypic sampling and acquiring song recordings, and ii) to establish the current threat level of the Príncipe thrush by estimating its population size (< 1000 = Vulnerable, < 250 = Endangered; IUCN 2001), distribution, breeding periods and habitat requirements. If this research shows that the population is declining then, in combination with its very restricted range, it will qualify to the highest threat category, 'Critically Endangered', independently of the population size (IUCN 2001).

5

PHYLOGENETIC RELATIONSHIPS OF THE GULF OF GUINEA *ALCEDO* KINGFISHERS (ALCEDINIDAE)

The phylogenetic affinities of the two pygmy kingfishers (Alcedinidae: Alcedininae) of São Tomé and Príncipe islands have always been a matter of debate. The forms in each island have been considered full species, different sub-species of the same coloniser, or different subspecies of different colonisers. Taxonomic classifications have used phenotypic traits that were recently shown to have little phylogenetic signal in this sub-family (Moyle *et al.* in press). Mitochondrial and nuclear sequences were used to clarify the origins and relationships of these taxa. A total of 2247 bp were obtained from the mitochondrial genome (ND2, ND3, ATP8, ATP6) and 684 bp from the nuclear genome (myoglobin intron-2). All genes had a compatible phylogenetic signal (as inferred both with maximum likelihood and Bayesian inference). Bayesian inference of the concatenated dataset provided a fully resolved tree. The two populations of São Tomé and Príncipe are very closely related to the mainland malachite kingfisher (*A. cristata*), from which they have diverged very little (uncorrected pairwise distances: 0.05 and 0.1 %). Based on the genetic data alone, the two island populations should be considered different populations of the same species, with phenotypic differentiation likely reflecting adaptation to local conditions.

INTRODUCTION

Our understanding of the processes behind the high levels of bird diversification in the Gulf of Guinea has been limited by the reliance on phylogenies inferred from phenotypic traits. These phylogenies often vary from researcher to researcher depending on the phylogenetic value that each attributes to different traits (reviewed in Jones & Tye 2006). This subjectivity derives from the scarcity of phylogenetically informative characters in birds (Bock 1963, 1967; Hafner *et al.* 1984) – a problem made evident by the fact that the commonly used morphological traits for bird classification, such as bill shape and body size, are also those most commonly used to infer the action of selection (Grant 1986; Schluter & Smith 1986; Smith & Girman 2000). Plumage colour in particular has been shown to be of little reliability for phylogenetic inference (Moreau 1957; Price & Birch 1996; Burns 1998; Omland & Lanyon 2000; Warren *et al.* 2005; Moyle *et al.* in press). This issue is compounded on islands where adaptation to the insular condition often leads to convergent changes both in metric and colour traits (Grant 1965c, 1968; Blondel 2000; Clegg & Owens 2002). The previous chapters illustrated this problem: relationships inferred from molecular data often provided a distinct picture from classic taxonomy.

In the Gulf of Guinea, phylogenetic uncertainties range from species that are so divergent from any mainland species that they cannot be placed in any family (the São Tomé short-tail *Amaurocichla bocagii* and Dohrn's thrush-babbler *Horizorhinus dohrni* are *incertae sedis*), to taxa that are very similar to mainland species. The latter might constitute incipient or recently diverged species, making them good models for the study of the processes driving speciation (Foster *et al.* 1998; Edwards *et al.* 2005). Such study must be based on reliable phylogenetic hypotheses; these must be derived from molecular markers not linked to adaptive phenotypic traits.

On São Tomé and Príncipe, the two representatives of the pygmy kingfishers (Alcedinidae: sub-family Alcedininae) are one example of taxa at the population/species border and demonstrate the difficulty of inferring phylogenies

based on phenotypic traits. Each island has one phenotypically distinct kingfisher form, which have been classified in all possible combinations (Table 5.1): different subspecies of the same mainland species, different subspecies of different mainland species, or distinct taxonomic species (derived from the same or different ancestors). The São Tomé kingfisher (form *thomensis*) is phenotypically close to the malachite kingfisher *Alcedo cristata*, whereas the Príncipe kingfisher (form *nais*) is intermediate between the white-bellied *A. leucogaster* and the malachite kingfishers (Fig. 5.1; Amadon 1953; Fry & Naurois 1985). These two mainland species have been shown to be closely related by molecular phylogenies (Moyle *et al.* in press). The white-bellied kingfisher occurs also on Bioko, the Gulf of Guinea island closest to the mainland (32 km) to which it was connected during glaciations. The most recent review of the birds of the Gulf of Guinea concluded that it is currently not possible to determine the affinities of these taxa, and they were kept as subspecies of the forms they most closely resemble only as 'an arrangement of convenience' (Jones & Tye 2006). The instability of the traditional taxonomic classifications derives from the fact that they were based in traits that have recently been shown to be non-informative in kingfishers (Moyle *et al.* in press).

In this study I use mitochondrial and nuclear DNA to clarify the phylogenetic position of the São Tomé and Príncipe kingfishers. This will contribute to the establishment of solid phylogenetic bases required to further our understanding of the processes behind the high diversity of endemic birds on the Gulf of Guinea. Additionally, this study will complete the recent molecular revision of the Alcedininae (Moyle *et al.* in press), which included representatives of all named species of the group except for the São Tomé and Príncipe forms.

Table 5.1.

Common classifications for the kingfishers from São Tomé and Príncipe.

Island	**Species**	**Subspecies of:**		
Príncipe	*nais*	*cristata*	*leucogaster*	*leucogaster*
São Tomé	*thomensis*	*cristata*	*leucogaster*	*cristata*
Sclater 1924		X		
Peters 1945		X		
Bannerman 1953		X		
Amadon 1953			X	
White 1965			X	
Praed & Grant 1970			X	
Snow 1978		X		
Wolters 1982	X			
Forshaw 1983	X			
Fry & Naurois 1985				X
Sibley & Monroe Jr 1990	X			
Dowsett & Dowsett-Lemaire 1993				X
Fry *et al.* 1992				X
Sttatersfield *et al.* 1998[1]	X			
Dickinson 2003				X
Sinclair & Ryan 2003	X			
Jones & Tye 2006[2]				X

Original classification - Príncipe: *Alcedo leucogaster nais* (Kaup 1848); São Tomé: *Corythornis* [=*Alcedo*] *thomensis* Salvadori 1902. [1]: Classification used by BirdLife International. [2]: Classification used as 'an arrangement of convenience'.

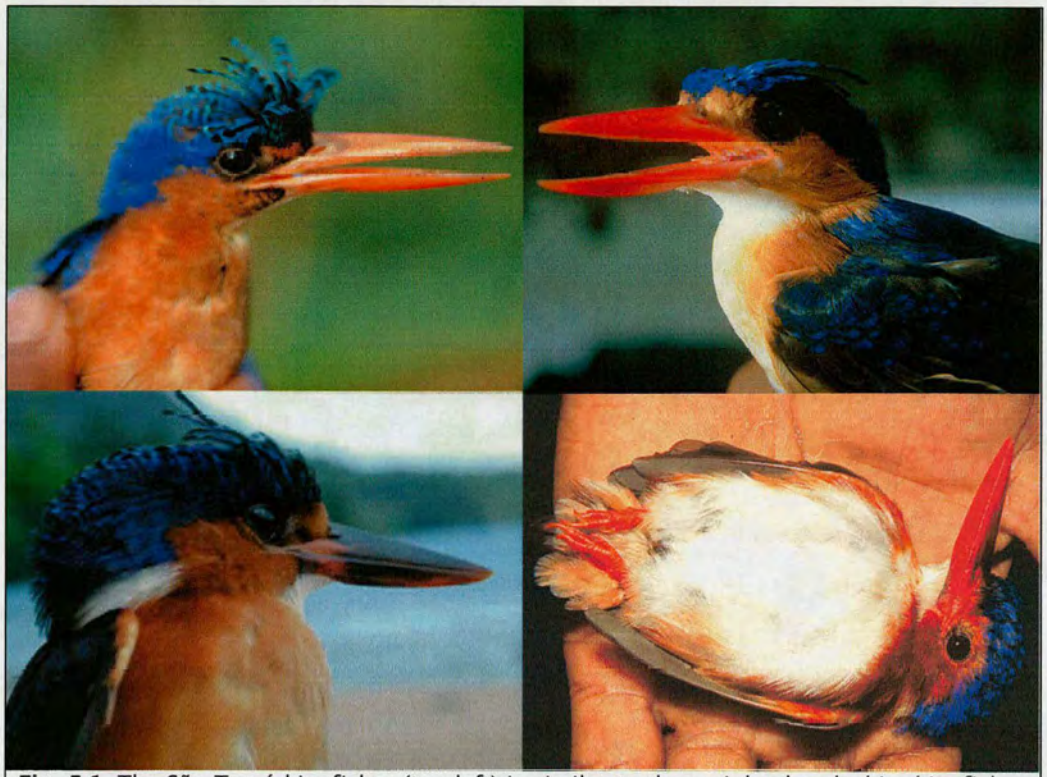


Fig. 5.1. The São Tomé kingfisher (top left) is similar to the mainland malachite kingfisher in that it has a prominent crest and a rufous belly. The Príncipe kingfisher (top right) shows a loss of rufous feathers in the belly. In this regard it has been considered close to the white-bellied kingfisher (bottom right, Bioko). Nevertheless, it differs from the white-bellied kingfisher in that the white colour is not so well delimited and rarely extends to the breast (like in the juvenile shown, bottom left). Additionally, birds from Príncipe have also a crest that they raise as does the malachite kingfisher. [Príncipe kingfisher juvenile photo by Dong Lin, California Academy of Sciences].

METHODS

Taxon and character sampling

Ingroup sampling included 16 individuals from nine taxa in the sub-family Alcedininae, representing the four clades from this subfamily (Moyle *et al.* in press; Appendix 5.1). Samples of *A. leucogaster* were obtained both from Bioko Island and the mainland. The outgroup comprised one representative from each of the other two sub-families (Cerylinae and Daceloninae) of the Alcedinidae. Samples consisted

either of blood samples collected non-destructively from the brachial vein of mist-netted wild individuals (stored in absolute ethanol), or of liver tissue. Total genomic DNA was extracted from the blood using a 'DNeasy Tissue Extraction Kit' (Qiagen), following the protocol for DNA extraction from animal tissues. Adaptation of this protocol for blood samples stored in absolute ethanol is described on Chapter 2. DNA was extracted from liver using a CTAB-based protocol (Winnepenninckx *et al.* 1993).

To ascertain the phylogenetic affinities of the São Tomé and Príncipe taxa, nucleotide sequence data from both the mitochondrial and nuclear genome were used. Mitochondrial sequences included the entire NADH dehydrogenase subunit 2 (ND2) and subunit 3 (ND3), and the entire ATP synthase subunit 6 (ATP6) and 8 (ATP8), in a total of 2247 base pairs (bp). From the nuclear genome, the entire sequence of the second intron of the myoglobin gene (Myo2) was obtained (684 bp). Overlapping sequence fragments were amplified using the primers and PCR conditions detailed in Appendix 5.2. PCR products were electrophoresed on a 1.5% agarose gel and visualized under UV light with ethidium bromide to check for the correct fragment size, and to control for the specificity of the amplifications. The PCR products were purified using either the 'QiaQuick PCR Purification Kit' or the 'Qiagen Gel Purification Kit' (Qiagen). Amplification primers were used for cycle sequencing both DNA strands for all samples using the 'CEQ Dye Terminator Cycle Sequencing' (Beckman Coulter, Inc, Fullerton, CA, USA) or the 'Big Dye' (Applied Biosystems Inc. Foster City, CA, USA) terminator chemistries and run on automated sequencers (CEQ2000 DNA Analysis System, ABI3100, ABI3730). Sequences were aligned with Sequence Navigator (Applied Biosystems) or Sequencher 4.1 (Genecodes) and by eye. The occurrence of single nucleotide polymorphisms (SNPs) in the nuclear locus was suggested by the presence of double peaks. These double peaks were coded using the appropriate IUPAC codes. Sequences were deposited in GenBank.

Phylogenetic analyses

Molecular phylogenies were estimated using model-based approaches (maximum likelihood, ML, and Bayesian inferences, BI), as implemented in PHYML v2.4 (Guindon & Gascuel 2003) and MRBAYES 3.1 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). Likelihood models were estimated with MRMODELTEST 2.0 (Nylander 2004) and the best-fit models were selected using the Akaike Information Criterion (Akaike 1973). Clade supports for the ML analyses were assessed by non-parametric bootstrapping (Felsenstein 1985; 100 replicates). For the Bayesian analyses, four incrementally heated Metropolis-coupled MCMC chains were run for two million generations with trees sampled every 100 generations. The first 200,000 generations (2000 trees) were discarded ('burn-in' period) and the posterior probabilities were estimated for the remaining sampled generations. Three independent Bayesian runs initiated from random starting trees were performed for each dataset, and the log-likelihood values and posterior probabilities were checked to ascertain that the chains had reached stationarity. The concatenated analyses were performed freeing the different parameters (base frequencies, rate matrix, shape parameter, proportion of invariable sites) to vary between the partitions (genes and codon positions) using the *unlink* command.

To assess whether the different genes had a compatible phylogenetic signal a parsimony based incongruence length differences (ILD) test (Farris *et al.* 1995) was performed in PAUP 4.0b10 (Swofford 2003), excluding invariant sites (Cunningham 1997), and ran for 1000 bootstrap repetitions. Because the utility of the ILD test to detect phylogenetic congruence has been criticised (Yoder *et al.* 2001; Barker & Lutzoni 2002), compatibility was also evaluated by comparing the topologies and nodal support of the individual gene trees obtained under the two analytical methods (ML, BI). Thresholds for incongruence were set at 70 % for bootstrap values (Hillis & Bull 1993) and at 0.95 for posterior probabilities (Huelsenbeck & Ronquist 2001).

The most adequate partitioning of the protein coding genes and concatenated data sets was evaluated using the Bayes Factors approach on biologically relevant

partitions (e.g., protein coding genes by codon position). The harmonic means of the likelihood values of the phylogenies sampled after convergence of a MRBAYES run were used as approximations of the marginal likelihood of the two compared models (M_1 et M_2) to calculate the respective Bayes factor (B_{12}). A value greater than 10 for $2\ln B_{12}$ was considered as a strong evidence against the alternative model, M_2 (Kass & Raftery 1995; Nylander *et al.* 2004).

RESULTS

Sequence characteristics

The pattern of sequence variation followed the expected pattern: the nuclear intron was the less variable marker (9% variable sites; 6% informative), contrasting with the mitochondrial markers that had levels of variation ranging from 26 to 37% (informative sites: 19 to 28%); ATP8 was the most variable mitochondrial marker; in the mitochondrial genes, the third codon position was the most variable, and the second was the least (Table 5.3). The nuclear intron had a relatively balanced base composition (Table 5.3), identical to that found in the study of all species from the sub-family Alcedininae (Moyle *et al.* in press). The mitochondrial markers had a deficiency of guanine, which was more severe at third positions, and an excess of adenine and cytosine (Table 5.3), a typical pattern of the avian mitochondrial genome (e.g., Baker & Marshall 1997; Moore & DeFilippis 1997; Voelker & Spellman 2003; Klicka *et al.* 2005), which is also found in kingfishers (Moyle 2006; Moyle *et al.* in press). Base frequencies of the combined dataset of each marker and of each codon position (mitochondrial genes) were homogeneous across taxa ($P \geq 0.98$ in all χ^2 tests).

Multiple lines of evidence suggest that all four mitochondrial genes are of mitochondrial origin rather than being nuclear pseudogenes, or *numts* (Sorenson & Quinn 1998; Bensasson *et al.* 2001). All chromatograms were of good quality and no double peaks, indicative of multiple copies, were observed; sequences of overlapping fragments were the same. All sequences translated appropriately, and the patterns of

sequence composition and variation just described are in agreement with the dynamics of the mitochondrial genome. All sequences gave the same topologies (variation only at the resolution level) with similar branch lengths. These topologies were in agreement with recent molecular phylogenies of the kingfishers (Marks & Willard 2005; Moyle 2006; Moyle *et al.* in press), even though DNA for this study was obtained mostly from blood (nuclear rich) whereas it was mostly of muscle origin (mitochondrial rich) in the previous studies. For all specimens, the ND3 gene had the 'silent base' described in several bird groups, an insertion in base 194 that does not disrupt the reading frame (Mindell *et al.* 1998). The ND2 gene presented a three base insertion on the three most closely related taxa: *A. cristata*, *A. thomensis*, and *A. nais* (see 'Phylogenetic inference'). This insertion translated into a new amino acid and did not disrupt the ND2 translation both up- and downstream. This insertion is very unlikely to be indicative of the amplification of a *numt* for the reasons listed above and also because it was found independently by Moyle (2004) and Moyle *et al.* (in press): in 55 kingfisher species sequenced for ND2, the same insertion was found only in *A. cristata*. These authors used DNA extracted from muscle, whereas in this study DNA came from blood. A two-codon insertion in the ND2 has also been described in the owl *Ninox novaseelandiae* (Harrison *et al.* 2004).

Table 5.3.

Sequence characteristics of the 5 genes based on all ingroup taxa.

	Sites	Var	Info	%A	%C	%G	%T	Model
Mitochondrial	2247	692	497	30.5	34.2	10.4	24.9	GTR+Γ+I
ND2								
All	1044	323	229	31.5	34.4	10.5	23.6	GTR+Γ
1st	348	79	20	36.9	27.9	15.2	20.0	GTR+Γ
2nd	348	21	14	17.7	35.4	8.5	38.5	HKY + I
3rd	348	223	156	39.9	40.0	7.7	12.4	GTR+Γ+I
ND3								
All	351	93	66	26.3	34.4	12.1	27.2	GTR+I
1st	117	22	16	24.4	31.4	18.9	25.4	K80+Γ
2nd	117	5	2	17.1	27.7	12.0	43.3	HKY + I
3rd	117	66	48	37.4	44.0	5.5	13.0	GTR
ATP8								
All	168	62	47	35.7	35.3	6.5	22.5	GTR+I
1st	56	11	10	44.5	25.4	3.5	26.4	GTR+Γ
2nd	56	14	10	21.7	39.2	10.5	28.7	GTR+I
3rd	56	37	27	40.8	41.3	5.5	12.4	GTR+I
ATP6								
All	684	214	155	29.9	33.7	10.4	26.1	GTR+Γ+I
1st	228	39	28	33.7	37.0	14.5	14.8	GTR+Γ
2nd	228	8	4	14.9	29.2	9.2	46.7	HKY
3rd	228	167	123	41.0	34.8	7.5	16.7	GTR+I
Nuclear (Myo2)	684	61	39	28.1	20.6	24.9	26.4	K2P+I
Combined	2931	753	536	29.9	31.1	13.7	25.2	GTR+Γ+I

Var: variable sites - Info: parsimony informative sites. Models of sequence evolution: GTR, General time reversible (Rodríguez *et al.* 1990); HKY, Hasegawa *et al.* (1985); K2P, Kimura 2 parameter (Kimura 1980); Γ = variation of nucleotide substitution rate among sites is described by a gamma distribution; I = a proportion of the sites are considered to be invariant.

Phylogenetic inference

The incongruence length difference test could not reject the null hypothesis that different genes had the same phylogenetic signal (congruence mitochondrial genes: $P = 0.86$; congruence mitochondrial v. nuclear: $P = 0.762$; congruence all genes: $P = 0.87$). Congruence of the phylogenetic signal was confirmed visually by inspection of ML and BI trees inferred from each marker: there were no supported nodes in conflict. Therefore all the markers could be combined for phylogenetic inference. Because the different genes followed different models of sequence evolution (Table 5.3), an analysis with the concatenated dataset was performed only under BI. BI can implement different models for different partitions (Nylander *et al.* 2004), whereas current ML methods cannot. The best data partition, as determined by the Bayes factor approach, was by gene and by codon position (13 partitions), followed by partition by codon position (4 partitions; $2\ln B_{12} = 85$). All other partition schemes were much worse ($2\ln B_{12} > 1490$).

Bayesian inference of the concatenated dataset provided a fully resolved tree (Fig. 5.2) that recovered the same relationships as previous molecular phylogenies of the kingfishers (Marks & Willard 2005; Moyle 2006; Moyle *et al.* in press). The two kingfishers from São Tomé and Príncipe grouped with *A. cristata*, from which they diverged very little (Table 5.4). For the 2247 bp mitochondrial dataset, *thomensis* differed in only 10 nucleotides from *cristata*; *nais* had 19 bp differences from *thomensis* and 21 from *cristata*. The 684 bp intron could not resolve the relationships of these taxa (only one and two bp differences). This close relationship is further supported by the synapomorphic insertion of an extra codon on the ND2. This '*cristata*' group was sister to *A. leucogaster*. There was no differentiation between the populations of *A. leucogaster* from mainland Cameroon and from Bioko Island (Fig 5.2).

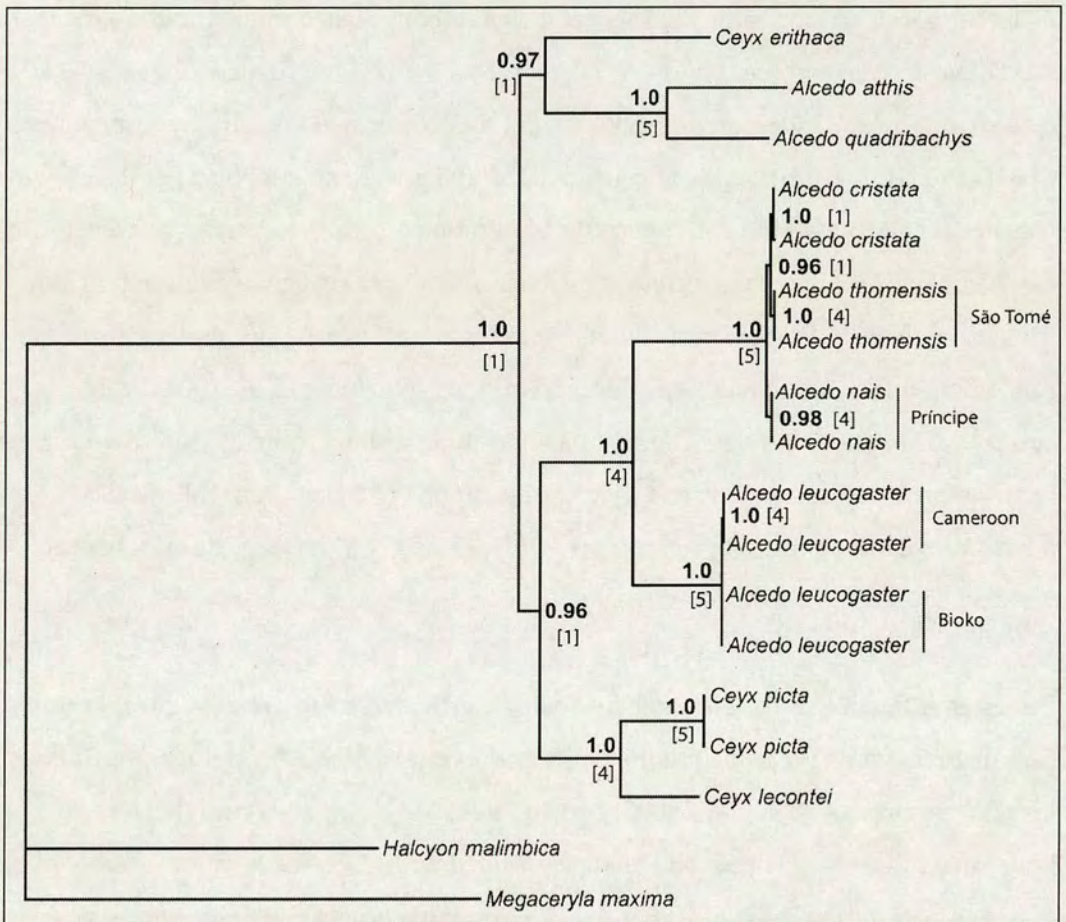


Fig. 5.2. Bayesian phylogeny estimated from the concatenated dataset (ND2, ND3, ATP8, ATP6, myo2) using a partition scheme by gene and codon position (models for each partition on Table 5.2). Posterior probabilities in bold. Numbers in square brackets indicate the number of single-gene trees estimated under ML that recovered each node with high bootstrap support (> 70%).

Table 5.4.

Uncorrected pairwise distances obtained from the mitochondrial dataset (ND2, ND3, ATP8, ATP6, above diagonal) and the myoglobin intron-2 (below diagonal). In bold: distances between São Tomé and Príncipe taxa and *Alcedo cristata*. Taxa names abbreviated from full names in Fig. 5.2.

	A. att	*A. cri*	*A. leu*	*A. nai*	*A. qua*	*A. tho*	*C. eri*	*C. lec*	*C. pic*
A. att	-	0.1560	0.1405	0.1582	0.1083	0.1587	0.1377	0.1381	0.1359
A. cri	0.0475	-	0.1070	**0.0093**	0.1484	**0.0045**	0.1480	0.1435	0.1444
A. leu	0.0450	0.0095	-	0.1102	0.1238	0.1090	0.1328	0.1285	0.1227
A. nai	0.0472	**0.0016**	0.0076	-	0.1489	0.0085	0.1484	0.1426	0.1444
A. qua	0.0092	0.0507	0.0460	0.0481	-	0.1498	0.1292	0.1332	0.1292
A. tho	0.0487	**0.0032**	0.0091	0.0015	0.0497	-	0.1484	0.1440	0.1453
C. eri	0.0466	0.0397	0.0346	0.0366	0.0477	0.0381	-	0.1275	0.1136
C. lec	0.0480	0.0271	0.0210	0.0258	0.0492	0.0273	0.0346	-	0.0820
C. pic	0.0480	0.0253	0.0180	0.0228	0.0491	0.0243	0.0316	0.0090	-

DISCUSSION

The 2931 bp genetic dataset used in this study settles the issue of the origins of the kingfishers of the oceanic islands of the Gulf of Guinea and suggests the need for the adoption of a new taxonomic classification. Both the São Tomé (*thomensis*) and the Príncipe (*nais*) kingfishers are closely related to the Malachite kingfisher (*A. cristata*), in agreement with some of the first classification schemes. This relationship is underlined by the fact that they share the only described case in kingfishers of a codon insertion in the ND2 gene. Based on genetic distances alone they do not constitute valid species, and even subspecific status would not be warranted. Considering that the *A. cristata* samples are from Malawi, it is likely that genetic similarity will be even greater with *cristata* birds from nearby mainland. Genetic differentiation is less between *thomensis* and *cristata*, in agreement with their higher phenotypic similarity, which has long been reflected in the wider acceptance of a link of the São Tomé population to *cristata* by traditional taxonomies. Amadon (1953) had proposed that the oceanic island taxa were derived from *leucogaster* because this species is present on the land-bridge island of Bioko that would have favoured a stepping-stone colonisation model, and because it is a

dense forest species and therefore pre-adapted to São Tomé and Príncipe habitats, whereas *cristata* is a species from more open habitats. This hypothesis can now be confidently refuted – and admittedly its arguments were weak: i) currently, Príncipe Island is only 10 km further from mainland than it is to Bioko, the opposite being true when sea levels were lower during glaciations (Jones & Tye 2006); and ii) forest species tend to be more sedentary than open habitat ones (Blondel 2000), which makes *cristata* a more likely coloniser. The Príncipe kingfisher was slightly more differentiated from *cristata/thomensis* but divergence was still only 0.1%. Apart from metric differences, birds in this population have lost most of the rufous in the belly, one of the main reasons for having been often put close to *leucogaster*. Nevertheless, the white colour of Príncipe birds is far from the clean and well-delimited white of *leucogaster*. It does not extend up to the breast (Fig. 5.1), and appears to result from heterogeneous pigment loss because the extent of white is variable and some rufous feathers may be retained within the white patch (pers. obs). This pattern is consistent with the common trend for colour loss on island populations (Grant 1965c), which has indeed occurred with most taxa of the Gulf of Guinea islands (Amadon 1953). Birds from Príncipe also differ from *cristata/thomensis*, being again more similar to *leucogaster* in this respect, in that the juveniles are similar to the adults rather than having a blackish plumage. These differences suggest that colonisation of Príncipe was earlier than colonization of São Tomé and/or that gene flow from mainland into Príncipe is more restricted than into São Tomé.

Based on the current molecular information, the insular morphs should be considered populations of the same species, with phenotypic differentiation reflecting local adaptation. Genetic similarity could be due to recent colonisation and/or continual gene flow between populations. The distances between islands and mainland (150-250 km) might not constitute a major barrier for kingfishers, as is demonstrated by the occasional occurrence of the larger pied kingfisher (*Ceryle rudis*) on the islands (Jones & Tye 2006). Additionally, on clear days São Tomé and Príncipe, or their cloud cover, can be seen from each other (pers. obs.; Jones & Tye 2006). An alternative hypothesis is that the genetic similarity of the three populations is a consequence of rare hybridization events between isolated populations that led to

introgression of neutral molecular markers. The populations could therefore have had a mostly independent evolutionary history since the islands were colonised, which would be reflected in differentiation of the genes underlying adaptive traits but not on neutral markers. This is well documented for the mitochondrial genome: because it is not linked to the nuclear genome, which contains the loci that determine individual fitness, it can penetrate the genome of a foreign population with very low levels of hybridisation (Barton & Jones 1983; Takahata & Slatkin 1984). This explains the introgression of mtDNA across species borders in many animal groups (Ferris *et al.* 1983; Goodman *et al.* 1999; Ting *et al.* 2000; Shaw 2002; Weisrock *et al.* 2005) including birds (Degnan 1993). Birds in particular can maintain the potential for hybridising successfully after million of years of separation (Grant & Grant 1992; Price & Bouvier 2002). However, the fact that the nuclear marker recovered the same phylogenetic relationships for all taxa and not only for those deemed to be hybridising suggests that more extensive gene flow is taking place.

Retention of the subspecific status for each insular population (São Tomé: *A. cristata thomensis*; Príncipe: *A. cristata nais*) could be justified as a way of highlighting that these populations have evolved fixed phenotypic differences (larger sample sizes will be required to confirm this). These may represent local adaptations that increase the functional diversity of the species, and are therefore relevant for the preservation of the evolutionary process (Lesica & Allendorf 1995; Crandall *et al.* 2000; Phillimore & Owens 2006). A population genetics approach is required to determine whether dispersal events between populations are frequent or if the demographic dynamics of the insular populations are determined primarily by birth and death rate rather than by immigration and emigration (Moritz 1994). Occasional secondary contact events would then bring additional genetic variation into the isolated gene pools with the potential to promote further diversification (e.g., Arnold & Emms 1998; Grant & Grant 1998; Petren *et al.* 2005).

6

WHAT DRIVES DIVERGENCE OF ALLOPATRIC POPULATIONS? DIFFERENTIATION OF AN ENDEMIC BIRD IN THREE ISLANDS: THE PRÍNCIPE SEEDEATER *Serinus rufobrunneus*

In most animals, speciation seems to require a phase of geographic isolation between the diverging populations. This is particularly true in birds. Nevertheless, the current 'ecological model' of speciation is based on cases where gene flow is not interrupted, and very few studies have addressed the speciation problem in allopatric situations. Speciation in allopatry could depart from the ecological model if genetic drift plays a role. Additionally, population isolation might allow a faster divergence of pre-mating isolating barriers driven by random processes. I describe the genetic and phenotypic patterns of variation of the Príncipe seedeater *Serinus rufobrunneus* between the islands of São Tomé, Príncipe and Boné de Jóquei, and between two habitats in São Tomé (forest and plantations) in order to identify the most likely processes involved in divergence of small isolated populations and to discuss how these could relate to the evolution of reproductive isolation. Characterization of genetic variation was based on 15 microsatellites. Phenotypic traits analysed included adaptive morphological traits linked to diet and flight performance, and mate recognition traits (plumage colour and song). I investigated the role of random processes on the evolution of allopatric populations by testing for concordance between phenotypic and neutral genetic variation. Only the evolution of the colour of the mantle could be accounted by random processes alone. Morphological shifts were in the direction predicted by differential selective pressures. Song divergence between islands did not depart much from the neutral pattern, but within São Tomé it was linked to habitat differences. Together with song, colour of the collar was the only other trait that diverged between São Tomé habitats. In sum, the three allopatric populations differentiated in all dimensions, whereas the parapatric populations diverged only in mate recognition traits. This result cannot be used to support the primacy of isolation over differential selection in the speciation process, since the allopatric populations have been diverging for thousands of years whereas the parapatric populations had only c. 300 years to diverge. However this result supports the potential for rapid divergence of mate recognition systems and for their implication in the first stages of the speciation process.

INTRODUCTION

For most of the 20th century, speciation was understood to be a by-product of the cessation of gene flow between populations in allopatry (Berlocher 1998). Ecological factors were considered the most important factors in population divergence, but they could lead to speciation only if preceded by a bout of drift in isolated populations (Mayr 1963). Genetic drift would change allele frequencies and allele linkage patterns, creating different gene pools where selection (ecology) would act and produce different results (Mayr 1942, 1954, 1963). The current interest in speciation research was partly triggered by a need to disprove Mayr's axiomatic view that full interruption of gene flow is a necessary condition for the splitting of lineages to occur. As detailed in Chapter 1, this new research programme led to a paradigm shift, hereafter designated as the 'ecological model': the role of drift has been dismissed, and speciation is now widely perceived as a deterministic process driven by selection (Morell 1999).

Nevertheless, to demonstrate that selection can override gene flow, the ecological view of speciation was mainly developed from models in parapatric and sympatric situations (Orr & Smith 1998). Mayr's model, by definition, only applies to allopatric cases. It is therefore essential to establish whether the speciation process as described by the ecological model can be generalised for cases of allopatric populations or if, in contrast, the mechanisms of speciation are dependent on the geographic context where it takes place. This is especially important as, despite all recent efforts to find counter-examples, speciation in animals appears to be a process that occurs overwhelmingly in allopatry or at least requires an allopatric phase (Barraclough & Vogler 2000; Coyne & Orr 2004; Fitzpatrick & Tureli 2006). This is a consensual view in birds (Grant & Grant 1997a; Coyne & Price 2000; Edwards *et al.* 2005), where only a single case of sympatric speciation can be safely assumed – a case with unique characteristics (cultural speciation during host-switching by brood parasitic indigobirds: Sorenson *et al.* 2003).

Interestingly, we know much less about the mechanisms of speciation in allopatric populations (the dominant pattern) than we know about those in sympatric and parapatric populations. The latter constitute the bulk of recent speciation research. Lack of interest in the allopatric model could be explained in part by the 'triviality' of the problem: given enough time any two isolated populations are expected to diverge to the point of being reproductively isolated even when divergence is fully neutral (Nei *et al.* 1983a; Turelli *et al.* 2001). But this assumed 'triviality' does not elucidate what really happens. Laboratory studies have supported the ecological model in isolated populations, with divergent selection being more important than drift in the evolution of reproductive isolation (Rice & Hostert 1993). Studies in natural systems are much scarcer and have often focused on determining whether vicariance or dispersal caused population isolation (Voelker 1999; Zink *et al.* 2000; Sanmartín & Ronquist 2004), rather than on the speciation process. At continental scales, recent studies have shown that adjacent populations in different habitats are phenotypically divergent, whereas populations isolated by thousands of kilometres but in the same habitat show no divergence (Smith *et al.* 1997; Schneider *et al.* 1999; Smith *et al.* 2005). These studies demonstrate clearly the primacy of ecology in promoting divergence and support speciation in parapatry. Nevertheless they cannot explain directly how speciation occurs in isolated populations (since no divergence was detected), but rather suggest indirectly that isolated populations must be under different selective pressures in order to speciate. Furthermore, these studies looked at large populations where any effect of drift, the distinguishing feature of the allopatric model, will always be weak (Barton & Charlesworth 1984). The conclusive evidence that drift plays a negligible or no role in speciation must come from situations where drift is more likely to be strong, i.e., in small isolated populations, like those on oceanic islands. This was indeed the most likely situation for speciation to occur according to Mayr (peripatric speciation), and led him to develop the founder-effect speciation model (Mayr 1954, 1963). This particular model has now been confidently refuted both by theory and laboratory experiments (Barton & Charlesworth 1984; Rice & Hostert 1993; Barton 1996; Rundle *et al.* 1998; Mooers *et al.* 1999; Rundle 2003), but very few studies have been carried out in natural

systems to determine the role of drift on the divergence and speciation of small isolated populations (Clegg *et al.* 2002b; Knowles & Richards 2005).

Even if the ecological model of speciation can be generalised to speciation in allopatry, geographical isolation of populations offers at least two routes for the speciation process that are not available in situations that permit interactions between diverging populations. Both processes concern the divergence of mate recognition systems. It is widely accepted that pre-mating behavioural barriers can be the most important isolating mechanism in animals (Coyne & Orr 2004). This is well established in birds, where pre-mating isolation explains the high diversification in this group, whereas the much slower evolution of post-zygotic isolating mechanisms explains the retention of hybrid viability and fertility for millions of years after species formation (Grant & Grant 1992, 1997a; Price & Bouvier 2002; Edwards *et al.* 2005).

Song is one of the most obvious traits implicated in pre-mating reproductive isolation in birds – it is used in territory defence and mate choice, and plays a major role in species recognition (Catchpole & Slater 1995; Martens 1996; Podos *et al.* 2004). The exact role of song in speciation is still a matter of debate, but due to its role in mate recognition the potential of song divergence in driving assortative mating is high (Slabbekoorn & Smith 2002; Edwards *et al.* 2005). Because song is a learned behaviour in many birds (in the oscine passerines or songbirds), it can evolve rapidly by cultural evolution, through the process of sexual imprinting on mate preferences (Grant & Grant 1979; Price 1998; Irwin & Price 1999). This cultural aspect of song evolution provides a mechanism, known as cultural drift, for the evolution of reproductive isolation that is restricted to (or much more likely in) isolated populations. Cultural drift includes the differential sampling of song types in isolated populations, and the spread of stochastic errors or innovations within each population and, contrary to genetic drift, it is a common feature of song evolution (Catchpole & Slater 1995; Edwards *et al.* 2005).

Geographic isolation can create the conditions for another mechanism for the rapid evolution of reproductive isolation. Mate recognition traits are the result of co-evolution between female preferences and male signals (Lande 1981, 1982; Pomiankowski & Isawa 1998). Therefore they are expected to be in continuous evolution: female preference leads to the intensification of a given signal in the males up to the point where evolutionary constraints or natural selection makes further intensification impossible, or when the signal is so common that it is no longer informative (Fisherian runaway sexual selection: Pomiankowski & Isawa 1998). Another signal/preference trajectory is then followed. Because of this cyclic fluctuation of male signal and female preference, it will be very likely for allopatric populations to fall out of phase with each other even if they are subject to the same selection regime (Isawa & Pomiankowski 1995). The presence of two isolated populations in environments with exactly the same signal transmission/reception characteristics is unlikely, and therefore this process will be speeded up by adaptation of mate preferences and signals to the local environment. It has been shown that even small habitat differences can have large effects in this process termed as 'sensory drive' (Boughman 2002). This process will affect all traits implicated in mate recognition and not only learned behaviours.

To summarize, speciation in allopatry could differ from the ecological model of speciation in three aspects: i) genetic drift could play a role, and pre-mating isolating mechanisms could diverge rapidly by ii) cultural drift and iii) the 'stepping out-of phase mechanism' just described coupled with sensory drive. In the current consensus, an important role for genetic drift seems unlikely, whereas the two other possibilities are considered to have the potential to play a determinant role in triggering the evolution of reproductive isolation in isolated populations.

In this study, I describe the genetic and phenotypic patterns of variation of the Príncipe seedeater *Serinus rufobrunneus*, an endemic songbird of the Gulf of Guinea with populations on the islands of São Tomé, Príncipe and Boné de Jóquei (see 'Study species'), in order to identify the most likely processes involved in divergence of small isolated populations and to discuss how these could relate to the evolution of

reproductive isolation. The three populations show sufficient differentiation in morphology and colour to have been classified as different subspecies (Naurois 1975b), and therefore constitute a useful model for the study of speciation (Edwards *et al.* 2005). Phenotypic traits analysed here include adaptive morphological traits linked to diet and flight performance, and mate recognition traits (plumage colour and song). I investigate both the role of random processes on the evolution of allopatric populations and the potential for rapid divergence of mate recognition systems to be implicated in the first stages of reproductive isolation.

To assess the role of random processes, the patterns of neutral genetic variation (15 microsatellites) were compared with the patterns of phenotypic differentiation. By definition, neutral genetic variation measures the amount of change due to drift. If the variation in phenotypic traits is concordant with neutral genetic variation the role of drift in phenotypic evolution cannot be excluded, whereas a lack of concordance indicates that other processes, like selection, were involved (Barrowclough 1983; Clegg *et al.* 2002b). The role of selection was also inferred by assessing whether there was a positive relation between habitat and phenotypic differentiation. Three broad habitat types are considered: mature forest (in São Tomé and Príncipe), oil palm forest (in Boné), and plantations with shade forest (in São Tomé). The initial study design also included shade forest plantations on Príncipe, but no seedeaters could be sampled in this habitat (see 'Study species'). If divergence were associated with degree of isolation independently of habitat type, then drift would be a sufficient explanation for the observed pattern. Predictions for the pattern of phenotypic divergence under the effect of selection follow directly from the different assumptions on habitat similarity. In relation to diet traits I assumed that the oil palm forest of Boné constitutes the most distinct habitat since it offers very abundant but not diverse food resources, and that São Tomé and Príncipe forests are the most similar habitats. In relation to habitat characteristics the transmission and reception of mating signals, I assumed that São Tomé and Príncipe forests are the most similar and that they are close to the Boné habitat (closed environments), whereas the more open São Tomé plantations constitute the most distinct habitat. The effects of the characteristics of the different habitats in relation to flight performance are more

difficult to predict, but one can safely assume that Boné offers the most distinct selection pressures – not because of habitat structure, but because of the very small island size (c. 600 x 900 m), which 'removes' the need for flight (the opposite prediction would be true if birds are adapted to cross the 3 km sea gap that separates this population from Príncipe).

This study design presents a bias favouring drift over selection, in that the shade forest plantations are very recent (c. 300 years) whereas the mature forests were present even before the origin of the seedeater. Therefore populations from the forest of São Tomé and Príncipe have had much more time to diverge than populations from São Tomé forests and plantations. Indeed, most of the evolutionary history of these latter populations must have happened in the forests. For this reason, this sampling design cannot be considered as equivalent to the ecotone-isolation framework where divergence is compared between allopatric and parapatric populations of the same age (Moritz *et al.* 2000). If divergence were greater between São Tomé forests and plantations than between Príncipe and São Tomé forests one could conclude that ecology is much more important than drift, but nothing can be concluded from the opposite result. This limitation can nevertheless be used to address the capacity of species to adapt to recent human made changes and to determine which, if any, phenotypic traits diverge first. A higher divergence of mate recognition traits would support the hypotheses that these traits can evolve faster than other phenotypic traits due to the processes of cultural evolution and sensory drive, and that they can play a role in the first steps of the speciation process.

METHODS

Study species

The Príncipe seedeater is a songbird endemic to the islands of São Tomé and Príncipe and to the small islet Boné de Jóquei (Jockey's Cap; hereafter: Boné), 3 km offshore from Príncipe. Each insular population has been described as a distinct subspecies on the basis of morphological and colour differences (Naurois 1975b).

From a dietary perspective the seedeater can be considered a generalist: it feeds on a wide range of fruits and seeds, but also on insects (Naurois 1975b; Christy & Clarke 1998; Jones & Tye 2006) and even pollen (pers. obs.). Fruits range from buds and small berries to the pulp of exotic fruits such as pawpaw *Carica papaya*. Correspondingly, the seedeater displays a large range of feeding techniques, which include probing the bark of trees in a woodpecker-like fashion and gleaning insects from twigs, leaves and rocks on streams (pers. obs., Jones & Tye 2006). Detailed observations will be required to assess whether this generalist foraging occurs at the individual level or if it results from the co-existence of individuals specialized in different food resources and/or feeding techniques (e.g., Grant 1986; Werner & Sherry 1987; Scott et al. 2003).

The ability to explore different resources has made the seedeater one of the most common and widespread species on São Tomé, where it breeds from the capital city at sea level up to the peak at 2024m. This pattern is not replicated on Príncipe where the species is mainly restricted to the southern mature forests (including the coastal fringe which still has coconut and oil palms planted during colonial times), but is rare in the northern cultivated areas. Only once during the course of the fieldwork for this thesis was a group of seedeaters observed in the north, in the palm oil plantations of the Sundi estate. This pattern remains inexplicable (Christy & Clarke 1998; Jones & Tye 2006). Many farmers are not familiar with the species, and others said that it only appears seasonally when trees are in flower or fruiting. On Boné, the seedeater is extremely abundant, with birds occurring in groups of c. 10 birds all over the islet (Fig. 6.1). Densities estimates range from 30-40 ha^{-1} (Naurois 1975b) to 60 ha^{-1} (J. Baillie in litt.). These extraordinary densities are sustained mostly by the oil palm *Elaeis guineensis* forest, the dominant vegetation cover. Seedeaters eat the pollen from the male inflorescences and the oil from the palm nuts (which they obtain by repeatedly hammering the hard nut). Interestingly, the palms on Boné produce cones with very large nuts (Fig. 6.1). This palm variety does not occur on the islands of São Tomé and Príncipe, and is apparently unique (William Baker, Kew Botanic Gardens, pers. comm.).

Fig 6.1. Boné de Jóquei Islet (1) is only about 600 x 900 m and lies 3 km off the southeast coast of Príncipe (2). It holds an endemic population of the Príncipe seedeater *Serinus rufobrunneus* (3), which occurs at very high densities, with birds forming large groups (4). On the islet, the seedeaters thrive on the resources provided by the oil palm *Elaeeis guineensis*, which constitutes the main vegetation cover. They feed both on the pollen of the male inflorescences (5) and on the oil held by the palm nuts (6, 7). The oil palms of Boné produce 'gigantic' palm nuts that appear to have no parallel worldwide - and are not found on neighbouring Príncipe. (6): large nuts from Boné in comparison with 'typical' nuts; 15 cm ruler shown. In picture 3, the collar (one of the feather areas sampled) can be clearly seen. Despite the proximity of Boné to Príncipe, gene flow is limited (this chapter).

Field sampling

Sampling took place during three field-seasons in the Gulf of Guinea region: October 2002 – March 2003, November 2003 – March 2004, and December 2004 – February 2005. The initial sampling design was aimed at measuring genetic and phenotypic variation between the three allopatric populations and between parapatric populations on São Tomé and Príncipe, with the objective of contrasting the relative importance of isolation and divergent selection across habitats in promoting diversification. Because seedeaters on Príncipe were apparently restricted to the forest habitat, parapatric sampling could be done only on São Tomé. Therefore the sampling groups used for this study were: 1) São Tomé forest, 2) São Tomé plantation, 3) Príncipe and 4) Boné.

Birds were caught with mist-nets, ringed on the right leg with individually numbered aluminium rings (AFRING, Avian Demography Unit, University of Cape Town), weighed, measured, and blood-sampled; feather samples were taken from a sub-sample of the individuals. In order to be able to detect dispersal between Príncipe and Boné populations (only 3 km apart), birds from Príncipe were also ringed with a red plastic ring on the left leg, and birds from Boné with either white, or yellow or orange colour-rings on the left leg. In the second and third field seasons, birds from Boné were individually colour ringed for future studies. The individual colour combination was put on the left leg, and the population colour was put on the right leg, over the metal ring. A total of 217 individuals were sampled in 12 sampling sites (Table 6.1, Fig. 6.2). Blood samples (50-100 µl) were taken from the brachial vein and stored in absolute ethanol. Mass was measured to the nearest 0.5 g with a 50-g Pesola spring balance. Wing and tail length were measured with a standard wing ruler to the nearest 0.5 mm. Bill and tarsus measurements were taken with a digital calliper to the nearest 0.1 mm. Measurements were taken as follows: wing length (flattened), from the carpal joint to the tip of the longest primary; tail length, from the uropygial gland to tip of the central rectrix; tarsus length, from the tibiotarsus joint to the distal end of the tarsometatarsus when the foot is held to the leg; upper mandible length, from when the culmen enters the feathers of the head to tip; bill

width and depth (height) at the anterior end of nares. Bill depth was not measured in the first field season. Feather samples for colour analyses were taken from 20 birds in each population from four body areas (head, collar, breast, and mantle) and kept separately in paper envelopes. The collar is a light feather patch on the throat (picture 3, Fig. 6.1.).

Song recordings were made at the following sites: Lagoa Amélia (São Tomé forest), Bom Sucesso site (São Tomé plantation), Rio Porco and Praia Cará (Príncipe), and Boné. Songs were recorded with a Marantz PMD222 tape recorder with Type II 60 min tapes and a Sennheiser ME66K6 directional microphone.

Table 6.1.

Sampling locations and sample sizes. M: males; F: females; T_S: Total per sampling site; T_C: Total per group used in analyses.

Island	Habitat	Site	Locality	Coordinates	Altitude m	*Sample size*			
						M	F	$T_.$	$T_.$
São Tomé	Forest	1	Lagoa Amélia	0°16'N 6°35'E	1487	11	9	20	**51**
		2	Pico Calvário	0°16'N 6°34'E	1587	5	2	7	
		3	Umbumgu	0°09'N 6°33'E	300	2	2	4	
		4	Quija	0°09'N 6°33'E	658	11	7	18	
		5	São Miguel	0°10'N 6°30'E	400	1	1	2	
	Plantation	6	Bom Sucesso	0°17'N 6°36'E	1158	10	9	19	**62**
		7	Monte Café	0°18'N 6°38'E	723	3	1	4	
		8	Contador	0°19'N 6°36'E	600	10	5	15	
		9	Alto Douro	0°12'N 6°41'E	150	13	11	24	
Príncipe	Forest	1	Rio Porco	1°32'N 7°22'E	128	4	0	4	**37**
		2	Praia Cará	1°33'N 7°21'E	27	27	6	33	
Boné	Oil Palms	1	Boné	1°30'N 7°25'E	5	41	26	67	**67**

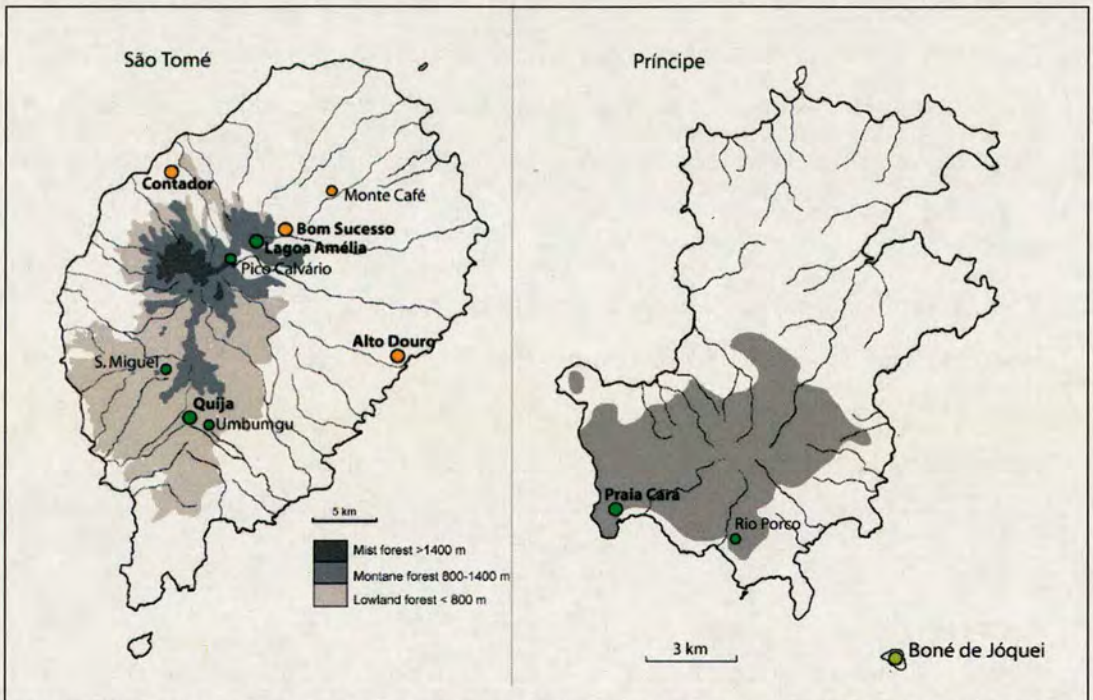


Fig 6.2. Sampling locations. Forest sites in dark green; plantation sites in orange; oil palm 'monoculture' of Boné in light green. Big dots indicate the main sampling sites. Maps adapted from Jones & Tye (2006).

Laboratory protocols

DNA extraction and sexing

Total genomic DNA was extracted from the blood using a DNeasy Tissue extraction kit (Qiagen), following the protocol for DNA extraction from animal tissues. Adaptation of this protocol for blood samples stored in absolute ethanol is described in Chapter 2. As the Príncipe seedeater shows no sexual dimorphism, birds were sexed with a molecular protocol (Griffiths *et al.* 1998), described on Chapter 2.

Microsatellite identification

Rather than isolating microsatellites in the Príncipe seedeater, 74 primer pairs developed for other passerine species were tested for cross-amplification (Melo & Hansson 2006). Because the Príncipe seedeater is quite divergent from the 'typical' *Serinus* species, being often placed in the monotypic genus *Poliospiza* (Jones & Tye 2006), the utility of the primers was also tested on 9 other *Serinus* species. The genus

Serinus, which includes the canaries, has several species used as models in evolutionary biology, and in particular in the study of communication and learning (Brainard & Doupe 2002). Therefore identifying microsatellites that work in this group can be useful for a wider research community. The BIRDMARKER database was used (http://www.shef.ac.uk/misc/groups/molecol/birdmarkers.html) to choose primers developed in closely related groups and/or those that have been shown to be successful in cross-amplifications.

Initially the primers were tested on three Príncipe seedeaters, one from each island. The PCR conditions are detailed in Appendix 6.1. The fluorescent-labelled PCR products were separated and alleles were detected in an ABI PRISM 3730 capillary sequencer (Applied Biosystems). Of the 74 primer pairs tested in the three test individuals, 21 were deemed as potentially useful since they had an apparently polymorphic product of expected size (in some cases additional unspecific products occurred). The remaining loci did not amplify any product (10 loci), or amplified multiple, complex and sometimes weak products (14), or had a single clear product of expected size (26) or had a weak but potentially polymorphic product of expected size and other much stronger unspecific products (3). The 21 apparently polymorphic loci were further tested in 96 individuals, and it became evident that one locus (Lswμ7) was monomorphic and that it was difficult to achieve a scoreable product for five others (Cuμ02, Gf02, LOX2, PmaD22 and Pma105). In total, 15 useful loci were identified and these were scored in another 17 individuals from São Tomé and tested on 33 individuals from nine other species from the genus *Serinus*. Gf10 is known to be Z-linked (Grant *et al.* 2001) and this was confirmed in this dataset as all females were 'homozygous' (hemizygous) for this locus. Results for all primers (including those that were monomorphic and failed to amplify) and species have been deposited in the BIRDMARKER database.

Microsatellite analysis

Descriptive statistics

The observed allele frequency and the allelic richness at each locus and in each population were calculated in FSTAT version 2.9.3.2 (Goudet 2002). The allelic

richness is a measure of allele diversity that accounts for differences in sample sizes through the use of rarefaction procedures (e.g., Petit *et al.* 1998). This is important in order to be able to compare genetic diversities of populations that were not equally sampled. The observed heterozygosity (H_O) and the expected heterozygosity (H_E = gene diversity: Nei 1987) were estimated in ARLEQUIN version 3 (Excoffier *et al.* 2005). For each locus and population, deviations from the Hardy-Weinberg equilibrium (HWE) were tested in FSTAT by assessing the significance of the deviation of F_{IS} from zero with 1000 permutations. F_{IS}, the inbreeding coefficient, describes the reduction in heterozygosity of an individual in relation to its population, and should be close to zero in randomly mating populations (Lowe *et al.* 2004). Departures from HWE are often in the direction of an excess of homozygotes, which in the case of microsatellites could indicate the presence of null alleles rather than a biologically meaningful departure (i.e., selection, inbreeding, non-random mating, or hidden population structure). The frequency of null alleles was estimated with the expectation maximization algorithm (EM; Dempster *et al.* 1977), as implemented in FREENA (Chapuis & Estoup in press). To test for non-random associations between loci, pairwise tests of linkage equilibrium were performed in ARLEQUIN. To control for type I error a sequential Bonferonni correction was applied (Rice 1989).

Population differentiation and gene flow

Population differentiation was estimated using several approaches: i) differentiation estimators based on F-statistics (Wright 1951) and their analogue developed for microsatellite data, R-statistics (Slatkin 1995); ii) analysis of molecular variance (AMOVA: Excoffier *et al.* 1992); iii) exact G-test of population differentiation; iv) a model-based clustering approach to identify population structure by assigning individuals to groups in Hardy-Weinberg and linkage equilibrium (Pritchard *et al.* 2000). The Z-linked locus was not included in the first three approaches. Gene flow between populations was estimated with the coalescent-based approach implemented in MIGRATE version 2.1.3 (Beerli & Felsenstein 1999, 2001). The assignment approach (iv) was also used to detect recent immigration events.

F-statistics describe the structuring of genetic variation at three hierarchical levels: the individual, the local population and the total population (Wright 1951). The degree of structuring at each level is identified by the extent of the deviation from the HWE. Genetic structuring at the population level is described by F_{ST}, the fixation index, which measures inbreeding within populations relative to all populations. F_{ST} assumes both that mutation follows the infinite allele model (IAM: Kimura & Crow 1964) and that mutation is negligible when compared to migration. These assumptions might not be met for microsatellites. In order to address this issue, Slatkin (1995) proposed the R_{ST} statistic based on the stepwise mutation model (SMM: Kimura & Otha 1978) thought to represent microsatellite evolution better. The most commonly used and least biased F_{ST} estimator is theta (θ: Weir & Cockerham 1984) based on an analysis of variance of allele frequencies. The analogue estimator for R_{ST} is rho (ρ: Michalakis & Excoffier 1996; Rousset 1996) based on the variance in repeat number. For the same dataset, the two estimators can differ markedly. Because the mutation dynamics of microsatellites are not well understood and vary between loci (Ellegren 2004), it is not clear which of the approaches is more suitable for microsatellite data (reviewed in: Gagiotti *et al.* 1999; Balloux & Goudet 2002; Balloux & Lugon-Moulin 2002). F_{ST} is sensitive to high mutation rates and will be deflated when very variable loci are used: even if two populations do not share alleles (highest level of structuring), the within-population homozygosity that is used in F_{ST} estimation will remain small and will result in small F_{ST} estimates (Hedrick 1999). R_{ST} is independent of mutation rate but only when microsatellite evolution follows a strict SMM model, which may only rarely be the case. Nevertheless, even if the SMM is not strictly followed, its assumption that the microsatellite mutation process has some memory (i.e., changes between alleles are not random) is likely to hold in most cases. Therefore, in theory R_{ST} should better represent microsatellite evolution, and it should be superior to F_{ST} when population differentiation is high (mutation >> migration). In practice, R_{ST} estimates tend to suffer from high variances, which can make F_{ST} estimates more accurate even in cases that follow a strict SMM model. Therefore, F_{ST} remains the most commonly used statistic for describing population structure even for microsatellite data. In this study both overall and pairwise θ and ρ were estimated in GENEPOP version 3.4

(Raymond & Rousset 1995). With the sample size (n = 217) and number of loci used (n = 15), ρ could better represent structuring between the three island populations (high differentiation), whereas θ could be more appropriate to assess structuring within São Tomé (low differentiation). In order to assess whether the presence of null alleles could constitute a problem in this dataset, θ was also estimated after correcting for the estimated frequency of null alleles in FREENA.

The analysis of molecular variance developed by Excoffier *et al.* (1992) uses a matrix of squared genetic distances among all haplotype pairs to estimate how the total genetic variance is distributed among-groups, among-populations, and within-populations. It estimates ϕ-statistics that are analogous to F-statistics, but have the advantage of not being dependent on any particular assumptions about the evolutionary dynamics of the loci used. ϕ-statistics were estimated in ARLEQUIN and their significance levels tested by non-parametric permutation of the data (10,000 replicates). The hierarchical design used was three groups (São Tomé, Príncipe, Boné) and two populations within São Tomé (forest and plantation).

Exact tests of population differentiation are independent of the method used for inference of population structure (Balloux & Lugon-Moulin 2002). Differences in allele frequencies among population pairs were tested in FSTAT with the exact G-test, considered the most powerful, particularly when samples are unbalanced as in this study (Goudet *et al.* 1996).

All the methods discussed above require the *a priori* definition of populations. As an alternative, the model-based clustering method implemented in STRUCTURE version 2 (Pritchard *et al.* 2000) was used. This method uses a Bayesian approach to infer the population of origin of each individual, the allele frequencies in all populations and the admixture proportions for each individual. The main assumptions of the implemented model are that the populations are in HWE and that loci are at linkage equilibrium between populations (but when linkage information is available it can be incorporated: Falush *et al.* 2003). For a given number of populations, *K,* individuals are assigned to each population (or several populations if admixed) in a way that

deviations from HW and linkage equilibrium are minimized within each population. STRUCTURE approximates the probability of the genotypic data (X) given the number of populations considered as $\ln P(X|K)$. From this value, the posterior probability of K, $P(K|X)$, can be computed. To determine the most likely clustering of populations, $P(K|X)$ was estimated for a range of K from 1 to 6 populations. The correlated allele frequency model was used to estimate K, as it performs better when there are mild departures from the assumptions of HW and linkage equilibrium. In order to guarantee that the parameter space was properly explored, replicate runs for each K were carried for different size runs, until the appropriate run length was found, i.e., the one for which $\ln P(X|K)$ were consistent between replicated runs. The appropriate run length consisted of 200,000 burn-in iterations followed by 10^6 iterations. After the best K was chosen, STRUCTURE was run again by making use of the information on the origin of the sampled individuals. This approach is useful to detect recent migration events by identifying individuals that are immigrants or that have immigrant ancestors. Because strong population differentiation (like the one expected between Príncipe and São Tomé) can mask more subtle levels of genetic structure, the São Tomé sample was also analysed separately. All 15 loci were included in the STRUCTURE analyses as the program can deal simultaneously with autosomal and sex-linked loci.

The coalescent-based method of Beerli & Felsenstein (1999, 2001), implemented in MIGRATE-N 2.1.3 (Beerli 2004), was used to estimate the effective population size and immigration rate for each population relative to the mutation rate. Parameters were estimated with the maximum-likelihood algorithm using a Markov chain Monte Carlo (MCMC) to integrate over all possible genealogies. The analysis was performed with 11 out of the 15 loci. Loci excluded were: Gf10 (Z-linked locus), Lswμ14 (20% missing data and null alleles frequencies > 10%), WBSW7 (monomorphic in Príncipe and Boné) and Gf08 (only three alleles). The Brownian motion model was used to describe microsatellite loci evolution. The initial starting tree was built from a distance matrix between the samples using the UPGMA algorithm. Due to CPU time constraints, the gamma model to describe variation of mutation rate of the different loci was not implemented. Even if a stepping-stone

model is more likely between the three islands this was not imposed on the search – rather a full migration model was implemented to check if the stepping-stone model was recovered by MIGRATE-N. An initial search was performed with the default parameters (10 short chains, genealogies sampled 20 steps apart, 500 genealogies sampled; 3 long chains, genealogies sampled 20 steps apart, 5000 genealogies sampled) and initial effective population size and migration parameters estimated by MIGRATE-N with a F_{ST} based measure. The effective population size and migration estimates of this run were then used as the initial values for a longer search: 10 short chains, genealogies sampled 20 steps apart, 2,500 genealogies sampled; 3 long chains, genealogies sampled 20 steps apart, 25,000 genealogies sampled. A heating scheme with 4 chains was implemented (temperatures: 1, 1.5, 3, 6). Profile likelihoods, used for the estimation of the 95% confidence intervals, were calculated using the 'fast' method.

Morphological analysis

In order to detect statistically any morphological differences between populations a multivariate analysis of variance (MANOVA) was performed on the seven morphological traits. Prior to analysis the data were log-transformed in order to reduce the influence of possible multivariate outliers (Quinn & Keough 2002). Normality and the presence of outliers in the transformed data were checked by visual inspection. Initially, three explanatory variables were used: sex, year, and population (São Tomé forest, São Tomé plantation, Príncipe, Boné). The factor 'year' was included to control for possible variations in measurements between years. There were significant differences between sexes for wing, tail, bill width and depth. Therefore subsequent analyses were carried on males only, for which the sample size was larger (130 males, 72 females). Post-hoc Tukey tests were conducted to determine which populations differed significantly for each variable. To visualize the morphological differences between populations, a principal component analysis (PCA) was conducted on the log-transformed data, with the components extracted from a covariance matrix. All statistical analyses were performed in SAS version 8.

Colour analysis

Colour is not a property of objects but a product of the light environment, the sensitivity of the receiver and of reflectance (Endler 1990; Andersson & Prager 2006). Reflectance on the other hand is a physical property of objects. Reflectance describes the ratio of reflected to incident light for a given object and measuring geometry (Andersson & Prager 2006): it is independent of light conditions (context-independent) and of the receiver (objective). It is inherent to each object, and should therefore not be confused with 'reflection', the optical product of reflectance dependent on incident light and sensitivity of receiver.

Colour measurements were conducted with an AVASPEC-2048 spectrometer (Avantes, Inc.) using a 200 μm fiber-optic probe at a 90° angle to the feather surface. Ambient light was excluded with a rectangular metal sheath affixed to the probe tip. The sheath was placed against the feather with the probe held at a fixed distance of 2 mm from the feather surface. The reading area was illuminated with both a deuterium bulb (UV light source) and a tungsten-halogen bulb (visible light source) with the AVALIGHT-DH-S light source. We generated reflectance data relative to a white standard and the dark current (black felt background). Each spectrum was averaged from 5 scans of 40 ms stabilized for maximal chroma and was analysed both with AVICOL version 1.0 (Doris Gomez) and PROJECT version 1.4 (Jean Marc Rossi).

A total of 80 individuals were sampled (20 from each population) for four colour patches: head, breast, mantle and collar. The collar consists in a patch of lighter feathers under the chin area (Fig. 6.1, photo 3). Measurements from each patch were carried out on five superimposed feathers to simulate the thickness of bird plumage. Whenever sufficient feathers were available, measurements were replicated in two 5-feather groups. Additionally for each replicate, three measurements were taken at different points of the patch and a repeatability analysis conducted following the method of Lessells & Boag (1987). Reflectance variation was converted to 'colour' (i.e., the perception of the receiver) by estimating the three general dimensions of vertebrate colour cognition (Box 6.1; Andersson & Prager 2006): spectral intensity

(brightness), spectral location (hue), and spectral purity (chroma). Differences in colour between populations were tested with MANOVA implemented in SAS. Only the samples from São Tomé had a sex ratio that allowed testing for colour differences with sex. No significant differences were detected, and therefore samples from both sexes were analysed together.

Box 6.1. Colorimetrics extracted from reflectance measurements. Adapted from Andersson & Prager (2006). *R*: reflectance; λ: wavelength.

Colorimetric	Definition	Reflectance property	Human perception
R_{sum}	*R* summed over λ interval	Spectral intensity	Brightness
R_{avg}	*R* averaged over λ interval		
λ_{R50}	λ halfway between R_{max} and R_{min}	Spectral location	Hue
C_{max}	(R_{max} and R_{min}) / R_{avg}	Spectral purity	Chroma

Song analysis

We sampled the song recordings on a PC with 16 bit accuracy and a sampling rate of 20,500 Hz using AVISOFT-SASLAB PRO version 4.3 (R. Specht, Berlin). Spectrograms were analysed with a fast Fourier transformation-size of 256 points, a 50% frame size and a temporal resolution overlap of 50% (Hamming window type), giving a frequency resolution of 86 Hz and a temporal resolution of 4.5 ms. Recordings that could be used came from 36 birds from São Tomé forest, 24 from São Tomé plantation, 51 from Príncipe, and 42 from Boné. The total number of songs recorded in each population was respectively: 242, 207 (total of 449 for São Tomé), 529, and 220. To describe variation in song between the different populations, sonograms were sorted by eye into different categories, the most sensitive method for this purpose (Nelson & Soha 2004). In a first stage, categorization of a subset of the songs was made independently by two persons (Claire Doutrelant and myself). Results were compared and a categorization scheme was developed (Box 6.2) and then applied to all 1198 songs by the same observer (myself). Song complexity was high in the Príncipe seedeater, with 116 song types described. Song diversity in each population was described by the Simpson diversity index (Magurran 1988), the

complement of the Simpson index, which gives the probability that songs randomly selected from a population will be different (0: no diversity; 1: highest diversity). The high song diversity was in part due to the variety of start and end notes framing the main theme of any given song, which made song sharing to be quite low even within populations. Therefore analysis of song differentiation was carried at the level of the main themes and songs restricted to a single individual were excluded. This allowed focusing on population differentiation rather than on individual variation.

To describe song differentiation between populations, a Bray-Curtis dissimilarity matrix (Bray & Curtis 1957) was calculated and used as the input for non-metric multi-dimensional scaling (MDS), a method that makes few assumptions about the distribution of the data (Cox & Cox 1994). The Bray-Curtis measure is the most appropriate to deal with data matrices containing many zeros as it ignores joint absences (such as species abundance data, or in this case song type abundance data: Clarke & Warwick 2001; Quinn & Keough 2002). Before estimation of dissimilarities, data were standardised to correct for the difference in sample sizes (as the duration of each recording per individual was different) and square root transformed in order to down-weight the importance of very abundant song types. Significance of any observed differentiation between populations was tested with the non-parametric Analysis of Similarities (ANOSIM). This is the appropriate analysis for 'species abundance' matrices, which violate most assumptions for the application of MANOVA (e.g., the high frequency of zero values makes it impossible to approximate the data to a normal distribution). The test statistic R compares the rank similarities between populations with the similarities within populations, and should be 0 when there is no differentiation and 1 for the maximum level of differentiation. All these analyses were performed in PRIMER version 5 (Primer-E Ltd).

Comparison of neutral genetic and phenotypic differentiation

Neutral genetic markers primarily reflect the effects of drift so that when selection is the major force driving divergence there will be no correspondence between patterns of molecular and phenotypic variation (Barrowclough 1983; Reed & Frankham 2001). Mantel tests were performed in FSTAT to test for the concordance between

genetic and phenotypic differentiation. Both pairwise F_{ST} and R_{ST} were used as genetic distances. For morphological and colour measurements, population pairwise Euclidean distances were calculated in PRIMER for the mean of each variable per population (of the log-transformed data). For the song data, a Bray-Curtis dissimilarity matrix of standardised and square root transformed data was used. It should be noted that the song distance reflects only the number of song types shared and does not hold information on acoustic distances (i.e., song duration, frequencies, number of syllables per song, etc). To assess correspondence further between phenotypic traits and neutral genetic variation, the distance matrices were used to build neighbour-joining trees (NJ) in PAUP* version 4.0b10 (Swofford 2003). Differences in topology and/or branch lengths would be indicative that non-random processes are involved in divergence.

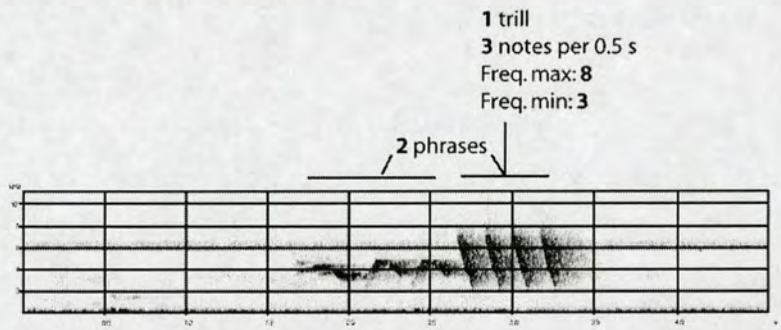
Box 6.2. Song classification scheme

Definitions
Element or note: a continuous trace on the sonogram
Trill: > 3 repeated identical elements
Phrase: cluster of elements separated by a time interval shorter than element lengths.
Song: cluster of phrases (or single phrase) separated from other songs by at least 1 second.
Main theme: trill or when no trill is present, the central phrase of the song.

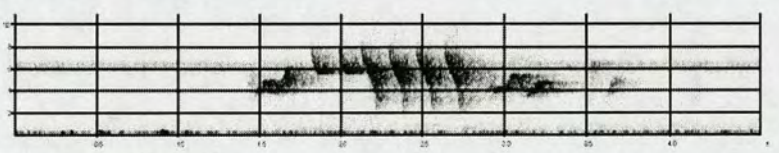
Classification 'algorithm'
Step 1. Number of trills.
Step 2. Number of notes of the main theme per 0.5 seconds.
Step 3. Frequency amplitude of the notes of the main theme.
Step 4. Shape: clear differences in note shape are identified by the last letters of the alphabet after the amplitude code.
Step 5. Number of phrases. In the case of trills, introductory and/or end notes are only quantified as additional phrases if their duration exceeds 0.5 seconds. Consistent patterns in phrase syntax were identified by the initial letters of the alphabet at the end of the code.

Song codification follows the steps of the algorithm. Such coding makes it easy to select the level of classification used in the analyses. In this study it was decided to perform the analysis of song differentiation based on the main theme of the songs, i.e., excluding start and end phrases or notes.

In the example below the first two songs were put in the same category because the main theme is the same. The main theme of the third song is the same except for the shape of the notes so it was put in a separate category.



song code: 1.3.83.2
song code that separates shape of trill notes and of extra phrase: **1.3.83x.2a**



song code: 1.3.83x.2b

song code: 1.3.83y.3

RESULTS

Microsatellite analysis

Descriptive statistics

The characteristics of the 15 polymorphic loci for the three island populations are summarised in Table 6.2 (separate results for São Tomé forest and São Tomé plantation are not shown for simplicity, as there was no differentiation between these populations – see below). The number of alleles ranged between three (Gf08 and Gf10) and 76 (LOX7), and the expected heterozygosity between zero (Cuμ04 and Gf08 in Boné and WSBW7 in Boné and Príncipe) and 0.98 (LOX3 and LOX7 in São Tomé). The highly variable loci from the Scottish crossbill *Loxia scotica* (Piertney *et al.* 1998) were also extremely variable in the Príncipe seedeater (40-76 alleles), which may be explained by the close relationship between these genera (Yuri & Mindell 2002). The genetic diversity (as measured by allelic richness corrected for sampling size) decreased with decreasing population sizes (São Tomé: largest population; Boné: smallest; Fig. 6.3).

In São Tomé, three loci (Ase 43, Lswμ14, Lswμ18) departed from Hardy-Weinberg equilibrium (HWE) after Bonferroni correction. In Boné another three loci (LOX1, LOX7, LOX8) departed from HWE. In Príncipe, only LOX8 departed from HWE. All departures were in the direction of a deficit of heterozygotes, and corresponded to null allele frequency estimates over 10% (maximum of 15%), except for Ase43 (7%). Only Lswμ14 had relatively high null allele frequencies in all populations (7-14%), suggesting that null alleles might constitute a problem with this locus – something supported by the fact that it was the locus with the lowest amplification success (174 products out of 217 reactions). Null allele frequencies were low for the other loci/populations, and overall the potential presence of null alleles did not induce biases in population differentiation estimates (analyses carried out with different loci sets; F_{ST} estimates corrected for null alleles in FREENA; results not shown). Tests of linkage equilibrium between all pairs of autosomal loci (i.e. excluding Gf10) revealed a single significant deviation after Bonferroni correction for São Tomé (between Lswμ14 and Lswμ18), two for Príncipe (Cuμ28 and LOX7,

Cuμ28 and LOX8), and four for Boné (Ase42 and Lswμ14, Cuμ28 and Lswμ18, LOX8 and Lswμ18, Lswμ18 and Pdoμ4). Because no instance of linkage disequilibrium (LD) was shared across populations, physical linkage of these loci can be dismissed, whereas the fact that LD cases increase with decreasing population size would suggest that drift might be involved. Overall LD within populations was weak in this set of mainly unlinked loci, and therefore loci were considered independent for the analyses.

All primers successfully amplified in nine other *Serinus* species, and in a sample of 13 *S. flaviventris* all loci were in fact polymorphic (Appendix 6.2). Senar *et al.* (2006) found another six primers that amplified polymorphic products in *Serinus citrinella*, and together these two panels of microsatellites constitute a valuable resource for future research in the genus *Serinus*.

Population differentiation

All methods gave the same results: there is strong genetic structure between islands but not between habitats within São Tomé. Analysis with the program STRUCTURE gave strong support for the structuring of genetic variation into three populations corresponding to the three oceanic islands (Table 6.3 & Fig. 6.4). A separate analysis for the São Tomé data did not detect any structure. Pairwise θ and ρ were high for between-islands comparisons, but close to zero between São Tomé habitats (Table 6.4). Accordingly, pairwise ϕ_{ST} values between islands were highly significant, but were non-significant between São Tomé habitats (permutation test). Under the MANOVA approach implemented in ARLEQUIN, 18.28% of the variation was distributed among islands and only 0.29% between São Tomé habitats. Except for the comparison Príncipe-Boné, ρ estimates were always higher than θ indicating that the stepwise mutation model might be more appropriate for this dataset (Hardy *et al.* 2003). Independently of the estimator used, fixation indices between São Tomé and Príncipe (150 km apart) were very similar to the values between Príncipe and Boné, only 3 km apart (Table 6.4). Finally, the three island populations were significantly different, but the difference between São Tomé habitats was non-significant (G test: significant P = 0.008, non-significant P = 0.017).

Table 6.2.

Characteristics of the 15 polymorphic passerine microsatellite loci used in the population genetic study of the Príncipe seedeater.

Locus	Population	*n*	*A*	H_O	H_E	HW P	F_{IS}	Null alleles	Allele size range
Ase42 7/208	São Tomé	111	7	0.811	0.736	0.987	-0.102	0.000	373-387
	Príncipe	34	2	0.412	0.331	1.000	-0.245	0.000	375-377
	Boné	63	3	0.143	0.176	0.363	0.100	0.051	375-379
Ase43 12/213	São Tomé	110	12	0.685	0.825	**0.001**	0.185	0.072	234-262
	Príncipe	37	5	0.811	0.743	0.867	-0.091	0.000	238-246
	Boné	66	4	0.576	0.698	0.025	0.175	0.081	240-246
Ase48 17/217	São Tomé	113	13	0.788	0.882	0.002	0.107	0.041	272-330
	Príncipe	37	10	0.730	0.687	0.863	-0.063	0.000	272-344
	Boné	67	6	0.701	0.729	0.333	0.038	0.006	272-350
Cuμ04 7/216	São Tomé	113	7	0.717	0.707	0.636	-0.014	0.000	126-140
	Príncipe	36	2	0.111	0.156	0.191	0.286	0.067	126-128
	Boné	67	1	0.000	0.000	na	na	0.001	128
Cuμ28 13/214	São Tomé	112	13	0.857	0.880	0.260	0.025	0.006	176-202
	Príncipe	36	8	0.694	0.802	0.071	0.134	0.041	184-198
	Boné	66	5	0.833	0.722	0.993	-0.155	0.000	180-194
Gf08 3/213	São Tomé	110	3	0.173	0.231	0.010	0.253	0.070	94-97
	Príncipe	37	2	0.027	0.027	1.000	0.000	0.000	94-95
	Boné	66	1	0.000	0.000	na	na	0.001	95
Gf10^{z} 3/216	São Tomé	66	3	0.318	0.343	0.500	0.080	0.023	204-210
	Príncipe	31	2	0.484	0.398	1.000	-0.200	0.000	204-210
	Boné	41	2	0.049	0.048	1.000	-0.013	0.000	204-210
LOX1 40/217	São Tomé	113	35	0.956	0.950	0.691	-0.006	0.000	286-446
	Príncipe	37	15	0.919	0.886	0.807	-0.037	0.000	314-416
	Boné	67	5	0.269	0.444	**0.001**	0.395	**0.147**	344-360
LOX3 74/215	São Tomé	111	71	0.991	0.979	0.257	-0.012	0.010	226-432
	Príncipe	37	20	0.838	0.931	0.045	0.100	0.032	258-410
	Boné	67	10	0.836	0.819	0.707	-0.021	0.018	234-334
LOX7 76/211	São Tomé	112	59	0.982	0.979	0.695	-0.003	0.000	278-436
	Príncipe	37	33	0.973	0.972	0.674	-0.001	0.000	294-432
	Boné	62	30	0.661	0.919	**0.001**	0.280	**0.113**	314-448
LOX8 71/214	São Tomé	112	62	0.920	0.968	0.005	0.060	0.013	350-540
	Príncipe	36	20	0.583	0.907	**0.001**	0.357	**0.127**	456-572
	Boné	66	9	0.424	0.616	**0.001**	0.311	**0.130**	478-566
Lswμ14 7/174	São Tomé	92	7	0.326	0.522	**0.001**	0.375	**0.132**	200-214
	Príncipe	37	4	0.216	0.313	0.023	0.309	0.092	204-212
	Boné	45	3	0.156	0.257	0.013	0.394	0.073	206-214
Lswμ18 12/214	São Tomé	111	8	0.509	0.687	**0.001**	0.266	0.092	208-222
	Príncipe	37	9	0.730	0.846	0.050	0.138	0.039	210-230
	Boné	66	2	0.015	0.015	1.000	0.000	0.000	222-224
Pdoμ4 40/216	São Tomé	113	28	0.770	0.819	0.077	0.060	0.000	198-338
	Príncipe	37	22	0.919	0.951	0.283	0.034	0.000	202-320
	Boné	66	9	0.758	0.824	0.114	0.081	0.033	268-328
WBSW7 4/216	São Tomé	113	4	0.566	0.549	0.708	-0.032	0.000	140-146
	Príncipe	37	1	0.000	0.000	na	na	0.001	142
	Boné	66	1	0.000	0.000	na	na	0.001	142

n, number of successful amplifications out of a total of 113 (ST), 37 (P), and 67 (B); *A*, number of alleles; H_O and H_E, observed and expected heterozygosity; HW *P*, *P* value of the Hardy-Weinberg test as implemented in FSTAT (Goudet 2001), adjusted nominal level = 0.001, departures from HWE in bold; Null alleles estimated in FreeNA (Chapuis & Estoup, in press), values > 10% in bold. ^{z}: Z-linked locus: only males were used for diversity estimates.

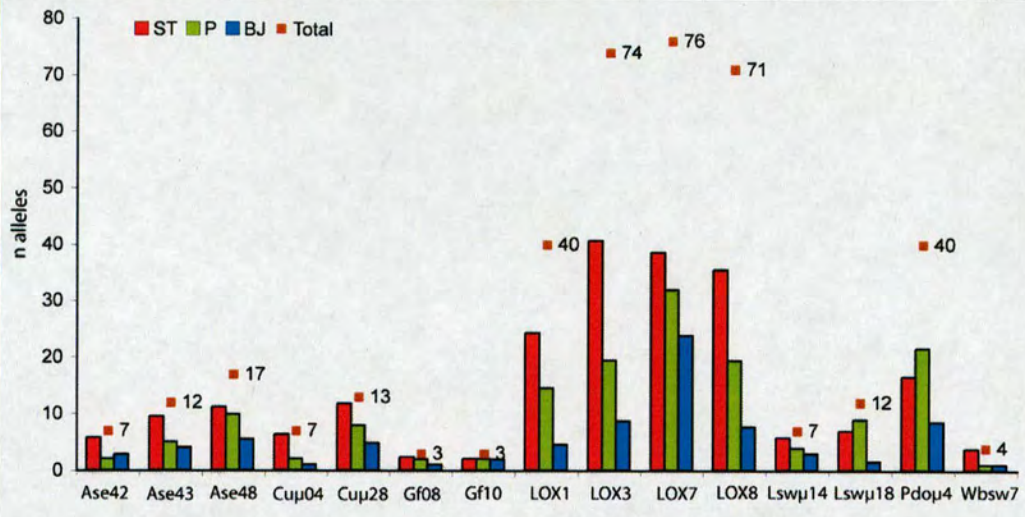


Fig 6.3. Allelic richness of each locus in each Príncipe seedeater population (ST: São Tomé; P: Príncipe; BJ: Boné de Jóquei) correcting for differences in sample sizes, and total number of sampled alleles.

Table 6.3.

Inference of the number populations (*K*) of the Príncipe seedeater with a Bayesian model-based clustering approach implemented in STRUCTURE (Pritchard *et al.* 2000).

K	log *P*(*X*|*K*)	*P*(*K*|*X*)
1	-13209	0
2	-11496	≈0
3	-10864	≈1
4	-11026	≈0
5	-11652	0
6	-11275	≈0

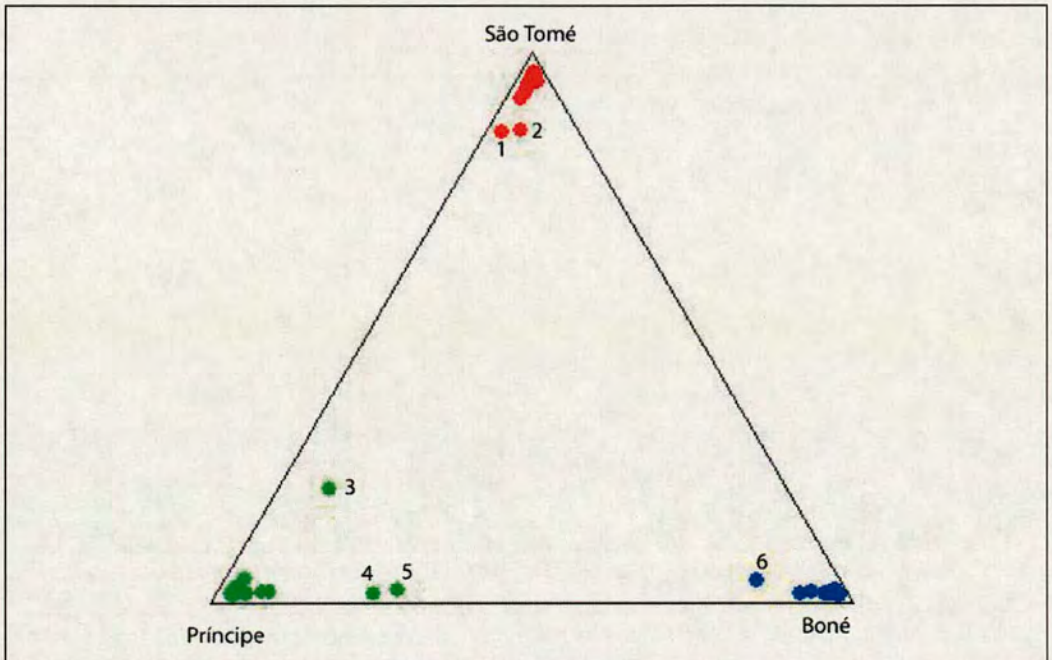


Fig 6.4. Summary of the clustering results of the program STRUCTURE based on 15 loci, and assuming three populations. Each point represents an individual and its location represents its mean estimated ancestry. Clustering was performed without providing information on the sampling location of the individuals. Overall the individuals from each island clustered together (red: São Tomé; green: Príncipe; blue: Boné). Numbers indicate possible outliers, for which the probability of being immigrants or descendents from immigrants was estimated (see text and Table 6.6).

Table 6.4.

Estimates of pairwise fixation indices using Weir & Cockerham's (1984) F_{ST} estimator (θ, below diagonal) and Rousset's (1996) R_{ST} estimator (ρ, above diagonal).

	ST-forest	**ST-plantation**	**Príncipe**	**Boné**
ST-forest	-	0.003	0.229	0.411
ST-plantation	0.001	-	0.218	0.396
Príncipe	0.152	0.145	-	0.112
Boné	0.282	0.267	0.150	-

Migration estimates with MIGRATE-N complemented the population differentiation analyses: gene flow between São Tomé habitats is high, whereas it is low between islands (Table 6.5). The MCMC sampling appears to have explored satisfactorily the parameter space as assessed by the acceptance ratios and the histograms of $\ln P(G|P)$ (probability of the genealogies given the parameters). Longer searches and with more realistic evolutionary models would have been preferable but computationally daunting. Estimates should therefore be interpreted more from a relative perspective (as they always should be for such a difficult parameter as past migration). Gene flow between islands was higher from São Tomé to Príncipe, followed from Boné to Príncipe. These were the only cases with mean migration estimates between islands slightly above the one-migrant-per-generation level that opposes population differentiation (Wright 1931; Spieth 1974). It is possible that longer searches would have provided lower values. Interestingly, the migration rate from Príncipe to Boné was very low. The effective population size estimator (theta: 4 x effective population size x mutation rate) confirmed that the population from São Tomé is much larger than the Príncipe and Boné populations, and at the same time revealed that these last two populations have similar sizes (Table 6.5).

The probability of recent migration events estimated in STRUCTURE by combining genetic data with information on the geographic origin of the sampled individuals is presented in Table 6.6. The highest probabilities of recent migration events were between Boné and Príncipe, with one individual from Príncipe having a high probability of being a descendent from a grandparent from Boné (Table 6.6).

Table 6.5.

Maximum likelihood estimates of theta ($4N_e\mu$: 4 x effective population size x mutation rate per generation) and N_em (number of effective migrants per generation) between the four study populations. Mean (bold) and 95% confidence interval presented. Estimates based on 11 loci out of the 15 used in this study (one Z-linked locus, one locus with 20% missing data and null alleles and two loci with very little variation excluded).

	N_em	*From*			
Theta	*To*	**ST-forest**	**ST-plantation**	**Príncipe**	**Boné**
3.11-**3.29**-3.49	**ST-forest**		9.52-**10.46**-11.48	0.43-**0.53**-0.64	0.41-**0.49**-0.58
3.70-**3.88**-4.06	**ST-plantation**	4.30-**4.79**-5.33		0.29-**0.35**-0.42	0.39-**0.46**-0.52
1.45-**1.54**-1.65	**Príncipe**	1.51-**1.86**-2.26	2.39-**2.85**-3.38		1.38-**1.59**-1.83
1.28-**1.33**-1.39	**Boné**	0.30-**0.38**-0.49	0.71-**0.85**-1.02	0.29-**0.35**-0.43	

Table 6.6.

Assessment of the probability that the 'outliers' in the STRUCTURE clustering (individual numbers from Fig 6.4) are immigrants or descendent from recent immigrants. Two different migration priors (ν) were used (highest values correspond to higher probability of migration). Values in bold indicate individuals with high probabilities of having migrant ancestry.

Individual	Geographic origin	Alternative source	ν	No immigrant ancestry	Immigrant	Immigrant parent	Immigrant grandparent
1	São Tomé	Príncipe	0.05	0.973	0.000	0.000	0.027
			0.10	0.939	0.000	0.000	0.061
2	São Tomé	Príncipe	0.05	0.842	0.000	0.007	0.152
			0.10	0.717	0.000	0.012	0.271
3	Príncipe	São Tomé	0.05	0.539	0.000	0.156	0.234
			0.10	0.354	0.000	0.220	0.327
4	Príncipe	Boné	0.05	0.516	0.000	0.001	0.483
			0.10	0.266	0.000	0.001	**0.733**
5	Príncipe	Boné	0.05	0.329	0.000	0.001	**0.659**
			0.10	0.131	0.000	0.001	**0.853**
6	Boné	Príncipe	0.05	0.534	0.002	0.000	0.401
			0.10	0.338	0.002	0.000	**0.568**

Morphological analysis

A two-way MANOVA revealed significant morphological differentiation between populations for all variables except bill depth, even when taking into account across-years measurement variation. Because bill depth was only measured in two out of the three field seasons and it did not differ significantly between populations, it was discarded from the dataset, and analyses were re-done on the larger six-variable dataset (130 birds), with all variables remaining significantly different between populations (Table 6.7). There were significant differences between years for bill length and width, and for tail length. Because measurements were always taken in the same period it is unlikely that these differences reflect seasonal variation in wear (Gosler 1986), but instead reflect consistent year biases in the measurement of these traits (bill measurements being particularly sensitive: Clegg *et al.* 2002b). Nevertheless, this is unlikely to affect the main conclusions because all populations were sampled in each year. Furthermore, in the dataset with bill depth (only two last years) only tail length remained significantly different, but the same results were obtained (not shown). PCA plots depicted a clear differentiation between the allopatric populations but extensive overlap for the parapatric populations from São Tomé (Fig. 6.5). The first principal component (PC1) was dominated by mass, and the second (PC2) by tail length. PC1 had positive and negative factor loadings, whereas all factor loadings were positive on PC2. Therefore PC1 might explain both body size and shape, and PC2 only body size. Together they accounted for 83% of the variance. PC3 was a 'bill axis', but accounted only for another 8% of the variance. Tukey tests confirmed that differentiation was significant only between allopatric populations, but not between the two São Tomé habitats (Fig. 6.6). The three island populations differed significantly from each other for most variables, but mass and bill width did not differ between Príncipe and São Tomé (Fig. 6.6)

Table 6.7.

Results of the two-way MANOVA between populations and across years for adult males. Significant *P* values in bold (****: $P < 0.0001$).

	Whole model			Population			Year		
Character	**d.f.**	***F***	***P***	**d.f.**	***F***	***P***	**d.f.**	***F***	***P***
Mass (g)	5	44.16	****	3	66.05	****	2	0.11	0.8974
Wing length (mm)	5	15.17	****	3	20.22	****	2	2.30	0.1053
Tail length	5	114.35	****	3	82.67	****	2	107.98	****
Tarsus length	5	14.68	****	3	22.70	****	2	0.29	0.7500
Bill length	5	8.19	****	3	10.26	****	2	5.77	**0.0042**
Bill width	5	20.06	****	3	24.03	****	2	4.04	**0.0205**
Bill depth[1]	4	0.41	0.8011	3	0.49	0.6937	1	0.07	0.7929
Wilk's lambda	18	26.38	****						

[1]: Bill depth was measured in two out of the three field seasons, hence the different degrees of freedom (d.f.).

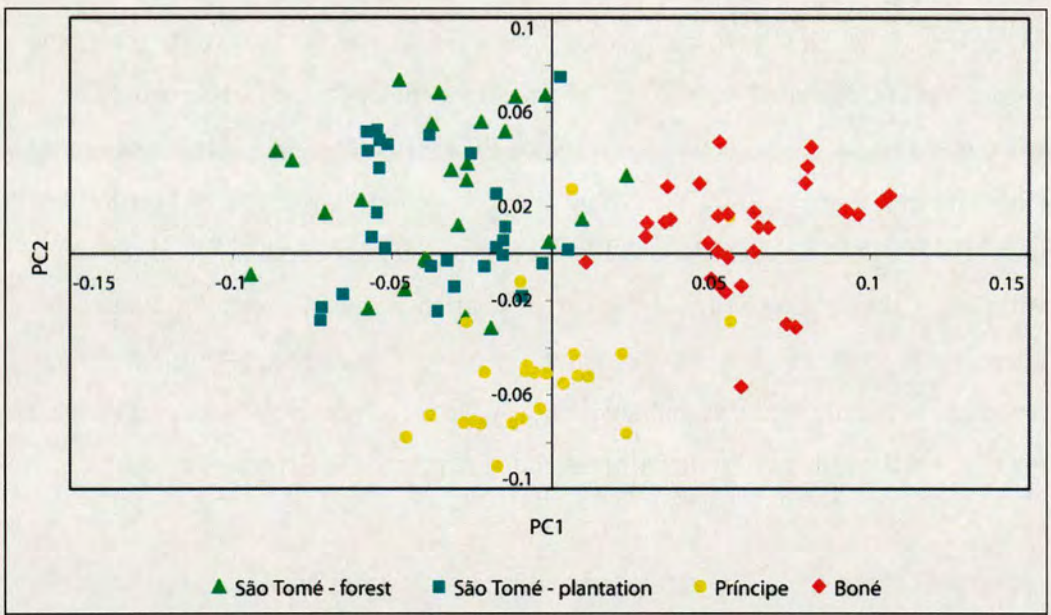


Fig. 6.5. Principal component plot of six morphological characters from 130 adult males. PC1 and PC2 explained approximately 83% of the variance. They were both related to size, but PC1 could also reflect shape.

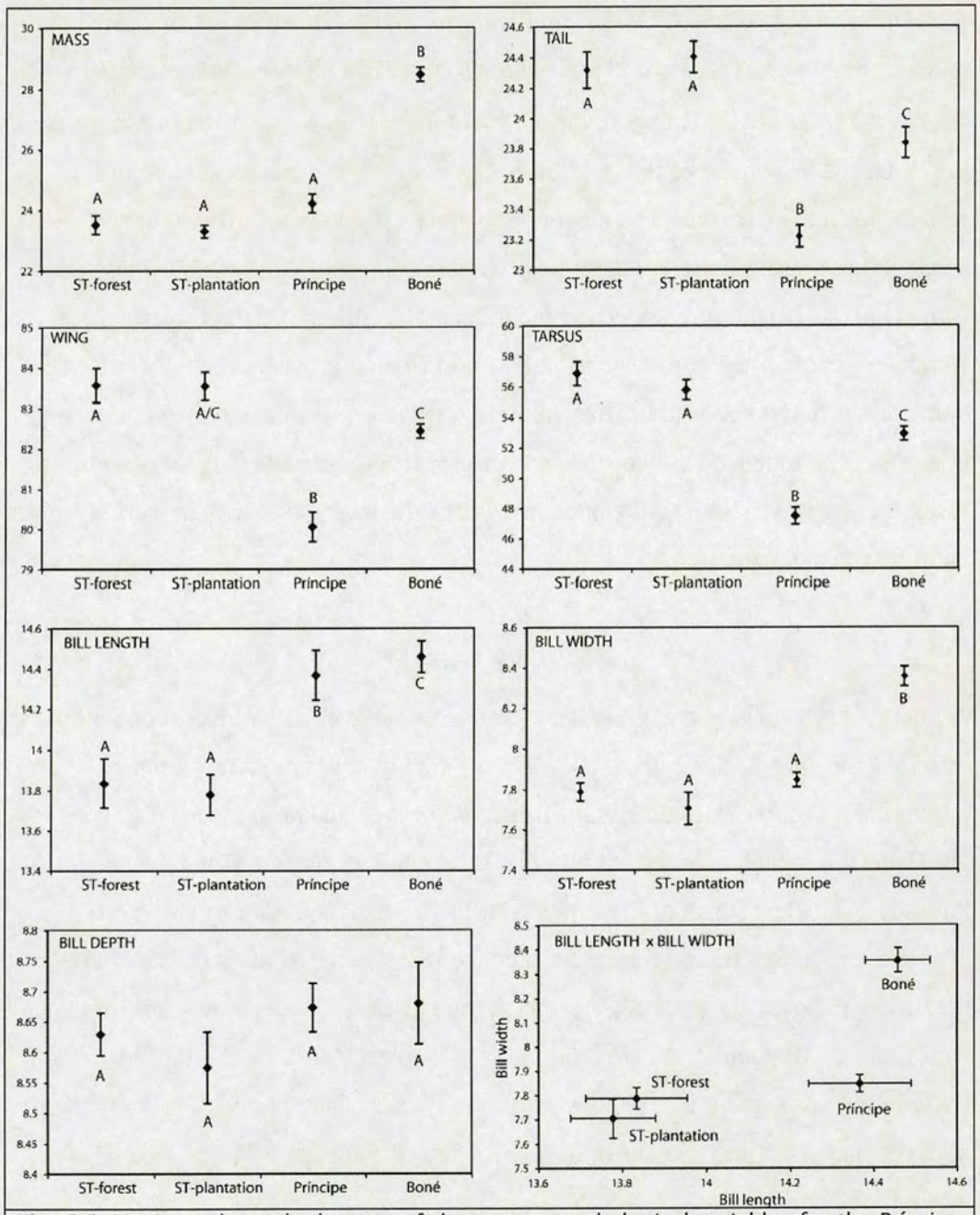


Fig. 6.6. Means and standard errors of the seven morphological variables for the Príncipe seedeater males in each study population (ST = São Tomé). Mass in g; all other in mm. Sample sizes: ST-forest, 30; ST-plantation, 33; Príncipe, 30; Boné, 36. Means with the same letter were not significantly different, as determined by post-hoc Tukey tests on the MANOVA results.

Colour analysis

Measurement repeatabilities were significant for all colorimetrics and in all feather patches, but overall repeatabilities were not particularly high (mean ± standard deviation $r = 0.58 \pm 0.11$; maximum $r = 0.76$ for the hue of the collar, minimum $r = 0.378$ for the brightness of the head), a consequence of the great sensitivity of reflectance measurements. The most common pattern of colour differentiation separated two groups: the populations from São Tomé and those from Príncipe and Boné (Fig. 6.7). This corresponds to the human perception of birds from Príncipe and Boné being brown-red and those from São Tomé being greyish. The only colour patch that differed between habitats in São Tomé was the collar (in brightness and hue; Fig. 6.7). Independently of the significance level, the birds from São Tomé forest were always closer to the populations of Príncipe and Boné than were the birds from São Tomé plantations.

Song analysis

A total of 1198 songs were used for the analysis (Table 6.7). These grouped into 116 different song types. Song diversity was very high overall and within each population (Table 6.7). Song structure differed between islands: most song types of São Tomé birds had at least one trill (99%), 79% of the songs of Boné also had trills, but only half of the songs of Príncipe had trills. By discarding information on elements or phrases framing the main theme of each song, 60 main themes were identified and used for the study of song differentiation. These themes discriminated well between all populations (Fig. 6.8, overall ANOSIM $R = 0.44$, $P < 0.001$). Differentiation was significant between the allopatric island populations and also between the São Tomé habitats (pairwise ANOSIM R between 0.34 and 0.52, $P < 0.001$ for all comparisons).

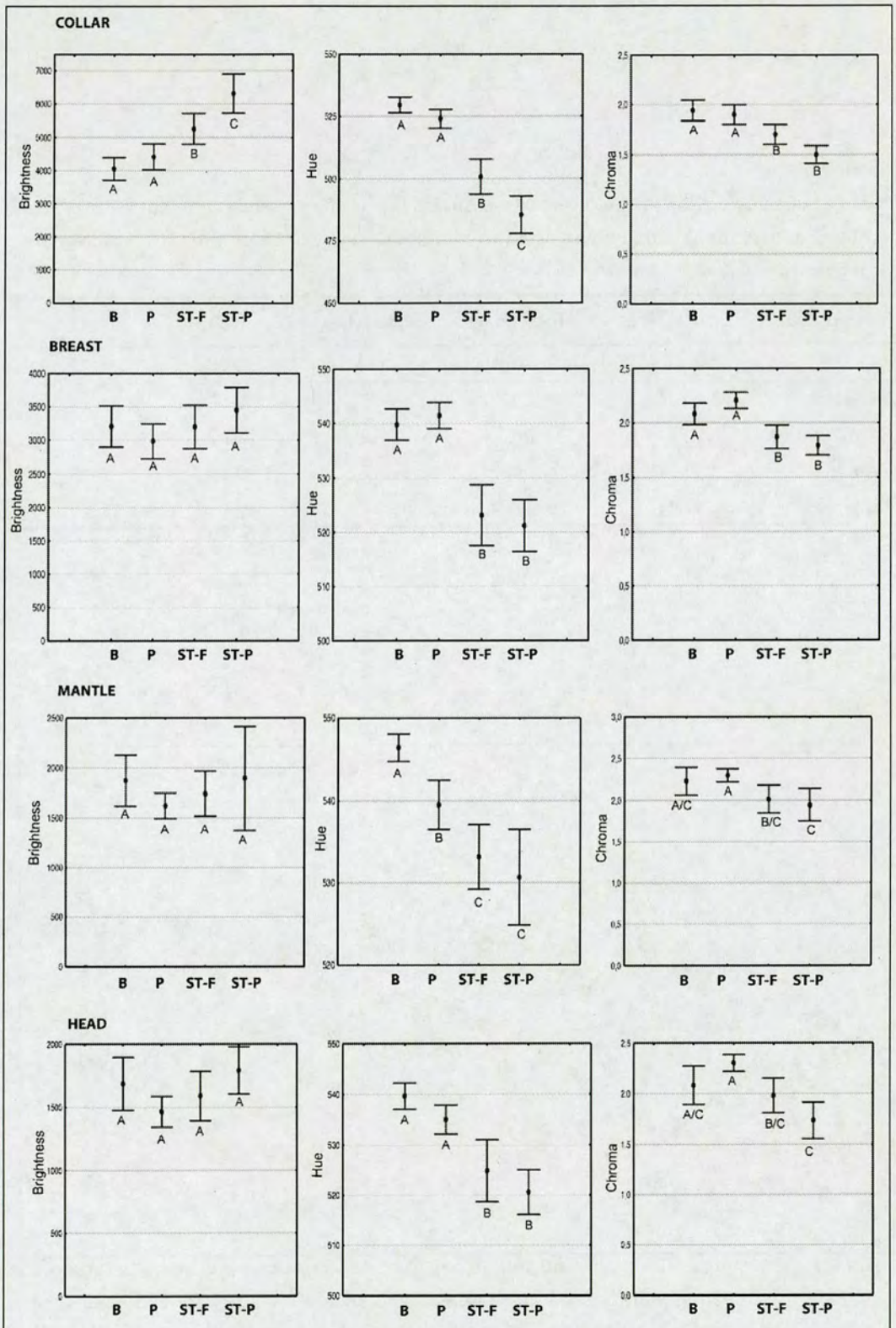


Fig. 6.7. Population means and 95% confidence intervals for the three colorimetrics measured in four feather patches in the Príncipe seedeater. B: Boné; P: Príncipe; ST: São Tomé, F: forest, P: plantation. Means with the same letter were not significantly different. Twenty individuals sampled per population, except for ST-P (19 individuals).

Table 6.7.

Number of sampled songs and diversity measures. 1-D = Simpson diversity Index - gives the probability that songs randomly selected from a population will be different (0: no diversity, 1: highest diversity).

Population	Total birds	Total songs	Song types	1-D	% Trills
ALL	153	1198	116	0.98	73
ST-forest	36	242	35	0.93	98
ST-plantation	24	207	30	0.92	99
Príncipe	51	529	37	0.96	49
Boné	42	220	27	0.93	79

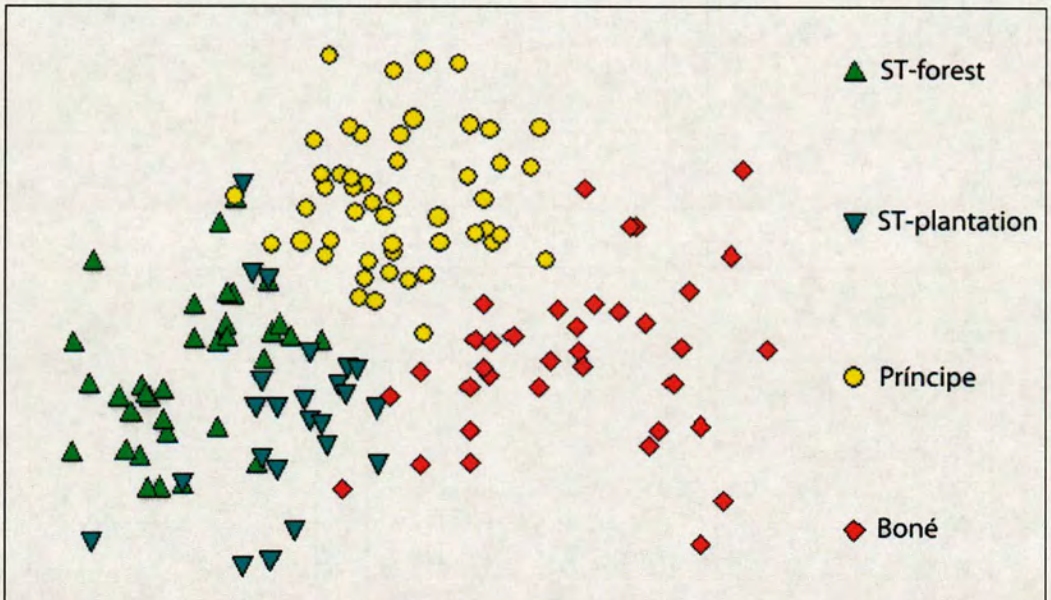


Fig. 6.8. Non-metric MDS plot of 60 song types from 153 Príncipe seedeaters, based on a Bray-Curtis similarity matrix of the standardized and square root transformed song type abundance data. (Stress = 0.08).

Comparison of genetic and phenotypic differentiation

The colour of the mantle was the only trait that was significantly correlated with genetic variation; song and morphological variation had also high correlations with genetic variation, very close to the 5% significance level for song and to the 10% level for morphology (Table 6.8). The colour of the collar, breast and head had weak and non-significant correlations with genetic variation. The last two traits had the lowest correlations with genetic variation, as both put the São Tomé populations closer to the Boné population rather than Príncipe (Fig. 6.9). The neighbour-joining trees derived from all other phenotypic traits inferred the same topological relationships between populations but inferred different amounts of change between populations in relation to the genetic data (Fig. 6.9). This is particularly evident for the colour of the breast where the population from São Tomé forest is almost equidistant (and far) from the populations of São Tomé plantation and Príncipe. In song, a trait very closely related to genetic variation according to the Mantel tests, the populations from the two São Tomé habitats are much more divergent than what is inferred by the genetic data.

Table 6.8. Results of the Mantel tests between the neutral genetic distance matrix (F_{ST}) and distance matrices computed for six phenotypic traits. Similar results were obtained with R_{ST}.

Trait	*R*	R^2 %	*P*
Morphology	0.75	56.66	0.106
Song	0.85	71.42	0.055
Colour: Collar	0.53	28.34	0.283
Colour: Breast	0.25	6.33	0.662
Colour: Mantle	0.93	86.05	0.013
Colour: Head	0.12	1.55	0.797

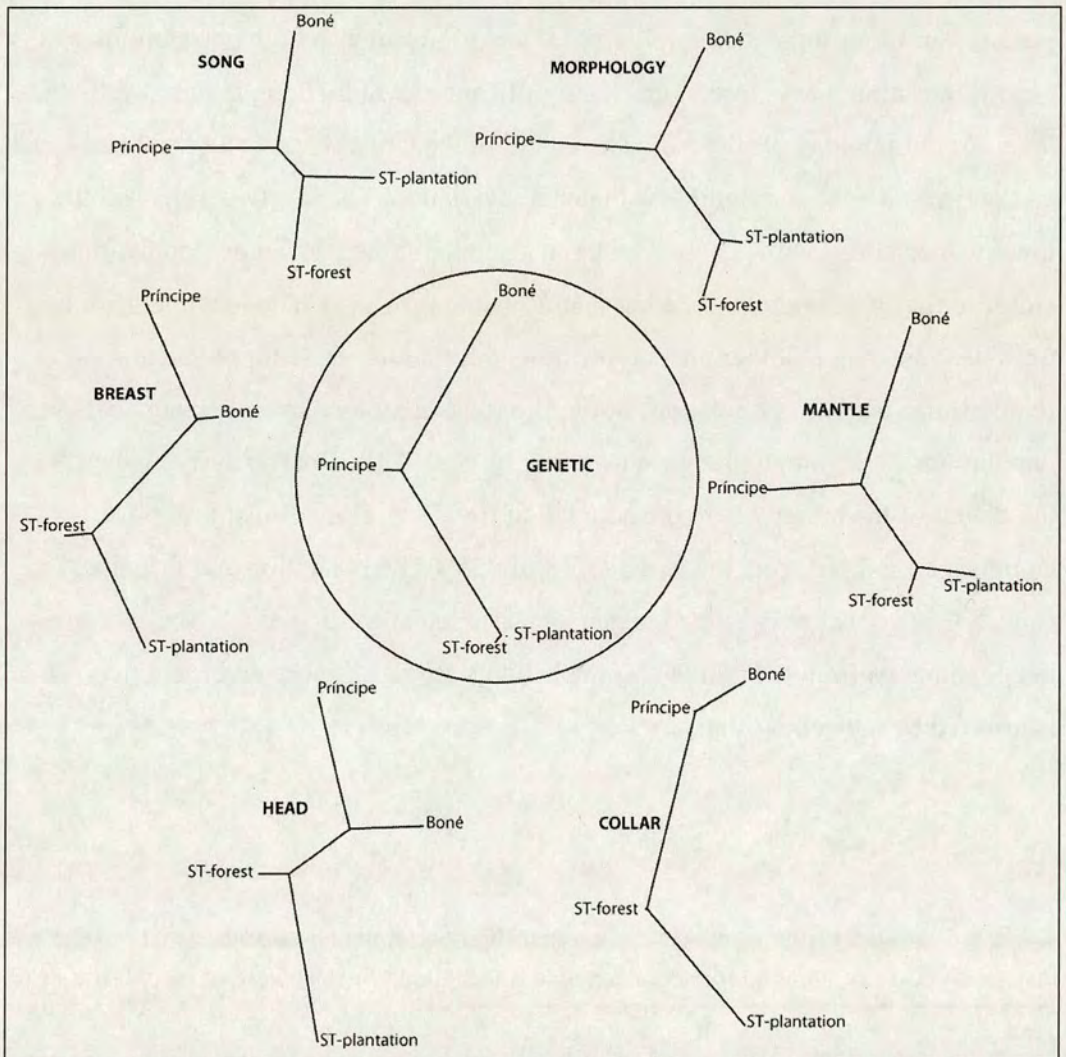


Fig. 6.9. Neighbour-joining trees depicting the relationships between the four study populations of the Príncipe seedeater as inferred from neutral genetic and phenotypic variation. Genetic distance: F_{ST} (same tree for R_{ST}); song distance: Bray–Curtis dissimilarity, where similarity reflects the number of song types shared; all others: normalized Euclidean distances.

DISCUSSION

The three allopatric populations of *S. rufobrunneus* differentiated in all dimensions analysed (neutral genetic markers, adaptive morphological traits, mate recognition traits). Their taxonomic classification, which attributes a sub-specific status for each insular population (Naurois 1975b), captures correctly this biological reality. In contrast to the allopatric populations, the parapatric populations in São Tomé diverged only in mate recognition traits. This result cannot be used to support a more important role of isolation over divergent selection in driving population divergence because the allopatric populations have been diverging since their separation when the islands were originally colonised, whereas the parapatric populations have had only c. 300 years to diverge (see Introduction). What it does support is the potential for pre-mating traits to evolve rapidly and therefore play a role in the initial stages of the evolution of reproductive isolation. Furthermore, the sister relationship between *S. rufobrunneus* and the São Tomé grosbeak (*Neospiza concolor*; Chapter 3) suggests, if anything, that the evolution of morphological divergence and speciation proceeded much faster without isolation.

The patterns of genetic and phenotypic variation in *S. rufobrunneus* are discussed next, and interpreted in light of the two main themes of this study: the role of drift in divergence and the potential for mate recognition traits to trigger speciation. A cautionary note is needed: this was an exploratory, correlative investigation with the aim of providing a bird's eye view of the likely diversification dynamics in this system. The approach used to distinguish between drift and selection has many limitations (Clegg *et al.* 2002b), and inferences of selection pressures consisted of what Schluter (2000) called 'educated guesses' – based on environmental correlates and the assumption that the studied traits are fitness correlated, rather then on direct measurement of selection on breeding or survival. Descriptions of habitat and habitat use by the seedeater are crude – based on the literature and personal observations rather than detailed ecological work. Nevertheless, in a system that offers so much potential for the study of speciation but where so little is known, this seemed a legitimate first step, from where other research lines can be drawn.

Genetic data

The evolutionary dynamics of each insular population have been determined mainly by isolation, as reflected by the high levels of population structuring and low estimates of migration events. They therefore constitute adequate models to study speciation in allopatry. The high differentiation between the populations from Boné and from Príncipe was surprising considering that only 3 km separate the islands, and none of the populations has lost the ability to fly. Other similar cases have been reported in birds and support the view that, even if flight capacity is not reduced or lost, the crossing of small water barriers may be impeded by behavioural barriers (Mayr 1963; Diamond 1981; Komdeur *et al.* 2004). This result supports vicariance, rather than dispersal, as the likely cause of separation of the two populations. Boné is separated from Príncipe by shallow seas, and was therefore part of Príncipe during glacial periods, being connected for the last time 12,000 – 10,000 years ago. The little gene flow that occurs between the two islands is dominated by migration from Boné to Príncipe. This 'back-colonization' pattern (i.e., island → mainland) was also found in the Bioko population of the Little Greenbul *Andropadus virens* which probably also originated by vicariance (Smith *et al.* 2005).

The relative estimates of effective population sizes (N_e) obtained from the program MIGRATE (i.e., N_e x mutation rate per generation) highlight the extraordinary densities on Boné islet, which had a total population estimate close to that of Príncipe. If one uses the estimate for the mutation rate for avian nuclear microsatellites of 10^{-4} (e.g., Petren *et al.* 2005), N_e of Boné would be of 3,325 birds, very close to the estimate for Príncipe of 3,850 birds (for São Tomé: 18,000). Interestingly, this estimate for Boné is close to a census size estimate from 1999 (3,487 birds: Baillie & Gascoigne 1999). This N_e estimate seems nevertheless inflated, as the high densities in such a restricted area (0.3 km^2) would surely limit access to breeding to a fraction of the population – therefore the census estimate would be substantially higher than 3,000 birds, which is unlikely.

No genetic structure was detected between the forests and plantations in São Tomé. This could reflect either extensive gene flow between forests and plantations, or it

could be due to the short timeframe available for differentiation by genetic drift. Movements between habitats are supported by the fact that the diet of the Príncipe seedeater is mainly fruit-based, which favours the evolution of a 'nomadic' behaviour required to search for fruiting patches (Mayr 1963). Evidence supporting a nomadic behaviour within São Tomé is still anecdotal: 106 individuals were captured during three field-seasons but no individual was ever recaptured, contrasting with 18 recaptures out of 276 captures for the insectivore São Tomé speirops *Speirops lugubris*. The alternative hypothesis of restricted gene flow cannot be rejected because the combination of large effective population sizes of the seedeaters together with the high variability of most of the sampled microsatellite loci makes differential fixation of alleles unlikely in such a short timeframe. The expected rate of loss of genetic heterozygosity per generation is $1/2N_e$ (Wright 1931; Lande 1980). Tropical birds are characterised by having slow life-histories, which probably includes delayed onset of reproduction (Martin 1996). Therefore, considering a generation time for the seedeater of 3 years, as documented for other tropical passerines (Clegg *et al.* 2002b), a N_e of c. 10,000 for each population, and a divergence period of 300 years, frequencies of neutral markers should differ by only 0.5% between habitats.

Morphology

The morphological traits used in this study have been shown to be highly heritable and correlated with performance and fitness through their link to flight and diet (Grant 1986; Schluter & Smith 1986; Smith & Girman 2000). Except for bill depth, there were significant morphological differences between the three island populations for all traits. No differences between São Tomé habitats were detected. Lack of differentiation between habitats can reflect high levels of gene flow or indicate that selective pressures on the measured traits do not differ significantly. As explained above, genetic differentiation may not yet be detectable because the habitat divergence is recent and the population sizes are large. Plant species composition differs strikingly between mature forests and shade forest plantations, suggesting that selective pressures on foraging are also different. Nevertheless, because the seedeater is a generalist and the structure of the two habitats is relatively similar (Jones & Tye 2006), with both offering a wide diversity of food resources, the selective differential

for traits associated with foraging (including flight) could be small. Lack of morphological differences between habitats would then reflect the 'generalist phenotype' of the seedeater. Evidently, field studies are required to test the causes of the lack of differentiation between habitats. Morphological divergence between the three allopatric populations showed a similar pattern to neutral genetic divergence, but the association was not significant. This suggests that neutral evolution may have played a role in the evolution of the isolated populations. Nevertheless, morphological shifts followed the directions predicted by the habitat. The population from Boné had the most divergent bill, in agreement with Boné offering the most distinct habitat: it is mainly a monoculture of oil palms, in contrast with the diverse habitats of São Tomé and Príncipe. The bill of Boné seedeaters is stouter (longest and widest) which could reflect an adaptation to hammer the very hard palm nuts in order to reach the oil located inside. Boné seedeaters were also the heaviest birds with the lowest wing/weight ratio. The large body size in Boné (the largest of the three islands) could be a consequence of the increase in bill size and/or driven by the high levels of intraspecific competition combined with little, if any, interspecific competition and no predation, in agreement with predictions for body size evolution on islands (MacArthur *et al.* 1972; Clegg & Owens 2002; Robinson-Wolrath & Owens 2003). The lowest wing/weight ratio is in agreement with the much-reduced need to fly in the small Boné islet, densely covered with palms.

Mate recognition traits

Plumage colour and song, together with behavioural displays, constitute the traits involved in mate recognition and choice in birds. Colour and song signals can be used in association for species or individual recognition (Baker & Baker 1990; Matyjasiak 2005). They are therefore responsible for pre-mating isolation between species, but their potential to promote assortative mating (and therefore the evolution of reproductive isolation) between diverging populations is still not resolved (Chapter 7). This study has showed that these traits can diverge rapidly and before other phenotypic traits, and therefore have the potential to play a role in the early stages of the speciation process.

Plumage colour

From the four areas of the body sampled for colour, mantle variation was the only characteristic strongly related to genetic variation, and therefore whose evolution could be accounted for by random factors alone. In contrast, the colour of the breast, head, and collar were the phenotypic traits with the highest departure from the neutral variation pattern. These results suggest that the mantle (the back of the birds) is a neutral trait, whereas the areas of the body that are visible in face-to-face interactions are under selection. They also confirm that plumage colour is a labile trait which can rapidly lose any phylogenetic signal (c.f. Chapter 2 and: Moreau 1957; Burns 1998; Omland & Lanyon 2000; Moyle *et al.* in press). The higher departure from the neutral pattern of evolution shown by colour characters than by any of the other phenotypic traits is in agreement with the view that plumage traits play a more important role in allopatric speciation in birds than morphological traits (Grant 2001). Nevertheless, even for the traits under selection, the phylogenetic signal is not completely lost when each colorimetric variable is analysed separately, because Príncipe and Boné often come close together.

Except for the mantle, the seedeaters from the São Tomé plantations were as or more divergent from the seedeaters of the São Tomé forests as these were from the other allopatric populations, as measured by multivariate Euclidean distances. This supports a role for natural selection in driving colour divergence in this system. When each colorimetric variable is taken separately, the values on the plantations are always the most extreme, even if differences are only significant for the collar (which could be due to lack of power and multiple comparisons). What is strongly suggestive is that the colour shifts are in agreement with differences in the light environment. Specifically, the closed habitats (São Tomé forests, Príncipe, Boné) had the highest hue values, and the plantations had the highest brightness values. High hue values reflect the long wavelengths that dominate closed habitats (Endler 1993; McNaught & Owens 2002), whether brightness could be more useful in open habitats (McNaught & Owens 2002 *contra* Endler 1993) because it maximizes contrast over long distances. The fact that the only significant plumage differences between São Tomé habitats were in the brightness and hue (of the collar) further

supports the light environment hypothesis. This result also suggests that in the Príncipe seedeater the collar is the most important area of the body used for communication. The collar consists of an area of feathers with distinct coloration from the surrounding matrix, making it a well-defined patch. Similar contrasting feather patches have been shown to play a role in bird communication (Lank & Dale 2001; McGraw *et al.* 2003; Badyaev & Young 2004).

The brownish colour of the seedeater plumage is likely to be based on the melanin pigment, whose production is under strong genetic control (Roulin & Dijkstra 2003). Therefore, the significant divergence of the colour of the collar between São Tomé habitats reflects genetic changes that most likely took place in the face of high gene flow (see 'Genetic data'), supporting an important role for selection. Despite being highly heritable, melanin traits can be an honest signal of phenotypic quality because their full expression depends on the acquisition of scarce minerals from the diet that are important co-factors in different physiological processes, and they could also reflect the ability to neutralise metals that are toxic in large concentrations (McGraw 2003; Nieeke *et al.* 2003; Griffith *et al.* 2006). One could then envisage that both sexual selection (choice of best quality males) and natural selection (increasing signal visibility in relation to the light environment) could participate in the fast evolution of this signal. Rapid evolution would be favoured by the simple genetic basis of melanin-based traits: it has been shown that a single point mutation in the *melanocortin-1 receptor* locus is implicated in colour polymorphisms of several species (Mundy 2005; Hoekstra *et al.* 2006).

Song

The song repertoire of the Príncipe seedeater was very diverse in all islands with no detectable effect of island or population size. This pattern goes against a common trend of song simplification with the increase in insularity (Baker & Jenkins 1987; Baker 1996; Hamao & Ueda 2000; Baker *et al.* 2001). This simplification is often associated with colonisation events by a few individuals carrying with them a small subset of the total song pool and with reduced genetic diversity. That this effect has not been detected even in the tiny Boné islet supports vicariance, rather than

dispersal, as the cause of isolation of this population. Alternatively, the high song diversity in the inbred Boné population could be a result of cultural evolution after colonization (e.g., Lynch & Baker 1994; Baker *et al.* 2003). The high *S. rufobrunneus* densities in Boné suggest that breeding opportunities are restricted and therefore mate recognition traits could be under strong sexual selection.

Together with plumage colour, song was the only other phenotypic trait that diverged between São Tomé habitats. This probably reflects adaptation to different acoustic environments (open and closed habitats), a well-documented process in birds (Slabbekoorn & Smith 2002; Slabbekoorn & Peet 2003; Podos *et al.* 2004; Seddon 2005; Ruegg *et al.* 2006). Song divergence between the allopatric populations closely followed the pattern of neutral evolution. This suggests that song among isolated populations in similar habitats can continue to diverge. Divergence could be driven by the male signal/female preference co-evolution cycle (Isawa & Pomiankowski 1995), which would fit a neutral model if it proceeds at the same rate in all populations, and/or by cultural drift. Even if divergence follows a neutral model, innovations are probably kept within the bounds imposed by selection for optimal song transmission. In this study, song divergence describes differences in song syntax (number, shape, and order of notes). These dimensions of song may change without changes in acoustic characteristics related to signal transmission. Quantification of the acoustic characteristics of the songs is therefore needed to test the role of selection: if selection is important they should differ between different acoustic environments even in parapatric situations, but not between allopatric populations occupying the same acoustic environments.

Implications for speciation

This study showed that the allopatric populations of the Príncipe seedeater are well advanced into the first stage of the 'archipelago radiation model' (Chapter 1): speciation starts with divergence of isolated populations and is completed (or aborted) when the diverging populations meet in sympatry. The mechanisms causing initial population divergence in allopatry are still a matter of debate: are they all selection-based or can random factors be important as well?

The evidence gathered in this study supports a primacy of divergent selection over random processes in the evolution of genetically based traits, while supporting a role for random change in song, a trait with a cultural component. But even the amount of random change in song is probably constrained by selection. Results support the current consensus that divergent selection is the main driver of phenotypic change and speciation, even in allopatric populations (Orr & Smith 1998; Rieseberg *et al.* 2002; Coyne & Orr 2004; Rundle & Nosil 2005). The same conclusions have been drawn from studies of the Darwin's finches in the Galápagos (Grant *et al.* 2000; Tonnis *et al.* 2005), an equivalent situation to the Gulf of Guinea islands. Nevertheless, it is still possible that genetic drift could have played a role in the initial stages, but that the signature of selection subsequently erased its effect. All models that postulate an effect of drift in speciation of small allopatric populations consider it to be important only in the first stages (Mayr 1954; Templeton 1980; Carson & Templeton 1984) – and therefore it might be impossible ever to solve this debate in a natural situation.

That population divergence is mainly driven by selection is a relatively trivial assertion – after all, it constitutes the fundamental basis of the theory of evolution. What is essential for the speciation process is that this divergence is followed by the evolution of reproductive isolation. For this to be possible, it is necessary that the divergent resource-use traits be genetically linked to mate recognition traits, leading to assortative mating between the incipient species (Kirkpatrick & Ravigné 2002). This association might be the most difficult condition for the speciation process, but it should be much easier to evolve in allopatry (Bolnick 2004). In this study, the three allopatric populations have been shown to diverge in genetic, morphological and mate recognition traits, indicating that the different pieces required for speciation are in place. The next step will be to test whether the differences in mate recognition traits are perceived as such by the birds, and have therefore the potential to act as reproductive barriers. In the next chapter, the potential of song divergence to constitute a reproductive barrier between populations was tested by means of playback experiments.

7

FROM POPULATION DIVERGENCE TO SPECIATION: RESPONSES OF ALLOPATRIC POPULATIONS OF THE PRÍNCIPE SEEDEATER *Serinus rufobrunneus* TO FOREIGN SONG DIALECTS

The Príncipe seedeater *Serinus rufobrunneus* is a songbird endemic to three Gulf of Guinea islands (São Tomé, Príncipe, Boné). Populations on each island have distinct song dialects. Playback experiments were conducted to determine how the different populations perceive the different dialects, and therefore if song divergence in allopatry might constitute a barrier to gene flow in case of secondary contact. Eight territories were tested in São Tomé, eight in Príncipe, and five in Boné. In each territory four playback experiments were carried out with songs from four categories: i) song from 'own' population; ii & iii) song from 'foreign' population (2 foreign populations); iv) song from a distinct but related species (thick-billed seedeater *Serinus burtoni*). All populations responded more strongly to their own song than to foreign or heterospecific songs. This pattern was significant only for São Tomé, where the response of birds to their 'own' song was different to their response to foreign songs, and the response to foreign songs was not different to response to heterospecific songs. Birds on Boné were more responsive to all stimuli in accordance with their inquisitive behaviour, probably derived from the high densities of this population and a likely concomitant loss of territoriality. Nevertheless they still responded more strongly to their own songs, and non-significance of the results could be an artefact of the small sample size. Birds on Príncipe responded little even to their own songs suggesting that playback experiments were carried in a period when territorial defence and/or mate attraction were not important. Overall, the results showed that birds do not recognise songs from different populations, and therefore that song may constitute a reproductive barrier in case of secondary contact. The outcome of secondary contact will depend also on the level of morphological divergence and the possibility of co-existence of the two diverging populations. Two outcomes are possible: song learning can dilute the differences evolved in allopatry, so breaking the reproductive barrier or, alternatively selection against unfit hybrids can drive further song divergence until the completion of the speciation process.

INTRODUCTION

In birds, song is a trait used in territory defence and mate choice, and it plays a major role in species recognition (Catchpole & Slater 1995; Martens 1996; Podos *et al.* 2004). Considering that reproductive isolation in birds is maintained primarily by pre-mating isolation mechanisms and only secondarily by post-mating incompatibilities (Grant & Grant 1997a; Price & Bouvier 2002; Edwards *et al.* 2005), song is often hypothesised as a central trait implicated in the first stages of reproductive isolation. The logic is simple: if different populations within a species evolve different songs, individuals from different populations might no longer be able to recognise each other and therefore be reproductively isolated. It has indeed been extensively documented that birds respond more to familiar (local) songs than to songs from foreign populations (Catchpole & Slater 1995; Baker 2001). It has also recently been shown that this preference for familiar songs has a genetic basis (Whaling *et al.* 1997; Braaten & Reynolds 1999; Nelson 2000). Because song is a trait that can evolve very rapidly, the evolution of distinct geographic song dialects is almost ubiquitous in birds (Krebs & Kroodsma 1980; Baker & Cunningham 1985). This is because it is a trait under both sexual and natural selection (Morton 1975; Searcy & Andersson 1986; Price 1998; Slabbekoorn & Smith 2002; Podos *et al.* 2004; Seddon 2005) and, because it is a learned behaviour in songbirds, divergence can proceed particularly fast under cultural evolution (Slater 1986; Payne 1996; Lachlan & Servedio 2004).

Therefore, it has been hypothesised that the high diversification of songbirds (c. 40% of the c. 10,000 bird species) is partly linked to the fast divergence of song and its role in reproductive isolation (Fitzpatrick 1988; Martens 1996; Vaneechoutte 1997); but see (Baptista & Trail 1992). Nevertheless, there is still some debate about whether song differentiation within species can lead to the evolution of reproductive isolation, or if song only acts as a barrier when differentiation is well established in other traits (Baker & Cunningham 1985; Baptista 1985; Zink 1985; Grant & Grant 1997a; Irwin & Price 1999; Slabbekoorn & Smith 2002; Edwards *et al.* 2005). Field

evidence for the importance of local song dialects in assortative mating has provided mixed results: song can either act as a barrier to gene flow (Salomon 1989; Grant & Grant 1996; Martens & Steil 1997; Irwin *et al.* 2001; Patten *et al.* 2004) or not (Fleischer & Rothstein 1988; Lougheed & Handford 1992; Wright & Wilkinson 2001), and its role remains equivocal in many cases (Payne & Westneat 1988; MacDougall-Shackleton & MacDougall-Shackleton 2001). The potential for song to trigger speciation has received strong support from the only known case of sympatric speciation in birds, in indigobirds *Vidua* spp., where reproductive isolation is entirely dependent on song (Payne *et al.* 1998; Payne *et al.* 2000; Sorenson *et al.* 2003).

The first condition for song dialects being implicated in the early stages of the evolution of reproductive isolation is that individuals from one dialect do not recognise songs from other dialects, or at least respond much less to foreign dialects. This would indicate the potential for song to act as a reproductive barrier. Here I use playback experiments to assess if this condition is met in the Príncipe seedeater *Serinus rufobrunneus*, an endemic songbird to three Gulf of Guinea islands (São Tomé and Príncipe islands, Boné de Jóquei islet; Chapter 6, Fig. 6.2). The three allopatric populations have clearly differentiated dialects (based on syllable type and song syntax) with hardly any overlap in song types between populations (Chapter 6), suggesting that the evolution of reproductive isolation might be taking place in this system.

METHODS

The playback experiments were carried out in December 2004: 6-10 December in São Tomé, 15-16 December in Boné de Jóquei, and 17-20 December in Príncipe. This month was chosen because, although this species appears to breed during most of the year, many of the records of breeding activity fall between November and January (Snow 1950; Naurois 1975b, 1994; Christy & Clarke 1998; Jones & Tye 2006). I observed both nest construction and nestlings in January in São Tomé, and singing activity was common from October to January. Fledglings were observed in December in Príncipe and Boné.

Experimental design

Four song categories were used in playback experiments: São Tomé, Príncipe, Boné and Burtoni. 'Burtoni' consisted of songs from the thick-billed seedeater *S. burtoni*, which occurs in Mt. Cameroon. Conspecific songs from allopatric populations were designated as 'foreign'. The 'Burtoni' category was used to assess the response to a song of a distinct but related species and in this way be able to interpret the response of birds to songs of foreign populations – i.e., to determine if the reaction was similar to that elicited by a distinct species. For any given territory, four playback experiments were carried out – one with each song category. A target of 10 territories for each of the three allopatric populations was set, totalling 120 experiments. This strategy – rather than performing a single playback experiment in each territory – could potentially be affected by problems such as non-independence of responses within a territory, but was chosen due to constraints in Príncipe and Boné. The Boné population is confined to the vegetated part of the 30 ha islet, most of it difficult to access or inaccessible, making it impossible to find 40 territories (instead of 10) in the short stays that are logistically feasible. The population of Príncipe is concentrated in the southern forests, only accessible by boat and where long stays or repeated visits are also problematic. For the same reasons, the minimum time interval between different playback experiments in the same territory in Boné and Príncipe was set at 15 minutes, whereas in São Tomé they were carried at a minimum of one day intervals. In total, 85 experiments were performed, covering eight territories on São Tomé, eight on Príncipe and five on Boné. In situations where the detailed observation of a bird's behaviour was difficult and no reaction to a song was detected, a song of a bird from the same population was played at the end of the experiment: if a bird approached, the lack of response during the previous experiment was considered a valid result; if not, the experiment was treated as having failed.

Playback tapes

To avoid pseudo-replication, efforts were made to have all playback recordings (n = 40) originating from different individuals (McGregor 2000). For *S. rufobrunneus*, this was achieved for São Tomé and Príncipe categories (n = 20), whereas only five

different individuals could be used for Boné because of lack of good quality recordings. All recordings were obtained in the previous 2002-03 and 2003-04 field seasons as described in Chapter 6. Recordings were chosen for their quality: sound level, distortion, background noise, and presence/absence of vocalisations from non-focal birds. Recordings from Príncipe and Boné had higher levels of background noise because these populations occur near to the sea. Playbacks of *S. burtoni* were made from recordings from only three individuals held at the British Library Sound Archive (Wildlife Section).

Each playback recording had a total length of 120s, and contained songs from only one individual. Songs were separated by 5s intervals, with a total of six to eight songs per minute. The number of different song types per minute varied between five and seven, with the total number of song types in the 120s stimulus ranging from six to nine. The song order followed the order from the original recording whenever possible, or was such as to agree with the above criteria. When a given song type was repeated in the 120s stimulus, the best quality one from the original recording was chosen to represent it. Fade in and fade out of c. 50 ms was inserted for each song. After assembly of a playback record it was normalised and then converted to 44.1 Hz. Playbacks were recorded onto CDs.

Field experiments

Playbacks were amplified with a Phillips portable CD-player connected by a 10 m cable to a Mineroff speaker/amplifier mounted at 1.5 m above ground on a tripod. On São Tomé, territories were identified by locating singing birds the day before the experiment. On Boné and Príncipe experiments were performed after identification of territories. The speaker was placed so as to offer perches close to the speaker and simultaneously allow good visibility of the surroundings by the observer. In each territory, song category order and respective playback was randomly chosen using the 'shuffle' function. Playback recordings that were used were 'discarded' from the set available for experiments in other territories. In the case of the Boné and Burtoni categories, these were put back into the available set after all available recordings had been used (after five and three territories, respectively).

An experiment started when an individual was present in a territory and not singing (birds calling intermittently with single notes were accepted). The experiment consisted of recording the behaviour of the focal bird, or any other responsive bird, for the 120s of the playback stimulus and the following 120s of silence. All behaviours were live-recorded into a portable Sony tape recorder, and transcribed later for analysis. The main measurements in the field comprised: start distance to speaker; number of birds present; distance to the speaker; general behaviour (preening, feeding, singing). From the field observations the following measurements were extracted: latency to first approach; minimum distance to speaker during and after playback; total time within 2 and 5 m from the speaker during and after playback; total time within 10 and 20 m from the speaker; total number of calls and total duration of song during and after playback.

In general, playbacks did not elicit strong responses besides approach to the speaker. Therefore, only this behaviour was analysed in the present study. This makes this analysis conservative in favour of no detection of differences between populations, because subtle differences in behavioural responses to songs from different dialects are not taken into account. On the other hand, differences in responsiveness will strongly support the hypothesis that the different populations do distinguish between the different dialects. Differences in responses to different song categories in each population were tested with the Fisher's exact test (as most cells had values below 10); differences in responses for the combined data set were tested with chi-square tests.

RESULTS

In all populations, birds approached the speaker more often to songs of their own population than to songs from conspecific foreign populations or from the distinct species, *S. burtoni* (Fig. 7.1). The 22 'Burtoni' playbacks only elicited two responses (one in São Tomé, another on Boné). Differences in responses across the three categories were significant for the São Tomé population and for the combined dataset but were not significant on Príncipe or Boné (Fisher's exact and chi-square

tests: Table 7.1). Differences in responses of the combined and São Tomé datasets were significant between 'Own' and 'Foreign' songs but not between 'Foreign' and 'Burtoni' songs, suggesting that birds perceived the songs from allopatric populations in the same way as they perceived the songs from other species (Table 7.1). No significant differences were found in any of these pairwise comparisons on Príncipe or Boné (Table 7.1). On Boné, a lack of significant differences reflects a tendency for Boné birds to respond more than the other populations to any stimulus offered, but may be partly an artefact of small sampling size (five territories tested), since birds responded twice as much to songs of their own population than to songs of foreign populations (Fig. 7.1). In Príncipe, lack of significant differences reflects an overall lack of responsiveness: only 37.5% of the Príncipe birds responded to their own playback, in contrast to 71% and 80% of approaches to their own playback by São Tomé and Boné birds.

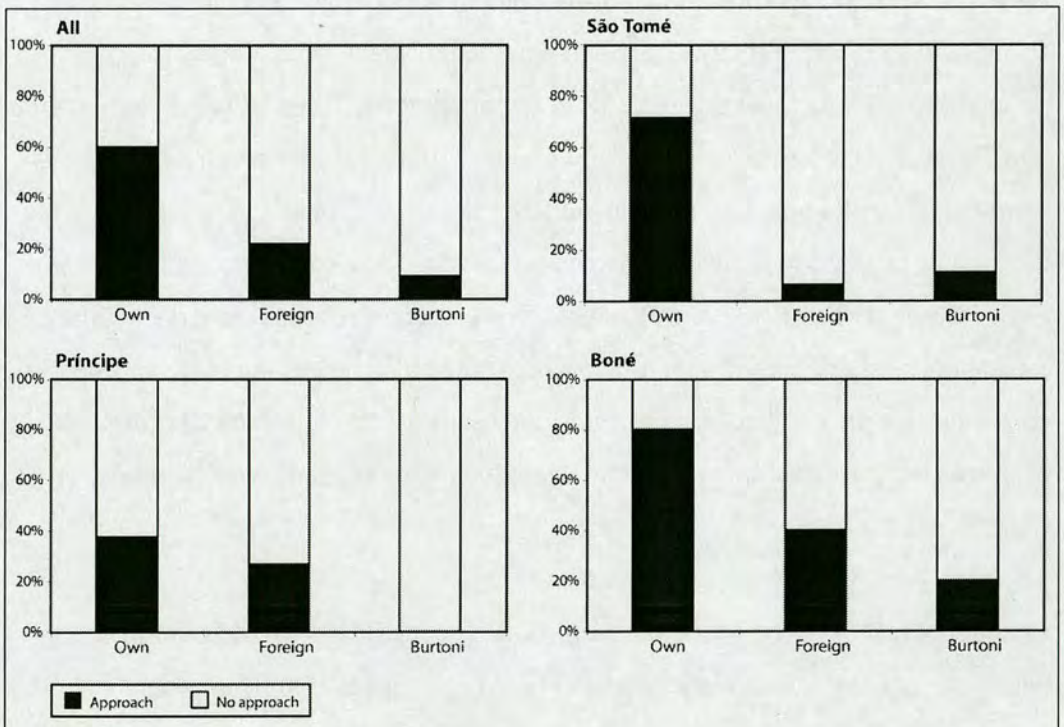


Fig. 7.1. Approach responses of *Serinus rufobrunneus* to playbacks of different categories: Own = song from the same population; Foreign = song from a different population; Burtoni = song from a different species, *S. burtoni*. Results are presented for each population and pooled together (All).

Table 7.1. P-values of the comparison of responses to the different song categories among *Serinus rufobrunneus* (chi-square tests for 'All'; Fisher's exact test for individual populations). Own = song from the same population; foreign = song from a different population; Burtoni = song from a different species, *S. burtoni*. Significant values in bold.

	Song categories compared		
Population	Own-Foreign-Burtoni	Own-Foreign	Foreign-Burtoni
All	**< 0.001**	**0.003**	0.200
São Tomé	**0.004**	**0.003**	1.000
Príncipe	0.214	0.657	0.257
Boné	0.182	0.282	0.600

DISCUSSION

Differences between populations

Playback experiments showed that birds from all populations responded more strongly to songs from their own population than to songs from allopatric populations. The fact that this pattern was significant only on São Tomé suggests that São Tomé birds discriminate more between the different song categories than birds from the other two populations. This may be the case for Boné, the population with the highest proportion of positive responses for each song category, but the low responsiveness of Príncipe birds makes interpretation of results for this population difficult. The significant results in São Tomé could also reflect the more appropriate conditions for the playback experiments: the four tests taken in each territory were performed with at least one-day intervals, reducing the possibility of 'stimulus-saturation'.

The higher responsiveness of birds from Boné is consistent with their unique behaviour in comparison with the other populations. Birds from Boné occur in high densities (Chapter 6, Fig. 6.1), and display a high level of tameness and curiosity for anything foreign: an observer can easily be surrounded by birds at c. 20 cm and, although not included in the experiments, the stimuli producing the strongest responses were sounds made by the observer. Therefore, and in accordance with theories of density compensation on islands, territoriality might have been lost in this

population (MacArthur *et al.* 1972). Song could nevertheless still play an important role in mate choice since the high densities suggest that breeding opportunities are limited and competition for mates is high. Despite their high overall curiosity, birds from Boné still responded more to their own song (four positive responses out of five tests), than to foreign or heterospecific songs. Additionally, birds from Boné might have evolved a mating system altogether different from the other populations. Individuals in Boné are closely related (as inferred from the low genetic diversity estimated from 15 microsatellites; Chapter 6) and move in large groups – a pattern consistent with either cooperative or polygamous breeding systems, for example. Differences in mating systems – where song would be just one of several elements – would likely constitute stronger reproductive barriers in case of secondary contact than differences in song alone.

In contrast to birds from Boné, birds from Príncipe were the least responsive of the three populations. This could indicate that Príncipe birds are less territorial than birds from the other populations, but it is more likely to indicate that the playback experiments were carried out in a period when territorial defence and/or mate attraction were not a priority. Several fledglings were observed, suggesting that the breeding season (or one breeding peak) was ending. Fledglings were also observed on Boné, but here the higher responses – even in non-breeding periods – may be explained by the inquisitive behaviour described above. Due to the weak response, experiments on Príncipe were not conclusive, and a new set should be performed earlier in the season.

Implications for speciation

The stronger response of birds to songs from their own populations than to foreign or heterospecific songs indicates that song has diverged to the point where it might act as a reproductive barrier in case of secondary contact. This inference was particularly supported in São Tomé, whereas songs from Príncipe and Boné populations elicited as little response as songs from a distinct species, *S. burtoni*. Therefore, in this system, divergence of mate recognition traits may play a role right at the start of the speciation process. These results conform with the classical model of allopatric

speciation where reproductive isolation arises as a by-product of adaptive changes in allopatry (Dobzhansky 1937; Mayr 1942; Rundle *et al.* 2005; Vines & Schluter 2006). A difference does exist, however, as those models postulate that reproductive barriers are themselves by-products of adaptive changes in other traits, whereas the evolution of song (the reproductive barrier) may not arise in this way. Although song divergence might be affected by adaptive changes in morphology (Podos 2001; Huber & Podos 2006), it is also driven directly by natural and sexual selection and by genetic and cultural drift (Chapter 6; Grant & Grant 1997a). Such process of gradual change of song in allopatry leading to reproductive isolation in sympatry has been demonstrated in different bird groups, and has been extensively documented in Darwin's finches (Grant & Grant 1997b, 1997c; Grant *et al.* 2000; Irwin *et al.* 2001; Grant & Grant 2002b).

The playback results do not guarantee that song will constitute a barrier upon secondary contact because they only test the behaviour of the receiver. Song learning that is responsible for the fast evolution of distinct geographic dialects might also allow immigrants to adapt their song to that of the host population, under the 'social adaptation hypothesis' (Payne 1981; Baptista 1985; Slabbekoorn & Smith 2002). For example, village indigo birds *Vidua chalybeata* are able to change their songs from year to year in order to match the songs of their neighbours (Payne 1985). The likelihood of this outcome will depend on when birds learn songs. Often, songs are imprinted in the first year of life, but this period can extend to the entire life in 'open-ended learners' (Catchpole & Slater 1995) – as is the case in indigobirds. Studies in captivity have shown that another *Serinus* species, the canary *Serinus canaria* is also an open-ended learner (Nottebohm & Nottebohm 1978; Nottebohm *et al.* 1986), suggesting therefore that the social adaptation hypothesis is likely to apply to *S. rufobrunneus*.

For assortative mating to be established upon secondary contact in *S. rufobrunneus*, a close linkage between mate recognition traits and population specific phenotypes must be present, and the phenotypic differences must be enough to allow exploitation of different resources (Kirkpatrick & Ravigné 2002) – and different resources must

be available. In such a situation, hybrids are expected to have low fitness, as their intermediate phenotypes will be less adapted than parental phenotypes in the exploitation of either of the parental niches. This will impose a cost on immigrants of learning the song of the host population. Therefore, low hybrid fitness could drive further the evolution of assortative mating by reinforcement, a process gaining increasingly theoretical and empirical support (Liou & Price 1994; Saetre *et al.* 1997; Coyne & Orr 1998; Nosil *et al.* 2003; Servedio & Noor 2003; Hoskin *et al.* 2005; Jiggins 2006). At the same time, resource-use phenotypic traits may also diverge further, through character displacement, as a way to reduce competitive interactions between the two phenotypes (Brown & Wilson 1956; Schluter & McPhail 1993). An alternative to this complex set of circumstances would be for full reproductive isolation to have already evolved in allopatry, which seems unlikely in this case. Despite conditions for speciation appearing restrictive, in the radiation of Darwin's finches the secondary contact stage has played a major role in pushing speciation to completion in many cases (Grant & Grant 2002a; Petren *et al.* 2005), in agreement with the 'archipelago speciation model' (Grant 2001; Rundle & Nosil 2005) discussed in Chapter 1, 6 and 8.

The outcome of a secondary contact in *S. rufobrunneus* will therefore be dependent on the evolution of sufficient morphological and song differences in allopatry (Grant & Grant 2002b) and on the availability of different resources for each phenotype. The three allopatric populations differ both in resource-use traits and mate recognition traits (Chapter 6), with song having diverged enough to be no longer recognised by birds from other populations. Whether the level of morphological divergence detected in this study is enough to allow the exploitation of distinct niches in sympatry is for now a matter of speculation. On Boné, co-existence is unlikely because there is only one resource, the oil palm *Elaeis guineensis*, capable of sustaining a viable population, and it is already being fully exploited by the native population. Príncipe, with its seedeater population mostly confined to the southern forests, offers a wide range of resources that are under-exploited or not exploited at all by its population, and therefore the ecological establishment of a diverging population would be dependent on inter- rather than intra-specific competition. São

Tomé, the largest island, might hold the answer to the likelihood of speciation after secondary contact, because its population of *S. rufobrunneus* co-exists with its sister species, the São Tomé grosbeak *Neospiza concolor* (Chapter 3). This suggests that interaction between diverging populations (the second stage of the 'archipelago speciation model') also played a very important role in speciation in this system (Chapter 3).

8

WHY ARE THERE SO MANY ENDEMIC SPECIES IN THE GULF OF GUINEA?

Molecular phylogenies have provided new insights into the relationships of the Gulf of Guinea birds. Levels of phenotypic differentiation had been wrongly assumed to reflect evolutionary time. In fact, on the oceanic islands the most divergent species were also the youngest. The resulting systematic rearrangements do not affect the total number of endemic species of the Gulf of Guinea system (33; 30 of which confined to the oceanic islands), but do affect the number of endemic genera: from six endemic genera, only one might be retained (*Amaurocichla*, for which no molecular data are yet available).

The high number of endemic bird species in the Gulf of Guinea islands is the result of speciation events rather than of the accumulation of relict species extinct on the mainland. Therefore, the Gulf of Guinea island system constitutes an important centre of bird speciation in Africa. Most of the speciation events followed the most 'trivial' route, i.e. full cessation of gene flow followed by independent divergence. This was due to the peculiar geographic situation of the islands, surrounded in the North and East by one of the richest centres of biodiversity in the world. As a result, the islands were colonised by several bird groups that, by occupying different niches, reduced the possibilities of radiations within the archipelago. Although 'trivial', this situation is very useful to test whether the current speciation paradigm – the 'ecological model' – can be generalised to speciation in allopatry. Data from the Príncipe seedeater *Serinus rufobrunneus* showed that selection rather than drift was the main driver of divergence in allopatry, thereby supporting the applicability of the ecological model of speciation to allopatric situations. The few species that speciated after establishing sympatry with related populations were also the most divergent, providing strong evidence for the importance of secondary contacts in promoting phenotypic diversification and speciation, and also supporting the latest version of the 'archipelago speciation model', whereby it is the small immigrant population that diverges rapidly upon secondary contact.

Overall, this study supported the view of speciation as a selection-driven process more likely to be completed in sympatry after an initial period of isolation. This is likely to constitute the most general model of speciation. As all general models have exceptions, it would be particularly interesting if a full sympatric origin for the São Tomé grosbeak *Neospiza concolor* could be demonstrated.

NEW INSIGHTS FROM MOLECULAR PHYLOGENIES

Molecular vs. phenotypic phylogenies

Molecular phylogenies are expected to constitute better hypotheses of evolutionary history than phenotype-based phylogenies. This is because the latter are often based on adaptive traits that can quickly lose their phylogenetic signal under selection. This study used molecular markers from the mitochondrial genome and from non-coding parts of the nuclear genome (microsatellites and introns). Albeit these markers may not be as neutral as previously thought – nuclear markers may be linked to regions under selection, and mitochondrial DNA may be under selection due to its role in the respiratory apparatus (Bazin *et al.* 2006) – they still provide phylogenetic inferences mostly independent from phenotypic traits (Bromham *et al.* 2002; Davies & Savolainen 2006). Furthermore, concordance between relationships inferred from nuclear and mitochondrial markers provides strong support for their phylogenetic utility, since selection would not affect all markers in the same way (Ballard & Whitlock 2004). In this study, mitochondrial and nuclear data were always used together and they always recovered the same general patterns – discrepancies were either due to the different resolution of the different markers or, as in Chapter 2, were likely to be associated with limitations of the nuclear dataset: in no case were the main conclusions affected. Therefore, in this thesis and in the following discussion, molecular phylogenies are considered to provide the best available hypotheses on the evolutionary history of the different groups. Discussion is centred on the three oceanic islands, where most speciation events took place.

Gulf of Guinea endemic birds: how many?

The molecular phylogenies inferred in this study provided a distinctly new picture for the history of the birds of the Gulf of Guinea (Tables 8.1 & 8.2). Previous systematic assessments exclusively based on phenotypic traits were misguided in two main ways: i) the most divergent (or 'aberrant') taxa were considered to be the oldest; and ii) 'aberrant' traits were thought to be shared by descent. This study showed both that 'aberrant' characters do not necessarily constitute synapormophies

(Chapter 2), and, most interestingly, that the level of phenotypic differentiation is not a function of time. On the contrary, on the oceanic islands the most divergent species were actually the youngest, as in the case of the endemic genera *Speirops* (Chapter 2) and *Neospiza* (Chapter 3). Phenotypic differentiation was therefore responsible for 'over-splitting', and this applies also to lesser levels of differentiation as in the *Alcedo* kingfishers (Chapter 5). The Gulf of Guinea thrush *Turdus olivaceofuscus* provided the exception to this trend: the two very distinct populations, almost certainly comprising two valid species, were formerly treated as subspecies only (Chapter 4). In a single case, molecular and phenotypic inferences differed for the opposite reason: little phenotypic differentiation corresponded to large genetic differences, suggesting a case of cryptic speciation (Príncipe white-eye *Zosterops ficedulinus*, Chapter 2). The new systematic arrangement will entail a revision of the global threat status (IUCN categories) for the new taxa as proposed on Table 8.1.

Table 8.1.

Taxonomic rearrangements based on the molecular phylogenies of this study, and preliminary proposals for change of IUCN Red list categories that they entail (changes highlighted in bold). IUCN Red List categories: - CR: critically endangered; EN: endangered; VU: vulnerable; nt: near threatened; lc: least concern. Current categories from (BirdLife International 2000).

TAXONOMY		IUCN	
Current	Rearrangement	Current	New
Alcedo leucogaster nais	*Alcedo cristata* [*nais*] [1]	lc	lc
Alcedo cristata thomensis	*Alcedo cristata* [*thomensis*] [1]	lc	lc
Turdus olivaceofuscus olivaceofuscus	*Turdus olivaceofuscus*	nt	**lc**
Turdus olivaceofuscus xanthorhynchus	*Turdus xanthorhynchus*	nt	**EN/VU**
Zosterops ficedulinus ficedulinus	*Zosterops* [*ficedulinus*] *ficedulinus* [2]	VU	VU
Zosterops ficedulinus feae	*Zosterops* [*ficedulinus*] *feae* [2]	VU	**lc**
Speirops lugubris	*Zosterops lugubris*	lc	lc
Speirops leucophaeus	*Zosterops leucophaeus*	VU	VU
Speirops brunneus	*Zosterops brunneus*	VU	VU
Speirops melanocephalus	*Zosterops melanocephalus*	VU	VU
Neospiza concolor	*Serinus concolor*	CR	CR

[1]: Maintenance of subspecific status is subjective: could be used to label phenotypic differentiation (Chapter 5)
[2]: Paraphyletic taxa likely to constitute distinct species but more data are needed. Their mitochondrial genetic divergence is 4.5% (corrected) and 3.7% (uncorrected).

Table 8.2. Revised list of the endemic bird species of the Gulf of Guinea island system. The likely source area for the oceanic island species is indicated (mainland or another island). In groups where inter-island dispersal is present, one of the species is randomly assigned a mainland origin. Subspecies of *Serinus rufobrunneus* accounted for (in brackets) as they are in the process of differentiation both in ecological and mate recognition traits.

SPECIES		SOURCE	
Common name	**Scientific name**	**Mainland**	**Island**
Dwarf olive ibis	*Bostrychia bocagei*	0-1[1]	0-1[1]
São Tomé green pigeon	*Treron sanctithomae*	0-1[2]	0-1[2]
São Tomé bronze-naped pigeon	*Columba malherbii*	1	0
Maroon pigeon	*Columba thomensis*	1	0
São Tomé scops owl	*Otus hartlaubi*	1	0
São Tomé spinetail	*Zoonavena thomensis*	1	0
São Tomé thrush	*Turdus olivaceofuscus*	1	0
Príncipe thrush	*Turdus xanthorhynchus*	0-1[3]	0-1[3]
São Tomé prinia	*Prinia molleri*	1	0
São Tomé short-tail	*Amaurocichla bocagii*	1	0
Annobón paradise flycatcher	*Terpsiphone smithii*	1	0
São Tomé paradise flycatcher	*Terpsiphone atrochalybeia*	1	0
Bioko batis	*Batis poensis*	-	-
Dohrn's thrush-babbler	*Horizorhinus dohrni*	1	0
Príncipe sunbird	*Anabathmis hartlaubii*	1	0
São Tomé sunbird	*Anabathmis newtonii*	1	0
Giant sunbird	*Dreptes thomensis*	1	0
Annobón white-eye	*Zosterops griseovirescens*	0	1
Príncipe white-eye	*Zosterops [ficedulinus] ficedulinus*	1	0
São Tomé white-eye	*Zosterops [ficedulinus] feae*	0	1
São Tomé speirops	*Zosterops lugubris*	0	1
Príncipe speirops	*Zosterops leucophaeus*	0	1
Fernando Po speirops	*Zosterops brunneus*	-	-
Mount Cameroon speirops	*Zosterops melanocephalus*	-	-
São Tomé fiscal	*Lanius newtoni*	1	0
São Tomé oriole	*Oriolus crassirostris*	1	0
Príncipe drongo	*Dicrurus modestus*	1	0
Príncipe glossy starling	*Lamprotornis ornatus*	1	0
Príncipe golden weaver	*Ploceus princeps*	1	0
Giant weaver	*Ploceus grandis*	1	0
São Tomé weaver	*Ploceus sanctithomae*	1	0
Príncipe seedeater	*Serinus rufobrunneus rufobrunneus*	1	0
	thomensis	0	(1)
	fradei	0	(1)
São Tomé grosbeak	*Serinus concolor*	0	1
	TOTAL	22-25	5-8 (+2)

Uncertainties - [1]: a subspecies of *B. olivacea*, to which *B. bocagei* is closely related, was formerly present in Príncipe. [2]: a subspecies of *T. australis*, to which *T. sanctithomae* is closely related, is present on Príncipe. [3]: cf. Chapter 4.

The systematic rearrangements proposed in this study do not affect the final number of endemic species in the region – which continues to be 33 if the species status for the two populations of *Zosterops ficedulinus* is accepted (Table 8.2). The status of some species may still change in the future: the three São Tomé endemic pigeons and the Príncipe drongo *Dicrurus modestus* may not be separate species from their mainland counterparts (Jones & Tye 2006), and the species status of the Bioko batis *Batis poensis* remains contested (Pérez del Val 1996; Borrow & Demey 2001). On the other hand the nominate race of the forest chestnut-winged starling *Onychognathus fulgidus*, endemic to São Tomé, may be a distinct species (Christy & Clarke 1998; Sinclair & Ryan 2003), and there is evidence that an unknown owl is present in the southern forests of Príncipe (Naurois 1975a; Melo 1998).

The systematic rearrangements do change, however, the number of endemic genera in the region. Until recently, six endemic genera were accepted: *Amaurocichla*, *Horizorhinus*, *Dreptes*, *Speirops*, *Thomasophantes* and *Neospiza* (Naurois 1994). All except *Speirops* are monotypic. The latest review of the birds of the region kept all endemic genera except for *Thomasophantes* (Jones & Tye 2006). This genus was placed in the large weaver genus *Ploceus* based on phenotypic evidence, a rearrangement recently confirmed by molecular data (Staffan Andersson, pers. comm.). The present study further showed that *Speirops* and *Neospiza* are synonymous with *Zosterops* and *Serinus* respectively. Preliminary data obtained for Dohrn's Thrush-babbler *Horizorhinus dohrni* (two cytochrome *b* sequences of 1043 nucletotides; M. Melo, unpublished) showed that its closest relative (from sequences available in GENBANK) is the African hill-babbler *Pseudoalcippe abyssinica*, from which it diverges by 6%. This is a cyt *b* divergence value typically found between congeneric bird species (Johns & Avise 1998), and therefore *Horizorhinus* is also unlikely to constitute a distinct genus. *Pseudoalcippe*, itself recently transferred to the large *Sylvia* genus (Cibois 2003), has populations both in the mountains of Bioko and on Mt. Cameroon (Borrow & Demey 2001). Preliminary molecular data also indicate that the sunbird genus *Dreptes* does not warrant distinct generic status (Rauri Bowie, pers. comm.). Molecular data are missing only for the São Tomé short-tail *Amaurocichla bocagii.* Due to its unique and primitive features, like the

protruding rachis of the tail feathers, the affinities of this species are unknown and it may well be the only species to be maintained in its monotypic genus. Therefore, from an initial complement of six endemic genera, only one genus may remain valid. A similar result was found recently among the endemic frogs. The São Tomé and Príncipe endemic genus of hyperoliid treefrogs *Nesionixalus* (Hyperoliidae) has been shown to nest within the larger *Hyperolius* genus (Drewes & Wilkinson 2004).

Old islands, young species

An unexpected pattern revealed by the molecular phylogenies is that, because the 'aberrant' species do not represent old lineages, most endemic species of the Gulf of Guinea islands likely represent recent speciation events, i.e., events that occurred from the late Pliocene (2.5 Ma ago) onwards (Fjeldså 1994; Fjeldså & Lovett 1997). In this study, *Turdus olivaceofuscus* (Chapter 4) might be the only exception as indicated by the large genetic divergence between the two populations and between these and their mainland relative. *Amaurocichla bocagii* may constitute another exception. Even without precise estimates of speciation times, it is safe to consider that the present species are all much more recent than the islands they inhabit. This pattern indicates therefore that the Gulf of Guinea constitutes a speciation centre rather than simply a refuge for species that went extinct on the mainland. The absence of old species is nevertheless surprising, and could be due to the very active volcanic history of the islands until the last 0.1 Ma (Lee *et al.* 1994), to extinctions driven by the rise in sea-levels after glacial periods, or reflect a high turn-over of species due to the relative proximity of the islands to the highly diverse Guinean forests, or most likely to a combination of these factors. The patterns of genetic variation of the lizard *Mabuya maculilabris* within São Tomé have been linked to the impact of volcanic activity, with the extent of this variation being much lower than expected from the age of the island (Jesus *et al.* 2005c). The surface areas of the oceanic islands could increase significantly with sea-level changes during glacial periods, (Jones & Tye 2006). During the last glaciation for example, c. 10,000 years ago, Príncipe and Annobón were almost 10 times as larger than at present (Fig. 1.2). This is likely to have led to extinctions, and therefore affects the interpretation of the evolutionary history of the birds of the region based on the current extant species.

SPECIATION IN THE GULF OF GUINEA

The importance of geography

To understand the origin and patterns of bird diversity in the Gulf of Guinea one must first acknowledge the very important role that the geographic situation of the islands has had in shaping the speciation process. The linear arrangement of the islands and the relatively large distances between them, which are similar to the distances to the mainland, favours independent colonisations from the species-rich mainland over dispersal events between islands. From the 30 species endemic to the three oceanic islands only 5 - 8 are derived from inter-island dispersal events (Table 8.2). Two additional inter-island dispersal events are responsible for the differentiation of the Príncipe seedeater *Serinus rufobrunneus* (Chapter 6), but the speciation process here is probably not yet completed. Colonisation by so many mainland species removes one of the major conditions necessary for radiation to take place, i.e., low interspecific competition (Schluter 2000). The 30 endemic species belong to 16 or 17 different families, and this diversity must certainly reduce the opportunities for diverging populations that meet in sympatry to speciate by adapting to different niches. This is supported by the fact that between 7 and 11 inter-island dispersal events of endemic species or subspecies did not result in speciation (Table 8.3).

'Passive speciation': allospeciation

The geographic situation of the Gulf of Guinea is such that although colonisations from the mainland are more likely than in other systems where endemic birds are present, distances to the mainland are still large enough to allow colonisers to be isolated from their source population. Speciation in this system will then have proceeded in many instances by the longest path, i.e., independent divergence in allopatry, or allospeciation (Mayr & Diamond 2001). In other words, only the first part of the 'archipelago radiation model' took place. This model would apply not only to all the species descended directly from independent mainland colonisations but also to colonisations of a new island from another one, as was the case with the

white-eyes. Allospeciation explains why, for many of these endemic species, phenotypic differentiation in relation to their counterparts is small. It has been shown that the allopatric phase of speciation is often characterised by phenotypic stasis, and that secondary contact between diverging populations is what triggers rapid phenotypic change (Petren *et al.* 2005). As was shown in the study of the Príncipe seedeater, divergence of allopatric populations, even if small, is driven by selection rather than drift. The same results were described for Galápagos warbler finches (Tonnis *et al.* 2005). In the Gulf of Guinea, the similarity between the mainland and island habitats (moist lowland and montane rainforest) is consistent with the generally low levels of phenotypic differentiation observed in allospecies (Table 8.2). *Amaurocichla bocagii* and *Turdus olivaceofuscus* constitute two exceptions likely to be explained by their older age (see above). Dohrn's thrush-babbler *Horizorhinus dohrni* may be the only exception of a species that, although apparently originated by allospeciation, shows a high degree of phenotypic differentiation. Until the preliminary results in this study placed it next to *Sylvia* [*Pseudoalcippe*] *abyssinica* its affinities, based on morphological traits, were uncertain.

Table 8.3.

Estimation of within-archipelago dispersal events for Gulf of Guinea endemic species and subspecies that did not lead to speciation. The source area for each population is indicated (mainland or another island). In groups with inter-island dispersal, one population is randomly assigned a mainland origin. ST: São Tomé, P: Príncipe, A: Annobón.

SPECIES		Distribution	SOURCE	
Common name	**Scientific name**		**Mainland**	**Island**
ST bronze-naped pigeon	*Columba malherbii*	ST, P, A	1	2
Lemon dove	*Apoplelia larvata simplex*	ST	1	0
	principalis	P	0-1	0-1
	hypoleuca	A	0-1	0-1
Emerald cuckoo	*Chrysococcyx cupreus insularum*	ST, P, A [1]	1-2	1-2
São Tomé spinetail	*Zoonavena thomensis*	ST, P	1	1
Little Swift	*Apus affinis bannermani*	ST, P	1	1
Malachite kingfisher	*Alcedo cristata (nais)*	P	1	0
São Tomé kingfisher	*Alcedo cristata (thomensis)*	ST	0-1	0-1
Príncipe seedeater	*Serinus r. rufobrunneus*	P	1	0
	thomensis	ST	0	1
	fradei	Boné	0	1
		TOTAL	7-11	7-11

[1]: There are only a few records of Emerald cuckoo in Annobón: and it is not known if they belong to ssp *insularum*.

'Active speciation': speciation in sympatry

The fact that the most divergent species in the oceanic islands are generally the youngest and derived from instances of secondary contact between diverging populations provides ample support for the importance of a sympatric phase in accelerating the diversification and speciation process, as postulated by the 'archipelago speciation model' (Lack 1947; Grant 2001).

Apart from *Amaurocichla* and *Horizorhinus* discussed above, all other highly divergent species are associated with secondary contact between diverging populations (this study) or at least between related species (for those species for which molecular data are lacking). This applies therefore to the following species pairs: the sunbirds *Dreptes thomensis* and *Anabathmis newtonii*; the weavers *Thomasophantes sanctithomae* and *Ploceus grandis*; the finches *Neospiza concolor* and *Serinus rufobrunneus*, and the two pairs of white-eyes *Speirops – Zosterops*. Nevertheless, *D. thomensis* and *A. newtonii* and *T. sanctithomae* and *P. grandis* may not represent cases of speciation after secondary contact, because preliminary molecular data indicate that they do not form sister species pairs (Rauri Bowie and Staffan Andersson, pers comm., respectively). In this case, competitive interactions between the related species in sympatry could still have driven phenotypic divergence, without being responsible for the speciation process – in an example of character displacement between species such as the one described in Hawaiian myzomelid honeyeaters (Diamond *et al.* 1989). *P. grandis* is nevertheless sister to the village weaver *P. cucullatus* (Staffan Andersson, pers. comm.), a species with a wide Afrotropical distribution, with which it co-occurs on São Tomé. The giant weaver could therefore have originated after secondary contact of the two diverging sister populations.

This thesis has provided strong evidence supporting the completion of the speciation process after diverging populations meet in sympatry, both for the *Zosterops* radiation and for the origin of the São Tomé grosbeak *Neospiza concolor* (Chapters 2 & 3). In the case of the São Tomé Speirops *S. lugubris* and for *N. concolor*,

phylogenetic evidence could not distinguish whether speciation happened fully in sympatry or if an allopatric phase had taken place before secondary contact. Whatever the reality, a sympatric phase was almost certainly implicated in the speciation process. Data from these two groups showed that the most 'aberrant' species (*Neospiza* and *Speirops* spp) represent the most recent speciation events, rather than old lineages as previously assumed. This is a particularly interesting result as it is concordant with the most recent refinement of the 'archipelago radiation model' (Petren *et al.* 2005). According to this proposal, little phenotypic divergence evolves in allopatric populations yet, once secondary contact is established, it is the immigrant population that changes rapidly, with the phenotype of the resident population remaining relatively unchanged. This new twist to the 'archipelago radiation model' is based on empirical data from many species of Darwin's finches (Petren *et al.* 2005), but the process of asymmetric divergence due to resource competition in sympatry had already received theoretical support (Doebeli & Dieckmann 2000). Such a process was shown to be more likely when a small population invades a larger, locally adapted, resident population – a situation that well describes the archipelago situation. What this revision also implies – but did not state, presumably because of its triviality – is that secondary contact must be made by a reasonably sized flock (i.e., a population) for this model to work. This may explain why the Zosteropidae were the only group that was able to radiate in this 'radiation-averse' archipelago: their gregarious behaviour allowing them to colonise islands in large flocks (Clegg *et al.* 2002a).

IN SUMMARY...

The high number of endemic bird species in the Gulf of Guinea islands is the result of speciation events rather than of the accumulation of relict species extinct on the mainland. On the contrary, most species are of recent origin. In this sense the Gulf of Guinea island system constitutes a very important centre of bird speciation in Africa, with almost half of the endemic bird species of the Guinean Forests' hotspot, which includes both the Upper and Lower Guinea forests (Bakkar *et al.* 1999). Most of the speciation events followed the most 'trivial' route: full cessation of gene flow followed by independent divergence. This was due to the unusual geographical situation of the islands that allowed colonisation by several bird groups, therefore reducing the possibilities of radiations within the archipelago. Data from the Príncipe seedeater *Serinus rufobrunneus* showed that selection rather than drift was the main driver of divergence in allopatry, thereby supporting the applicability of the ecological model of speciation to allopatric situations. The few species that speciated after establishing sympatry with related populations were also the most divergent, providing strong evidence for the importance of secondary contacts in promoting phenotypic diversification and speciation, and also supporting the latest version of the 'archipelago radiation model', whereby it is the small immigrant population that diverges rapidly upon secondary contact (Petren *et al.* 2005).

This study supported the view of speciation as a selection-driven process (the ecological model), which proceeds faster following the sequence of events described by the 'archipelago radiation model': sympatric interactions are the main driver of speciation but only after some period of isolation. The combination of the ecological and the archipelago models is likely to constitute the most general model of speciation, even in mainland situations, and has been designated as the 'two stage speciation model' (Grant & Grant 1997a; Rundle & Nosil 2005). As all general models have exceptions, I suggest that the São Tomé grosbeak may be the most likely candidate to have had, (almost) uniquely among birds, a full sympatric origin.

References

Akaike H (1973) Information theory as an extension of the maximum likelihood principle. In: *Second International Symposium on Information* (eds. Theory Petrov BN, Csaki F). Akademiai kiado, Budapest.

Alatalo RV, Gustafsson L, Lundberg A (1986) Interspecific competition and niche changes in tits (*Parus* spp.): evaluation of nonexperimental data. *American Naturalist,* **127**, 819-834.

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

Amadon D (1965) Position of the genus *Neospiza* Salvadori. *Ibis,* **107**, 395-396.

Andersson S, Prager M (2006) Quantification of coloration. In: *Bird Coloration. Vol. 1: Mechanisms and Measurements* (eds. Hill GE, McGraw KJ), pp. 41-89. Harvard University Press, Cambridge.

Arbogast BS, Drovetski SV, Curry RL, *et al.* (2006) The origin and diversification of Galapagos mockingbirds. *Evolution,* **60**, 370-382.

Arnaiz-Villena A, Álvarez-Tejado M, Ruíz-del-Valle V, *et al.* (1999) Rapid radiation of canaries (Genus *Serinus*). *Molecular Biology and Evolution,* **16**, 2-11.

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

Atkinson P, Peet N, Alexander J (1991) The status and conservation of the endemic bird species of São Tomé and Príncipe, West Africa. *Bird Conservation International,* **1**, 255-282.

Avise JC (1989) A role for molecular genetics in the recognition and conservation of endangered species. *Trends in Ecology and Evolution,* **4**, 279-281.

Avise JC, Ball RMJ (1990) Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology,* **7**, 45-67.

Badyaev AV, Young RL (2004) Complexity and integration in sexual ornamentation: an example with carotenoid and melanin plumage pigmentation. *Journal of Evolutionary Biology,* **17**, 1317-1327.

Baillie JM, Gascoigne A (1999) Extrinsic and intrinsic factors associated with the present conservation status of the endemic birds of Príncipe Island (Gulf of Guinea). *Unpublished manuscript.*

Baker AJ, Jenkins PF (1987) Founder effect and cultural evolution of songs in an isolated population of Chaffinches, *Fringilla coelebs*, in the Chatham Islands. *Animal Behaviour,* **35**, 1793-1803.

Baker AJ, Marshall HD (1997) Mitochondrial control region sequences as tools for understanding evolution. In: *Avian Molecular Evolution and Systematics* (ed. Mindell DP), pp. 51-82. Academic Press, San Diego.

Baker MC (1996) Depauperate meme pool of vocal signals in an island population of singing Honeyeaters. *Animal Behaviour,* **51**, 853-858.

Baker MC (2001) Bird song research: the past 100 years. *Bird Behavior,* **14**, 3-50.

Baker MC (2006) Differentiation of mating vocalizations in birds; acoustic features in mainland and island populations and evidence of habitat-dependent selection on songs. *Ethology,* **112**, 757-771.

Baker MC, Baker EM (1990) Reproductive behavior of female buntings: isolating mechanisms in a hybridizing pair of species. *Evolution,* **44**, 332-338.

Baker MC, Baker EM, Baker MSA (2001) Island and island-like effects on vocal repertoire of singing Honeyeaters. *Animal Behaviour,* **62**, 767-774.

Baker MC, Baker MSA, Baker EM (2003) Rapid evolution of a novel song and an increase in repertoire size in an island population of an Australian songbird. *Ibis,* **145**, 465-471.

Baker MC, Cunningham MA (1985) The biology of bird-song dialects. *Behavioral and Brain Sciences,* **8**, 85-133.

Bakkar ML, Bailey B, Myers N, *et al.* (1999) Guinean Forests. In: *Hotspots: Earth's Richest and Most Endangered Terrestrial Ecoregions* (eds. Mittermeier RA, Myers N, Robles Gil PR, Mittermeier CG), pp. 238-253. CEMEX, Mexico City.

Balinsky JB (1981) Adaptation of nitrogen metabolism to hyperosmotic environment in Amphibia. *Journal of Experimental Zoology,* **215**, 335-350.

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

Balloux F, Goudet J (2002) Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. *Molecular Ecology,* **11**, 771-783.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology,* **11**, 155-165.

Bannerman DA (1953) *The Birds of West and Equatorial Africa.* Oliver & Boyd, Edinburgh.

Baptista LF (1985) Bird-song dialects: social adaptation or assortative mating? *Behavioral and Brain Sciences,* **8**, 100-101.

Baptista LF, Trail PW (1992) The role of song in the evolution of passerine diversity. *Systematic Biology,* **41**, 242-247.

Barker FK, Lutzoni FM (2002) The utility of the incongruence length difference test. *Systematic Biology,* **51**, 625-637.

Barraclough TG, Nee S (2001) Phylogenetics and speciation. *Trends in Ecology and Evolution,* **16**, 391-399.

Barraclough TG, Vogler AP (2000) Detecting the geographical pattern of speciation from species-level phylogenies. *American Naturalist,* **155**, 419-434.

Barrowclough GF (1983) Biochemical studies of microevolutionary processes. In: *Perspectives in Ornithology* (eds. Brush AH, Clark Jr GA). Cambridge University Press, Cambridge.

Bartlein PJ, Prentice IC (1989) Orbital variations, climate and paleoecology. *Trends in Ecology and Evolution,* **4**, 195-199.

Barton NH (1996) Natural selection and random genetic drift as causes of evolution on islands. In: *Evolution on Islands* (ed. Grant PR), pp. 102-123. Oxford University Press, Oxford.

Barton NH (2001) Speciation. *Trends in Ecology and Evolution,* **16**, 325.

Barton NH, Charlesworth B (1984) Genetic revolutions, founder effects, and speciation. *Annual Review of Ecology and Systematics,* **15**, 133-164.

Barton NH, Jones JS (1983) Mitochondrial DNA: new clues about evolution. *Nature,* **306**, 317-318.

Bazin E, Glémin S, Galtier N (2006) Population size does not influence mitochondrial genetic diversity in animals. *Science,* **312**, 570-572.

Beerli P (2004) MIGRATE: documentation and program, part of LAMARC. Version 2.1.3. Distributed over the internet: http://evolution.gs.washington.edu/lamarc.html.

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, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences USA,* **98**, 4563-4568.

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

Bensch S, Helbig AJ, Salomon M, Seibold I (2002) Amplified fragment length polymorphism analysis identifies hybrids between two subspecies of warblers. *Molecular Ecology,* **11**, 473-481.

Berlocher SH (1998) Origins: a brief history of research on speciation. In: *Endless Forms: Species and Speciation* (eds. Howard DJ, Berlocher SH). Oxford University Press, New York.

BirdLife International (2000) *Threatened Birds of the World.* Lynx Edicions and BirdLife International, Barcelona and Cambridge.

Blom A, Schipper J (2004) Mount Cameroon and Bioko montane forests. In: *Terrestrial Ecoregions of Africa and Madagascar: a Conservation Assessment* (eds. Burgess N, D' Amico Hales J, Underwood E, *et al.*), pp. 232-234. Island Press, Washington.

Blondel J (2000) Evolution and ecology of birds on islands: trends and prospects. *Vie et Milieu,* **50**, 205-220.

Bocage JVB (1888) Oiseaux nouveaux de l'île St. Thomé. *Jornal de Sciencias Mathematicas, Physicas e Naturaes de Lisboa,* **12**, 229-232.

Bocage JVB (1904) Contribution à la faune des quatre îles du Golfe de Guinée. *Jornal de Sciencias Mathematicas, Physicas e Naturaes de Lisboa - Série 2,* **7**, 65-96.

Bock WJ (1963) Evolution and phylogeny in morphologically uniform groups. *American Naturalist,* **97**, 265-285.

Bock WJ (1967) The use of adaptive characters in avian classification. *Proceedings of the XIV International Ornithological Congress*, 61-74.

Bolnick DI (2004) Waiting for sympatric speciation. *Evolution,* **58**, 895-899.

Borrow N, Demey R (2001) *Birds of Western Africa.* Christopher Helm, London.

Boughman JW (2002) How sensory drive can promote speciation. *Trends in Ecology and Evolution,* **17**, 571-577.

Bowie RCK, Bloomer P, Clancey PA, Crowe TM (2003) The Karoo Thrush (*Turdus smithi* Bonaparte 1850), a southern African endemic. *Ostrich,* **74**, 1-7.

Bowie RCK, Voelker G, Fjeldså J, *et al.* (2005) Systematics of the olive thrush *Turdus olivaceus* species complex with reference to the taxonomic status of the endangered Taita thrush *T. helleri*. *Journal of Avian Biology,* **36**, 391-404.

Braaten RF, Reynolds K (1999) Auditory preferences for conspecific song in isolation-reared zebra-finches. *Animal Behaviour,* **58**, 105-111.

Brainard MS, Doupe AJ (2002) What songbirds teach us about learning. *Nature,* **417**, 351-358.

Bray JR, Curtis JT (1957) An ordination of the upland forest communities of Southern Wisconsin. *Ecological Monographs,* **27**, 325-349.

Bremer K (1992) Ancestral areas: a cladistic reinterpretation of the center of origin concept. *Systematic Biology,* **41**, 436-445.

Bromham L, Woolfit M, Lee MSY, Rambaut A (2002) Testing the relationship between morphological and molecular rates of change along phylogenies. *Evolution,* **56**, 1921-1930.

Brown WLJ, Wilson EO (1956) Character displacement. *Systematic Zoology,* **5**, 49-64.

Brumfield RT, Jerningan RW, McDonald DB, Braun MJ (2001) Evolutionary implications of divergent clines in an avian (*Manacus*: Aves) hybrid zone. *Evolution,* **55**, 2070-2087.

Bull JJ, Huelsenbeck JP, Cunningham CW, Swofford DL, Waddell PJ (1993) Partitioning and combining data in phylogenetic analysis. *Systematic Biology,* **42**, 384-397.

Burke K (2001) Origin of the Cameroon Line of volcano-capped swells. *Journal of Geology,* **109**, 349-362.

Burns KJ (1998) Molecular phylogenetics of the genus *Piranga*: implications for biogeography and the evolution of morphology and behavior. *Auk,* **115**, 621-634.

Bush G (1969) Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera, Tephritidae). *Evolution,* **23**, 237-251.

Carson HL, Clague DA (1995) Geology and biogeography of the Hawaiian Islands. In: *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago* (eds. Wagner WL, Funk VA), Washington.

Carson HL, Templeton AR (1984) Genetic revolutions in relation to speciation phenomena. *Annual Review of Ecology and Systematics,* **15**, 97-131.

Case TJ, Sidell R (1983) Pattern and chance in the structure of model and natural communities. *Evolution,* **37**, 832-849.

Catchpole CK, Slater PJB (1995) *Bird Song: Biological Themes and Variations.* Cambridge University Press, Cambridge.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics,* **19**, 233-257.

Chapuis MP, Estoup A (in press) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution.*

Cheke RA, Mann CF (2001) *Sunbirds: a Guide to the Sunbirds, Flowerpeckers, Spiderhunters and Sugarbirds of the World.* Christopher Helm, London.

Chesser RT (1999) Molecular systematics of the rhinocryptid genus *Pteroptochos. Condor,* **101**, 439-446.

Chirhart SE, Honeycutt RL, Greenbaum IF (2005) Microsatellite variation and evolution in the *Peromyscus maniculatus* species group. *Molecular Phylogenetics and Evolution,* **34**, 408.

Christy P (2001) São Tomé and Príncipe. In: *Important Bird Areas in Africa and Associated Islands: Priority Sites for Conservation* (eds. Fishpool LDC, Evans MI), pp. 727-731. Pisces Publications and BirdLife International, Newbury and Cambridge.

Christy P, Clarke WV (1998) *Guide des Oiseaux de São Tomé e Príncipe.* ECOFAC, São Tomé.

Cibois A (2003) Mitochondrial DNA phylogeny of babblers (Timaliidae). *Auk,* **120**, 33-54.

Cicero C, Johnson NK (2001) Higher-level phylogeny of New World Vireos (Aves: Vireonidae) based on sequences of multiple mitochondrial DNA genes. *Molecular Phylogenetics and Evolution,* **20**, 27-40.

Clarke KR, Warwick RM (2001) *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation.* PRIMER-E, Ltd, Plymouth.

Clegg SB, Owens IPF (2002) The 'island rule' in birds: medium body size and its ecological explanation. *Proceedings of the Royal Society of London B,* **269**, 1359-1365.

Clegg SM, Degnan SM, Kikkawa J, *et al.* (2002a) Genetic consequences of sequential founder events by an island-colonizing bird. *Proceedings of the National Academy of Sciences USA,* **99**, 8127-8132.

Clegg SM, Degnan SM, Moritz C, *et al.* (2002b) Microevolution in island forms: the roles of drift and directional selection in morphological divergence of a passerine bird. *Evolution,* **56**, 2090-2099.

Clement P, Harris A, Davis J (1993) *Finches and Sparrows.* Christopher Helm, London.

Clement P, Hathway R (2000) *Thrushes: Helm Identification Guides.* Christopher Helm, London.

Collar N (1997) Taxonomy and conservation. *Bulletin of the British Ornithologists' Union,* **117**, 122-136.

Collar N, Stuart SN (1988) *Key Forests for Threatened Birds In Africa.* International Council for Bird Preservation & IUCN, The World Conservation Union, Cambridge.

Cox TF, Cox MAA (1994) *Multidimensional Scaling.* Chapman & Hall, London.

Coyne JA (1994) Ernst Mayr and the origin of species. *Evolution,* **48**, 19-30.

Coyne JA, Orr HA (1998) The evolutionary genetics of speciation. *Philosophical Transactions of the Royal Society of London B,* **353**, 287-305.

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

Coyne JA, Price TD (2000) Little evidence for sympatric speciation in island birds. *Evolution,* **54**, 2166-2171.

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

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution,* **15**, 290-295.

Crochet P-A, Chen JZ, Pons J-M, *et al.* (2003) Genetic differentiation at nuclear and mitochondrial loci among large white-headed gulls: sex-biased interspecific gene flow? *Evolution,* **57**, 2865-2878.

Crowe TM, Ryan PG, Essop MF, *et al.* (1994) Species as the 'currency' of conservation: The Karoo/Dune/Red Lark complex of south-western Africa. In: *Systematics and Conservation Evaluation* (eds. Forey PL, Humphries CJ, Vane-Wright RI), pp. 229-234. Claredon Press, London.

Cunningham CW (1997) Is congruence between data partitions a reliable predictor of phylogenetic accuracy? Empirically testing an iterative procedure for choosing among phylogenetic methods. *Systematic Biology,* **46**, 464-478.

Dallimer M, King T, Leitão P (2003) New records of the São Tomé Grosbeak *Neospiza concolor. Bulletin of the African Bird Club,* **10**, 23-25.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection.* John Murray, London.

Darwin C (1860) *Journal of Researches Into the Natural History and Geology of the Countries Visited During the Voyage Round the World of H.M.S. Beagle.* John Murray, London.

Daugherty CH, Cree A, Hay JM, Thompson MB (1990) Neglected taxonomy and continuing extinctions of tuatara (*Sphenodon*). *Nature,* **347**, 177-179.

Davies TJ, Savolainen V (2006) Neutral theory, phylogenies and the relationship between phenotypic change and evolutionary rates. *Evolution,* **60**, 476-483.

Dawson DA BIRDMARKER homepage. NERC Sheffield Molecular Genetics Facility. http://www.shef.ac.uk/misc/groups/molecol/deborah-dawson-birdmarkers.html.

de Queiroz K (1998) The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In: *Endless Forms: Species and Speciation* (eds. Howard DJ, Berlocher SH), pp. 57-75. Oxford University Press, New York.

de Queiroz K (2005) Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences USA,* **102**, 6600-6607.

Degnan SM (1993) The perils of single gene trees - mitochondrial versus single-copy nuclear DNA variation in white-eyes (Aves: Zosteropidae). *Molecular Ecology,* **2**, 219-225.

Degnan SM, Robertson BC, Clegg SB, Moritz C (1999) Microsatellite primers for studies of gene flow and mating systems in white-eyes (Zosterops). *Molecular Ecology,* **8**, 159-160.

deMenocal PB (1995) Plio-Pleistocene African climate. *Science,* **270**, 53-58.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B,* **39**, 1-38.

DeSalle R, Absher R, Amato G (1994) Speciation and phylogenetic resolution. *Trends in Ecology and Evolution,* **9**, 297-298.

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.

Diamond AW (1984) Biogeography of Seychelles land birds. In: *Biogeography and Ecology of the Seychelles* (ed. Stoddart DR), pp. 487-504. W Junk, The Hague.

Diamond JM (1973) Distributional ecology of New Guinea birds. *Science,* **179**, 759-769.

Diamond JM (1981) Flightlessness and fear of flying in island species. *Nature,* **293**, 507-508.

Diamond JM, Pimm SL, Gilpin ME, LeCroy M (1989) Rapid evolution of character displacement in Myzomelid honeyeaters. *American Naturalist,* **134**, 675-708.

Dickinson EC, ed. (2003) *The Howard and Moore Complete Checklist of the Birds of the World.* Third edn. Princeton University Press, Princeton.

Dieckmann U, Doebeli M (1999) On the origin of species by sympatric speciation. *Nature,* **400**, 354-357.

Dillon S, Fjeldså J (2005) The implications of different species concepts for describing biodiversity patterns and assessing conservation needs for African birds. *Ecography,* **28**, 682-692.

Dobzhansky T (1937) *Genetics and the Origin of Species.* Columbia University Press, New York.

Doebeli M, Dieckmann U (2000) Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *American Naturalist,* **156**, S77-S101.

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

Drewes RC, Stoelting RE (2004) The California Academy of Sciences Gulf of Guinea Expedition (2001) II. Additions and corrections to our knowledge of the endemic amphibians of São Tomé and Príncipe. *Proceedings of the California Academy of Sciences,* **55**, 573-587.

Drewes RC, Wilkinson JS (2004) The California Academy of Sciences Gulf of Guinea São Tomé and Príncipe Expedition (2001) I. The taxonomic status of the genus *Nesionixalus* Perret, 1976 (Anura: Hyperollidae), treefrogs of São Tomé and Príncipe, with comments on the genus *Hyperolius*. *Proceedings of the California Academy of Sciences,* **55**, 393-405.

Drummond AJ, Rambaut A (2003) BEAST. Available from http://evolve.zoo.ox.ac.uk.

Dupont LM, Jahns S, Marret F, Ning S (2000) Vegetation change in equatorial West Africa: time slices for the last 150 ka. *Paleogeography, Paleoclimatology, Paleoecology,* **155**, 95-122.

Dutton J, Haftt J (1996) Distribution, ecology and status of an endemic shrew, *Crocidura thomensis*. *Oryx,* **30**, 195-201.

Eck S (1995) Ergäzendes zu morphologie und systematik der *Speirops*-arten (Zosteropidae). *Mitteilungen aus dem Zoologischen Museum Berlin,* **71 Suppl Ann. Orn. 19**, 101-107.

Edwards SV (1997) Relevance of microevolutionary processes to higher level molecular systematics. In: *Avian Molecular Evolution and Systematics* (ed. Mindell DP), pp. 251-278. Academic Press, San Diego.

Edwards SV, Kingan SB, Calkins JD, *et al.* (2005) Speciation in birds: genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences USA,* **102 Suppl. 1**, 6550-6557.

Einsentraut M (1965) Rassenbildung bei Saugetiernen und Vogeln auf der Insel Fernando Poo. *Zoologischer Anzeiger,* **174**, 37-53.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics,* **5**, 435-445.

Ellegren H (2005) The avian genome uncovered. *Trends in Ecology and Evolution,* **20**, 180-186.

Ellis JR, Pashley CH, Burke JM, McCauley DE (2003) High genetic diversity in a rare and endangered sunflower as compared to a common congener. *Molecular Ecology,* **15**, 2345-2356.

Emerson BC (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology,* **11**, 951-966.

Endler JA (1982) Pleistocene forest refuges: Fact or fancy? In: *Biological Diversification in the Tropics* (ed. Prance GT), pp. 641-657. Columbia University Press, New York.

Endler JA (1990) On the measurement and classification of colour in studies of animal colour patterns. *Biological Journal of the Linnean Society,* **41**, 315-352.

Endler JA (1993) The color of light in forests and its implications. *Ecological Monographs,* **63**, 1-27.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1**, 47-50.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics,* **131**, 479-491.

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

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.

Fa JE, Juste BJ (1994) Biodiversity conservation in the Gulf of Guinea islands. *Biodiversity and Conservation,* **3**, 757-758.

Falush D, Stephens M, Pritchard JK (2003) Infrence of population structure: Extensions to linked loci and correlated allele frequencies. *Genetics,* **164**, 1567-1587.

Farris JS, Källersjö M, Kluge AG, Bult C (1995) Constructing a significance test for incongruence. *Systematic Biology,* **44**, 570-572.

Feder JL, Berlocher SH, Roethele JB, *et al.* (2003) Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences USA,* **100**, 10314-10319.

Feiler A, Nadler T (1992) Über evolutive beziehungen zwischen den brillenvögeln der Gattung *Speirops* (Aves, Zosteropidae) von West-Kamerun und der Inseln im Golf von Guinea. *Bonner Zoologische Beiträge,* **43**, 423-432.

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

Felsenstein J (1981b) Skepticism towards Santa Rosalia, or Why are there so few kinds of animals? *Evolution,* **35**, 124-138.

Felsenstein J (1985) Confidence limits of phylogenies: an approach using the bootstrap. *Evolution,* **39**, 783-791.

Felsenstein J (2004) *Inferring Phylogenies.* Sinauer Associates, Sunderland, MA.

Ferris SD, Sage RD, Huang C-M, *et al.* (1983) Flow of mitochondrial DNA across a species boundary. *Proceedings of the National Academy of Sciences USA,* **80**, 2290-2294.

Figueiredo E (1994) Diversity and endemism of angiosperms in the Gulf of Guinea islands. *Biodiversity and Conservation,* **3**, 785-793.

Figueiredo E (1998) The Pteridophytes of São Tomé e Príncipe. *Bulletin of the Natural History Museum, London (Botany),* **28**, 41-66.

Fishpool LDC, Evans MI, eds. (2001) *Important Bird Areas in Africa and Associated Islands: Priority Sites for Conservation.* Pisces Publications and BirdLife International, Newbury and Cambridge.

Fitton JG (1980) The Benue trough and Cameroon line - a migrating rift system in West Africa. *Earth and Planetary Science Letters,* **51**, 132-138.

Fitton JG (1987) The Cameroon line, West Africa: a comparison between oceanic and continental alkaline volcanism. In: *Alkaline Igneous Rocks* (eds. Fitton JG, Upton BGJ), pp. 273-291. Geological Society Special Publication No 30.

Fitzpatrick BM (2002) Molecular correlates of reproductive isolation. *Evolution,* **56**.

Fitzpatrick BM, Tureli M (2006) The geography of mammalian speciation: mixed signals from phylogenies and range maps. *Evolution,* **60**, 601-615.

Fitzpatrick JW (1988) Why so many passerine birds? A response to Raikow. *Systematic Zoology,* **37**, 72-77.

Fjeldså J (1994) Geographical patterns of neoendemic relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodiversity and Conservation,* **3**, 107-226.

Fjeldså J, Bayes MK, Bruford MW, Roy MS (2005) Biogeography and diversification of Africa forest faunas: implications for conservation. In: *Tropical Rainforests: Past, Present, and Future* (eds. Bermingham E, Dick C, Moritz C), pp. 127-147. University of Chicago Press, Chicago.

Fjeldså J, Lovett J (1997) Geographic patterns of old and young species in African forest biota: The significance of specific montane areas as evolutionary centres. *Biodiversity and Conservation,* **6**, 325-346.

Fjeldså J, Zuccon D, Irestedt M, Johansson US, Ericson PGP (2003) *Sapayoa aenigma:* a New World representative of 'Old World suboscines'. *Proceedings of the Royal Society of London B (Suppl.),* **270**, S238-S241.

Fleischer RC, McIntosh CE, Tarr CL (1998) Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K-Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. *Molecular Ecology,* **7**, 533-545.

Fleischer RC, Rothstein SI (1988) Known secondary contact and rapid gene flow among subspecies and dialects in the brown-headed cowbird. *Evolution,* **42**, 1146-1158.

Forshaw JM (1983) *Kingfishers and Related Birds. Vol 1: Alcedinidae Ceryle to Cittura.* Lansdowne Editions, Sydney.

Foster SA, Scott RJ, Cresko WA (1998) Nested biological variation and speciation. *Philosophical Transactions of the Royal Society of London B,* **353**, 207-218.

Franck P, Garnery L, Solignac M, Cornuet J-M (1998) The origin of west European subspecies of honeybees (*Apis mellifera*): new insights from microsatellite and mitochondrial data. *Evolution,* **52**, 1119-1134.

Fredoux A (1994) Pollen analysis of a deep-sea core in the Gulf of Guinea: vegetation and climate changes during the late 225,000 years. *Paleogeography, Paleoclimatology, Paleoecology,* **109**, 317-330.

Frentiu FD, Lange CL, Burke T, Owens IPF (2003) Isolation of microsatellite loci in the Capricorn silvereye, *Zosterops lateralis chlorocephalus* (Aves: Zosteropidae). *Molecular Ecology Notes,* **3**, 462-464.

Fry CH (2000) Family Zosteropidae: the white-eyes. In: *The Birds of Africa* (eds. Fry CH, Keith S, Urban EK), pp. 305-326. Academic Press, London.

Fry CH, Fry K, Harris A (1992) *Kingfishers, Bee-eaters and Rollers.* Christopher Helm, London.

Fry CH, Keith S, eds. (2004) *The Birds of Africa. Vol VII.* Christopher Helm, London.

Fry CH, Naurois Rd (1985) *Corythornis* systematics and character release in the Gulf of Guinea. *Proceedings of the 5th Pan-African Ornithological Congress*, 47-61.

Fuchs J, Ohlson JI, Ericson PGP, Pasquet E (in press) Molecular phylogeny and biogeographic history of the piculets (Piciformes: Pucumninae). *Journal of Avian Biology.*

Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from Animal mitochondrial DNA. *Annual Review of Ecology and Systematics,* **34**, 397-423.

Gagiotti OE, Lange O, Rasmann K, Gliddon C (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology,* **8**, 1513-1520.

García-Moreno J (2004) Is there a universal mtDNA clock for birds? *Journal of Avian Biology,* **35**, 465-468.

Gascoigne A (1994) The biogeography of land snails in the islands of the Gulf of Guinea. *Biodiversity and Conservation,* **3**, 794-807.

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

Gibbs HL, Tabak LM, Hobson K (1999) Characterization of microsatellite DNA loci for a neotropical migrant songbird, the Swainson's thrush (*Catharus ustulatus*). *Molecular Ecology,* **8**, 1551-1552.

Gitzendanner MA, Soltis PS (2000) Patterns of genetic variation in rare and widespread plant congeners. *American Journal of Botany,* **87**, 783-792.

Goldstein DB, Linares AR, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics,* **139**, 463-471.

Goldstein DB, Pollock DD (1997) Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of Heredity,* **88**, 335-342.

Goldstein DB, Schlötterer C (1999) *Microsatellites: Evolution and Applications.* Oxford University Press, Oxford.

Goodman SJ, Barton NH, Swanson G, Abernethy K, Pemberton JM (1999) Introgression through rare hybridization: a genetic study of a hybrid zone between red and sika deer (Genus Cervus) in Argyll, Scotland. *Genetics,* **152**, 355-371.

Gosler AG (1986) Pattern and process in the bill morphology of the Great Tit *Parus major*. *Ibis,* **129**, 451-476.

Goudet J (2002) FSTAT: a computer program to estimate and test gene diversities and fixation indices. http://www2.unil.ch/popgen/softwares/fstat.htm.

Goudet J, Raymond M, de Meeüs T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics,* **144**, 1933-1940.

Gower DJ, Kupfer A, Oomen OV, *et al.* (2002) A molecular phylogeny of ichthyophiid caecilieans (Amphibia: Gymnophiona: Ichthyophiidae): Out of India or out of South East Asia? *Proceedings of the Royal Society of London B,* **269**, 1563-1569.

Graham CH, Smith TB, Languy M (2005) Current and historical factors influencing patterns of species richness and turnover of birds in the Gulf of Guinea highlands. *Journal of Biogeography,* **32**, 1371-1384.

Grant BR, Grant PR (1979) Darwin's Finches: population variation and sympatric speciation. *Proceedings of the National Academy of Sciences USA,* **76**, 2359-2363.

Grant BR, Grant PR (1996) Cultural inheritance of song and its role in the evolution of Darwin's finches. *Evolution,* **50**, 2471-2487.

Grant BR, Grant PR (1998) Hybridization and speciation in Darwin's finches: the role of sexual imprinting on a culturally transmitted trait. In: *Endless Forms: Species and Speciation* (eds. Howard DJ, Berlocher SH), pp. 404-422. Oxford University Press, New York.

Grant BR, Grant PR (2002a) Lack of premating isolation at the base of a phylogenetic tree. *American Naturalist,* **160**, 1-19.

Grant BR, Grant PR (2002b) Simulating secondary contact in allopatric speciation: an empirical test of premating isolation. *Biological Journal of the Linnean Society,* **76**, 545-556.

Grant PR (1965a) Adaptive significance of some size trends is island birds. *Evolution,* **19**, 355-367.

Grant PR (1965b) Ecological compatibility of bird species on islands. *American Naturalist,* **100**, 451-462.

Grant PR (1965c) Plumage and the evolution of birds on islands. *Systematic Zoology,* **14**, 47-52.

Grant PR (1968) Bill size, body size, and the ecological adaptations of bird species to competitive situations on islands. *Systematic Zoology,* **17**, 319-333.

Grant PR (1986) *Ecology and Evolution of Darwin's Finches.* Princeton University Press, Princeton.

Grant PR (1998) Patterns on islands and microevolution. In: *Evolution on Islands* (ed. Grant PR), pp. 1-17. Oxford University Press, Oxford.

Grant PR (2001) Reconstructing the evolution of birds on islands: 100 years of research. *Oikos,* **92**, 385-403.

Grant PR, Grant BR (1992) Hybridization of bird species. *Science,* **256**, 193-197.
Grant PR, Grant BR (1995) Predicting microevolutionary responses to directional selection on heritable variation. *Evolution,* **49**, 241-251.
Grant PR, Grant BR (1997a) Genetics and the origin of bird species. *Proceedings of the National Academy of Sciences USA,* **94**, 7768-7775.
Grant PR, Grant BR (1997b) Hybridization, sexual impriting and mate choice. *American Naturalist,* **149**, 1-28.
Grant PR, Grant BR (1997c) Mating patterns of Darwin's Finch hybrids determined by song and morphology. *Biological Journal of the Linnean Society,* **60**, 317-343.
Grant PR, Grant BR (2002c) Unpredictable evolution in a 30-year study of Darwin's finches. *Science,* **296**, 707-711.
Grant PR, Grant BR, Petren K (2000) The allopatric phase of speciation: the sharp-beaked ground finch (*Geospiza difficilis*) on the Galápagos islands. *Biological Journal of the Linnean Society,* **69**, 287-317.
Grant PR, Grant BR, Petren K (2001) A population founded by a single pair of individuals: establishment, expansion, and evolution. *Genetica,* **112-113**, 359-382.
Gray GR (1862) Descriptions of a few West-African birds. *Annual Magazine Natural History,* **(3) 10**, 443-445.
Griffith SC (2000) High fidelity on islands: a comparative study of extrapair paternity in passerine birds. *Behavioral Ecology,* **11**, 265-273.
Griffith SC, Parker TH, Olson VA (2006) Melanin- versus carotenoid-based sexual signals: is the difference really so black and red? *Animal Behaviour,* **71**, 749-763.
Griffiths R, Double MC, Orr K, Dawson RJG (1998) A DNA test to sex most birds. *Molecular Ecology,* **7**, 1071-1075.
Guedes ME, Peiriço NM (1997) *Carbonários, Operação Salamandra: Chioglossa lusitanica Bocage 1865.* Edições Contraponto, Palmela.
Guindon S, Gascuel O (2003) A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology,* **52**, 696-704.
Gíslason D, Ferguson M, Skúlason S, Snorrason SS (1999) Rapid and coupled phenotypic and genetic divergence in Icelandic Artic char (*Salvelinus alpinus*). *Canadian Journal of Fisheries and Aquatic Sciences,* **56**, 2229-2234.
Hafner MS, Remsen JV, Jr, Laynon SM (1984) Birds versus mammal morphological differentiation. *Evolution,* **38**, 1154-1156.
Haldane JBS (1922) Sex-ratio and unisexual sterility in hybrid animals. *Journal of Genetics,* **12**, 101-109.
Hall BP, Moreau RE, eds. (1970) *An Atlas of Speciation in African Passerine Birds.* Trustees of the British Museum of Natural History, London.
Hamao S, Ueda K (2000) Simplified song in an island population of the Bush Warbler *Cettia diphone. Journal of Ethology,* **18**, 53-57.
Hansson B, Åkesson M, Slate J, Pemberton JM (2005) Linkage mapping reveals sex-dimorphic map distances in a passerine bird. *Proceedings of the Royal Society of London B,* **272**, 2289-2298.
Hardy O, Charbonnel N, Freeville H, Heuertz M (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics,* **163**, 1467-1482.
Harr B, Weiss S, David JR, Brem G, Schlotterer C (1998) A microsatellite-based multilocus phylogeny of the *Drosophila melanogaster* species complex. *Current Biology,* **8**, 1183.
Harrison GLA, McLenachan PA, Phillips MJ, *et al.* (2004) Four new avian mitochondrial genomes help get to basic evolutionary questions in late Cretaceous. *Molecular Biology and Evolution,* **21**, 974-983.
Hartlaub G (1852) Zweiter Beitrag zur Ornithologie Westafrica's. *Abhandlungen und Gebiet der Naturwissenschaften in Hamburg,* **2**, 57-68.

Hasegawa M, Kishino H, Yano T-A (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution,* **22**, 160-174.

Hayward A, Stone GN (2006) Comparative phylogeography across two trophic levels: the oak gall wasp *Andricus kollari* and its chalcid parasitoid *Megastigmus stigmatizans*. *Molecular Ecology,* **15**, 479-489.

Hazevoet CJ (1996) Conservation and species limits: taxonomic neglect promotes the extinction of endemic birds, as exemplified by taxa from eastern Atlantic islands. *Bird Conservation International,* **6**, 181-196.

Hedrick PW (1999) Highly variable loci and their interpretation in evolution and conservation. *Evolution,* **53**, 313-318.

Heim de Balsac H, Hutterer R (1982) Les Soricidae (Mammifères Insectivores) des îles du Golfe de Guinée: faits nouveaux et problèmes biogeographiques. *Bonner Zoologische Beiträge,* **33**, 133-150.

Helbig AJ, Martens J, Seibold I, *et al.* (1996) Phylogeny and species limits in the Paleartic chiffchaff *Phylloscopus collybita* complex: mitochondrial genetic differentiation and bioacoustic evidence. *Ibis,* **138**, 650-666.

Hendry AP, Wenburg JK, Bentzen P, Volk EC, Quinn TP (2000) Rapid evolution of reproductiver isolation in the wild: evidence from introduced salmon. *Science,* **290**, 516-518.

Heras P, Infante M, Ondó CO, Gascoigne A (2002) Vegetacíon de la isla de Annobón (República de Guinea Ecuatorial). *Estudios del Museo de Ciencias Naturales de Alava,* **17**, 115-123.

Heslewood MM, Elphinstone MS, Tidemann SC, Baverstock PR (1998) Myoglobin intron variation in the Gouldian Finch *Erythrura gouldiae* assessed by temperature gradient gel electrophoresis. *Electrophoresis,* **19**, 142-151.

Hey J (2001) *Genes, Categories, and Species: the Evolutionary and Cognitive Causes of the Species Problem.* Oxford University Press, New York.

Hey J (2006) On the failure of modern species concepts. *Trends in Ecology and Evolution,* **21**, 447-450.

Hillis DM, Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology,* **42**, 182-192.

Hoekstra HE, Hirschmann RJ, Bundley RJ, Insel P, Crossland JP (2006) A single amino acid mutation contributes to adaptive color pattern in beach mice. *Science,* **313**, 101-104.

Hoelzer GA, Melnick DJ (1994) Patterns of speciation and limits to phylogenetic resolution. *Trends in Ecology and Evolution,* **9**, 104-107.

Holder M, Lewis PO (2003) Phylogeny estimation: traditional and bayesian approaches. *Nature Reviews Genetics,* **4**, 275-284.

Hoskin CJ, Higgie M, McDonald KR, Moritz C (2005) Reinforcement drives rapid allopatric speciation. *Nature,* **437**, 1353-1356.

Howard DJ, Berlocher SH, eds. (1998) *Endless Forms: Species and Speciation.* Oxford University Press, New York.

Huber SK, Podos J (2006) Beak morphology and song features covary in a population of Darwin's finches (*Geospiza fortis*). *Biological Journal of the Linnean Society,* **88**, 489-498.

Huelsenbeck JP, Bollback JP (2001) Empirical and hierarchical Bayesian estimation of ancestral states. *Systematic Biology,* **50**, 351-366.

Huelsenbeck JP, Larget B, Alfaro ME (2004) Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Molecular Biology and Evolution,* **21**, 1123-1133.

Huelsenbeck JP, Larget B, Miller RE, Ronquist F (2002) Potential applications and pitfalls of Bayesian inference of phylogeny. *Systematic Biology,* **51**, 673-688.

Huelsenbeck JP, Rannala B (2004) Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. *Systematic Biology,* **53**, 904-913.

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogeny. *Bioinformatics,* **17**, 754-755.

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science,* **294**, 2310-2314.

Inskipp T, Gillet HJ, eds. (2005) *Checklist of CITES Species and Annotated CITES Appendices and Reservations.* CITES Secretariat and UNEP-WCMC, Geneva and Cambridge.

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**, 23.

Irwin DE, Bensch S, Price TD (2001) Speciation in a ring. *Nature,* **409**, 333-337.

Irwin DE, Price TD (1999) Sexual imprinting, learning and speciation. *Heredity,* **82**, 347-354.

Isawa Y, Pomiankowski A (1995) Continual changes in mate preferences. *Nature,* **377**, 420-422.

IUCN (2001) *IUCN Red List Categories and Criteria: Version 3.1.* IUCN Species Survival Commission, Gland and Cambridge.

IUCN (2006) *2006 IUCN Red List of Threatened Species.*http://www.iucnredlist.org, downloaded June 2006.

Jahns S (1996) Vegetation history and climate changes in West Equatorial Africa during the late Pleistocene and Holocene, based on a marine pollen diagram from the Congo fan. *Vegetation History and Archaeobotany,* **5**, 207-213.

Jensen FP, Stuart SN (1986) The origin and evolution of the Cameroon montane forest avifauna. In: *Conservation of Cameroon Montane Forests* (ed. Stuart SN), pp. 28-37. International Council for Bird Preservation, Cambridge.

Jesus J, Brehm A, Harris DJ (2005a) Phylogenetic relationships of *Hemidactylus* geckos from the Gulf of Guinea islands: patterns of natural colonizations and anthropogenic introductions estimated from mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics and Evolution,* **34**, 480-485.

Jesus J, Brehm A, Harris DJ (2005b) Relationships of scincid lizards (*Mabuya* spp.) from the islands of the Gulf of Guinea based on mtDNA sequence data. *Amphibia-Reptilia,* **26**.

Jesus J, Harris DJ, Brehm A (2005c) Phylogeography of *Mabuya maculilabris* (Reptilia) from São Tomé Island (Gulf of Guinea) inferred from mtDNA sequences. *Molecular Phylogenetics and Evolution,* **37**, 503-510.

Jiggins CD (2006) Speciation: Reinforced butterfly speciation. *Heredity,* **96**, 107-108.

Jiggins CD, Bridle JM (2004) Speciation in the apple maggot fly: A blend of vintages? *Trends in Ecology and Evolution,* **19**, 111-114.

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

Johnson KP, Seger J (2001) Elevated rates of nonsynonymous substitution in island birds. *Molecular Biology and Evolution,* **18**, 874-881.

Jones PJ (1994) Biodiversity in the Gulf of Guinea: an overview. *Biodiversity and Conservation,* **3**, 772-784.

Jones PJ, Burlison J, Tye A (1992) The status of endemic birds and their habitats in São Tomé and Príncipe. *Proceedings of the 7th Pan-African Ornithological Congress,* 453-459.

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

Jordan DS (1905) The origin of species through isolation. *Science,* **22**, 545-562.

Juste BJ, Fa JE (1994) Biodiversity conservation in the Gulf of Guinea islands: taking stock and preparing action. *Biodiversity and Conservation,* **3**, 759-771.

Juste BJ, Ibañez C (1994) Bats of the Gulf of Guinea: faunal composition and origins. *Biodiversity and Conservation,* **3**, 837-850.

Kaneshiro KY, Gillepsie RG, Carson HL (1995) Chromosomes and male genitalia of Hawaiian *Drosophila*: tools for interpreting phylogeny and geography. In: *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago* (eds. Wagner WL, Funk VA). Smithsonian Institution Press, Washington.

Kankare M, Shaw MR (2004) Molecular phylogeny of Cotesia Cameron, 1891 (Insecta: Hymenoptera: Braconidae: Microgastrinae) parasitoids associated with Melitaeini butterflies (Insecta: Lepidoptera: Nymphalidae: Melitaeini). *Molecular Phylogenetics and Evolution,* **32**, 207.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association,* **90**, 773-795.

Kaup (1848). In: *Verhandlungen Naturhistorischer Verein fuer das Grossherzogthum Hessen und Umgebung - Heft 2*, p. 72, Darmstadt.

Keast A (1970) Adaptive evolution and shifts in niche occupation in island birds. *Biotropica,* **2**, 61-75.

Keith S, Urban EK (1992) A summary of present knowledge of the status of thrushes in the *Turdus olivaceus* species complex. *Proceedings of the 7th Pan-African Ornithological Congress*, pp. 249-260.

Kim JH (1996) General inconsistency conditions for maximum parsimony: effects of branch lengths and increasing numbers of taxa. *Systematic Biology,* **45**, 363-374.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution,* **16**, 111-120.

Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. *Genetics,* **49**, 725-738.

Kimura M, Otha T (1978) Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences USA,* **75**, 2868-2872.

Kirkpatrick M, Ravigné V (2002) Speciation by natural and sexual selection: models and experiments. *American Naturalist,* **159**, S22-S35.

Klicka J, Voelker G, Spellman GM (2005) A molecular phylogenetic analysis of the "true trushes" (Aves: Turdinae). *Molecular Phylogenetics and Evolution,* **34**, 486-500.

Knowles LL, Richards CL (2005) Importance of genetic drift during Pleistocene divergence as revealed by analyses of genomic variation. *Molecular Ecology,* **14**, 4023-4032.

Komdeur J, Piersma T, Kraaijeveld K, Kraaijeveld-Smit F, Richardson DS (2004) Why Seychelles Warblers fail to recolonize nearby islands: unwilling or unable to fly there? *Ibis,* **146**, 298-302.

Krebs JR, Kroodsma DE (1980) Repertoires and geographical variation in bird song. In: *Advances in the Study of Behavior vol. 11* (eds. Rosenblatt JS, Hinde RA, Beer C, Busnel M-C), pp. 143-177. Academic Press, New York.

Krüger K, Gaillard C, Stranzinger G, Rieder S (2005) Phylogenetic analysis and species allocation of individual equids using microsatellite data. *Journal of Animal Breeding and Genetics,* **122**, 78-86.

Lachlan RF, Servedio MR (2004) Song learning accelerates allopatric speciation. *Evolution,* **58**, 2049-2063.

Lack D (1944) Ecological aspects of species formation in passerine birds. *Ibis,* **86**, 260-286.

Lack D (1947) *Darwin's Finches.* Cambridge University Press, Cambridge.

Lack D (1971) *Ecological Isolation in Birds.* Blackwell Scientific, Oxford.

Lack D (1976) *Island Biology Illustrated by the Land Birds of Jamaica.* Blackwell Scientific, Oxford.

Lambert DM, Ritchie PA, Millar CD, *et al.* (2002) Rates of evolution in ancient DNA from Adélie penguins. *Science,* **295**, 2270-2273.

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

Lande R (1980) Genetic variation and phenotypic evolution during allopatric speciation. *American Naturalist,* **116**, 463-479.

Lande R (1981) Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences USA,* **78**, 3721-3725.

Lande R (1982) Rapid origin of sexual isolation and character divergence in a cline. *Evolution,* **36**, 213-223.

Langerhans RB, Layman CA, Langerhans AK, Dewitt TJ (2003) Habitat-associated morphological divergence in two Neotropical fish species. *Biological Journal of the Linnean Society,* **80**, 689-698.

Lank DB, Dale J (2001) Visual signals for individual indentification: the silent "song" of Ruffs. *Auk,* **118**, 759-765.

Lee D-C, 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.

Lemmon AR, Moriarty EC (2004) The importance of proper model assumption in Bayesian phylogenetics. *Systematic Biology,* **53**, 47-67.

Lesica P, Allendorf FW (1995) When are peripheral populations valuable for conservation. *Conservation Biology,* **9**, 753-760.

Lessells CM, Boag PT (1987) Unrepeatable repeatabilities: a common mistake. *Auk,* **104**, 116-121.

Liou LW, Price T (1994) Speciation by reinforcement of prezygotic isolation. *Evolution,* **48**, 1451-1459.

Lougheed SC, Handford P (1992) Vocal dialects and the structure of geographic variation in morphological and allozymic characters in the rufous-collared sparrow, *Zonotrichia capensis*. *Evolution,* **46**, 1443-1456.

Lovett JC (1993) Climatic history and forest distribution in eastern Africa. In: *Biogeography and Ecology of the Rain Forests of Eastern Africa* (eds. Lovett JC, Wasser SW), pp. 23-29, 33-55. Cambridge University Press, Cambridge.

Lovette IJ (2004) Mitochondrial dating and mixed support for the "2% rule" in birds. *Auk,* **121**, 1-6.

Lovette IJ, Bermingham E (1999) Explosive radiation in the New World *Dendroica* warblers. *Proceedings of the Royal Society of London B,* **266**, 1629-1636.

Lowe A, Harris S, Asthon P (2004) *Ecological Genetics: Design, Analysis, and Application.* Blackwell Publishing, Oxford.

Lynch A, Baker AJ (1994) A population memetics approach to cultural evolution in Chaffinch song: differentiation among populations. *Evolution,* **48**, 351-359.

Lézine A-M, Tastet J-P, Leroux M (1994) Evidence of atmospheric paleocirculation over the Gulf of Guinea since the last Glacial Maximum. *Quaternary Research,* **41**, 390-395.

MacArthur RH, Diamond JM, Karr JR (1972) Density compensation in island faunas. *Ecology,* **53**, 330-342.

MacDougall-Shackleton EA, MacDougall-Shackleton SA (2001) Cultural and genetic evolution in mountain white-crowned sparrows: song dialects are associated with population structure. *Evolution,* **55**, 2568-2575.

Maddison WP, Maddison DR (2005) Mesquite: a modular system for evolutionary analysis. http://mesquiteproject.org.

Magurran AE (1988) *Ecological Diversity and its Measurement.* Princeton University Press, Princeton.

Magurran AE, May TM, eds. (1999) *Evolution of Biological Diversity.* Oxford University Press, Oxford.

Maley J (1996) The African rain forest - main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh,* **104B**, 31-73.

Mallet J (2001) The speciation revolution. *Journal of Evolutionary Biology,* **14**, 887-888.

Markland HM, Lovette IJ (2005) Phylogenetic affinities and inter-island differentiation in the Vitelline Warbler *Dendroica vitellina*, a West Indian endemic. *Ibis,* **147**, 764-771.

Marks BD, Willard DE (2005) Phylogenetic relationships of the Madagascar Pigmy Kingfisher (Ispidina madagascariensis). *Auk,* **122**, 1271-1280.

Marshall HD, Baker AJ (1999) Colonization history of Atlantic island common chaffinches revealed by mitochondrial DNA. *Molecular Phylogenetics and Evolution,* **11**, 201-212.

Martens J (1996) Vocalizations and speciation of Paleartic birds. In: *Ecology and Evolution of Acoustic Communication in Birds* (eds. Kroodsma DEM, Miller EH), pp. 221-240. Cornell University Press, Ithaca.

Martens J, Steil B (1997) Territorial songs and species differentiation in the lesser whitethroat subspecies *Sylvia* (*curruca*). *Journal of Ornithology,* **138**, 1-23.

Martin TE (1996) Life history evolution in tropical and south temperate birds: What do we really know? *Journal of Avian Biology,* **27**, 263-272.

Matyjasiak P (2005) Birds associate species-specific acoustic and visual cues: recognition of heterospecific rivals by male blackcaps. *Behavioral Ecology,* **16**, 467-471.

Mayden RL (1997) A hierarchy of species concepts: the denouement in the saga of the species problem. In: *Species: the Units of Biodiversity* (eds. Claridge MF, Dawah HA, Wilson MR), pp. 381-424. Chapman & Hall, London.

Mayr E (1940) Speciation phenomena in birds. *American Naturalist,* **74**, 249-278.

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

Mayr E (1954) Change of genetic environment and evolution. In: *Evolution as a Process* (eds. Huxley HS, Hardy AC, Ford EB), pp. 157-180. Allen and Unwin, London.

Mayr E (1963) *Animal Species and Evolution.* Belknap Press, Cambridge.

Mayr E (1965) Avifauna: turnover on islands. *Science,* **150**, 1587-1588.

Mayr E (1992) A local flora and the Biological Species Concept. *American Journal of Botany,* **79**, 222-238.

Mayr E (1999) Introduction, 1999. In: *Systematics and the Origin of Species From the Viewpoint of a Zoologist* (ed. Mayr E), pp. xiii-xxxv. Harvard University Press, Cambridge.

Mayr E (2000a) The Biological Species Concept. In: *Species Concepts and Phylogenetic Theory* (eds. Wheeler QD, Meier R), pp. 17-29. Columbia University Press, New York.

Mayr E (2000b) A critique from the Biological Species Concept perspective: What is a species, and what is not? In: *Species Concepts and Phylogenetic Theory* (eds. Wheeler QD, Meier R), pp. 93-100. Columbia University Press, New York.

Mayr E (2000c) A defense of the Biological Species Concept. In: *Species Concepts and Phylogenetic Theory* (eds. Wheeler QD, Meier R), pp. 161-166. Columbia University Press, New York.

Mayr E, Diamond JM (2001) *The Birds of Northern Melanesia: Speciation, Ecology, and Biogeography.* Oxford University Press, Oxford.

Mayr E, O'Hara RJ (1986) The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. *Evolution,* **40**, 55-67.

McAllister DE, Schueler FW, Roberts CM, Hawkins JP (1994) Mapping and GIS analysis of the global distribution of coral reef fishes on an equal-area grid. In: *Mapping the Diversity of Nature* (ed. Miller RI), pp. 155-176. Chapman & Hall, London.

McCracken KG, Sorenson MD (2005) Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear genome trees in the stiff-tailed ducks (*Nomonyx-Oxyura*)? *Systematic Biology,* **54**, 35-55.

McGraw KJ (2003) Melanins, metals and mate quality. *Oikos,* **102**, 402-406.

McGraw KJ, Dale J, Mackillop EA (2003) Social environment during molt predicts the expression of melanin-based plumage pigmentation in male house sparrows (*Passer domesticus*). *Behavioural Ecology and Sociobiology,* **53**, 116-122.

McGregor PK (2000) Playback experiments: design and analysis. *Acta Ethologica,* **3**, 3-8.

McNaught MK, Owens IPF (2002) Interspecific variation in plumage colour among birds: species recognition or light environment? *Journal of Evolutionary Biology,* **15**, 505-514.

McRae SB, Amos W (1999) Characterization of hypervariable microsatellites in the cooperatively breeding white-browed sparrow weaver *Plocepasser mahali*. *Molecular Ecology,* **8**, 903-904.

Measey GJ, Vences M, Drewes RC, *et al.* (2007) Freshwater paths into the ocean: molecular phylogeny of the frog *Ptychadena newtoni* gives insights into amphibian colonisation of oceanic islands. *Journal of Biogeography*, **in press**, doi: 10.1111/j.1365-2699.2006.01589.x.

Mees GF (1957) A systematic review of the Indo-Australian Zosteropidae, Part I. *Zoologische Verhandelingen,* **35**, 1-204.

Mees GF (1961) A systematic review of the Indo-Australian Zosteropidae, Part II. *Zoologische Verhandelingen,* **50**, 1-168.

Mees GF (1969) A systematic review of the Indo-Australian Zosteropidae, Part III. *Zoologische Verhandelingen,* **102**, 1-390.

Melo M (1998) Differentiation between Príncipe Island and mainland populations of the African Grey Parrot *Psittacus erithacus* - First field expedition report: 28 October - 3 December 1998. Percy FitzPatrick Institute, University of Cape Town, Cape Town.

Melo M, Hansson B (2006) Identification of 15 polymorphic microsatellite loci in the Príncipe seedeater (*Serinus rufobrunneus*) and assessment of their utility in nine other *Serinus* species (Fringillidae, Aves). *Molecular Ecology Notes,* **in press**, doi: 10.1111/j.1471-8286.2006.01510.x.

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*, **in press**.

Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference to microsatellite loci. *Genetics,* **142**, 1061-1064.

Minch E, Ruiz-Linares A, Goldstein DB, Feldman M, Cavalli-Sforza LL (1995) MICROSAT: a computer program for calcultating various statistics on microsatellite allele data. University of Standford, Standford.

Mindell DP, Sorenson MD, Dimcheff DE (1998) An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. *Molecular Biology and Evolution,* **15**, 1568-1571.

Mindell DP, Thacker CE (1996) Rates of molecular evolution: phylogenetic issues and applications. *Annual Review of Ecology and Systematics,* **27**, 279-303.

Mooers AØ, Rundle HD, Whitlock MC (1999) The effect of selection and bottlenecks on male mating success in peripheral isolates. *American Naturalist,* **153**, 437-444.

Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution,* **49**, 718-726.

Moore WS, DeFilippis VR (1997) The window of taxonomic resolution for phylogenies based on mitochondrial cytochrome b. In: *Avian Molecular Evolution and Systematics* (ed. Mindell DP), pp. 84-119. Academic Press, San Diego.

Moreau RE (1957) Variation in the western Zosteropidae (Aves). *Bulletin of the British Museum (Natural History), Zoology,* **4**, 309-433.
Moreau RE (1960) Conspectus and classification of the Ploceine Weaver-Birds. *Ibis,* **102**, 298-321.
Moreau RE (1962) Ploceinae. In: *Check-list of the Birds of the World, Vol. 15.* (eds. Mayr E, Paynter RA). Museum of Comparative Zoology, Cambridge.
Moreau RE (1966) *The Bird Faunas of Africa and its Islands.* Academic Press, London.
Morell V (1999) Ecology returns to speciation studies. *Science,* **284**, 2106-2108.
Moritz C (1994) Defining 'evolutionary significant units' for conservation. *Trends in Ecology and Evolution,* **9**, 373-375.
Moritz C, Patton JL, Schneider CJ, Smith TB (2000) Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics,* **31**, 533-563.
Morrone JJ, Crisci JV (1995) Historical biogeography: introduction to methods. *Annual Review of Ecology and Systematics,* **26**, 373-401.
Morton ES (1975) Ecological sources of selection on avian sounds. *American Naturalist,* **109**, 17-34.
Moulton MP (1985) Morphological similarity and coexistence of congeners: an experimental test with introduced Hawaiian birds. *Oikos,* **44**, 301-305.
Moyle RG (2006) A molecular phylogeny of kingfishers (Alcedinidae) with insights into early biogeographic history. *Auk,* **123**, 487-499.
Moyle RG, Fuchs J, Pasquet E, Marks B (in press) Feeding behavior, toe count, and the phylogenetic relationships among alcedinine kingfishers. *Journal of Avian Biology.*
Mundy N (2005) A window on the genetics of evolution: MC1R and plumage colouration in birds. *Proceedings of the Royal Society of London B,* **272**, 1633-1640.
Mustapha M, McCall PJ, Cheke RA, Post RJ (2006) The blackflies (Diptera: Simuliidae) of Bioko (Republic of Equatorial Guinea) and the Gulf of Guinea with a description of the larvae of the 'Pomeroy' form of *Simulium cervicornutum. Systematic Entomology,* **Unpublished for the purposes of zoological nomenclature**, doi: 10.111/j.1365-3113.2006.00330.x.
Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature,* **403**, 853-858.
Naurois Rd (1975a) Le "Scops" de l'Ile de São Tomé *Otus hartlaubi* (Giebel). *Bonner Zoologische Beiträge,* **26**, 319-355.
Naurois Rd (1975b) Les Carduelinae des îles de São Tomé et Principe (Golfe de Guinée). *Ardeola,* **21 especial**, 903-931.
Naurois Rd (1984) Les *Turdus* des îles de São Tomé et Principe: *T. o. olivaceofuscus* (Hartlaub) et *T. olivaceofuscus xanthorhynchus* Salvadori (*Aves Turdinae*). *Revue de Zoologie Africaine,* **98**, 403-423.
Naurois Rd (1988) *Neospiza concolor* (Bocage, 1888): Endémique de l'île de São Tomé (Golfe de Guinée). *Bollettino del Museo Regionale di Scienze Naturali (Torino),* **6**, 321-339.
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.
Nei M (1987) *Molecular Evolutionary Genetics.* Columbia University Press, New York.
Nei M, Maruyama T, Wu C-I (1983a) Models of evolution of reproductive isolation. *Genetics,* **103**, 557-579.
Nei M, Tajima F, Tateno Y (1983b) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution,* **19**, 153-170.
Nelson DA (2000) A preference for own-subspecies' song guides vocal learning in a song bird. *Proceedings of the National Academy of Sciences USA,* **97**, 13348-13353.
Nelson DH, Soha JA (2004) Perception of geographic variation in song by male Puget Sound white-crowned sparrows. *Animal Behaviour,* **68**, 395-405.

Neumann K, Wetton JH (1996) Highly polymorphic microsatellites in the house sparrow *Passer domesticus*. *Molecular Ecology*, **5**, 307-309.

Newton I (2003) *The speciation and biogeography of birds*. Academic Press, London.

Nicholls R (2001) Gene trees and species trees are not the same. *Trends in Ecology and Evolution*, **16**, 358-364.

Nieeke M, Rothlaender S, Roulin A (2003) Why do melanin ornaments signal individual quality? Insights from metal analysis of Barn Owl feathers. *Oecologia*, **137**, 153-158.

Nosil P, Crespi BJ, Sandoval CP (2003) Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement. *Proceedings of the Royal Society of London B*, **270**, 1911-1918.

Nottebohm F, Nottebohm ME (1978) Relationship between song repertoire and age in the canary *Serinus canaria*. *Zeitschrift fur Tierpsychologie*, **46**, 298-305.

Nottebohm F, Nottebohm ME, Crane L (1986) Developmental and seasonal changes in canary song and their relation to changes in the anatomy of song control nuclei. *Behavioral and Neural Biology*, **46**, 445-471.

Nunn GB, Stanley SE (1998) Body size effects and rates of cytochrome *b* evolution in tube-nosed seabirds. *Molecular Biology and Evolution*, **15**, 1360-1371.

Nussbaum RA, Pfrender ME (1998) Revision of the African caecilian genus *Schistometopum* Parker (Amphibia: Gymnophiona: Caecilidae). *Miscellaneous Publications of the Museum of Zoology of the University of Michigan*, **I-IV**, 1-32.

Nylander JAA (2004) MRMODELTEST. Program distributed by the author. Evolutionary Biology Centre, Uppsala University. http://www.ebc.uu.se/systzoo/staff/nylander.html.

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

Ogden R, Thorpe RS (2002) Molecular evidence for ecological speciation in tropical habitats. *Proceedings of the National Academy of Sciences USA*, **99**, 13612-13615.

Ogonovszky M (2003) *Endémisme et Phytogéographie des Plantes de São Tomé et Príncipe*. MSc Thesis, Laboratoire de Botanique Systématique et de Phytosociologie - Université Libre de Bruxelles, Bruxelles.

Olson D, Dinerstein E, Abell R (2000) *The Global 200: A Representation Approach to Conserving the Earth's Distinctive Ecoregions*. World Wide Fund Conservation Science Program, Washington DC.

Omland KE, Lanyon SM (2000) Reconstructing plumage evolution in orioles (*Icterus*): repeated convergence and reversal in patterns. *Evolution*, **54**, 2119-2133.

Orr HA (1997) Haldane's rule. *Annual Review of Ecology and Systematics*, **28**, 195-218.

Orr MR, Smith TB (1998) Ecology and speciation. *Trends in Ecology and Evolution*, **13**, 502-506.

Pagel M, Meade A (2004) A phylogenetic mixture-model for detecting pattern-heterogeneity in gene sequence or character-state data. *Systematic Biology*, **53**, 571-581.

Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology*, **53**, 673-684.

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

Patten MA, Rotenberry JT, Zuk M (2004) Habitat selection, acoustic adaptation, and the evolution of reproductive isolation. *Evolution*, **58**, 2144-2155.

Payne RB (1981) Population strucuture and social behavior: models for testing the ecological significance of song dialects in birds. In: *Natural Selection and Social Behavior: Recent Research and New Theory* (eds. Alexander RD, Tinkle DW). Chiron Press, New York.

Payne RB (1985) Behavioral continuity and change in local song populations of village indigobirds *Vidua chalybeata*. *Zeitschrift fur Tierpsychologie,* **70**, 1-44.

Payne RB (1996) Song traditions in Indigo Buntings: origin, improvisation, dispersal and exctinction in cultural evolution. In: *Ecology and Evolution of Acoustic Communication in Birds* (eds. Kroodsma DE, Miller EH), pp. 198-220. Cornell University Press, Ithaca.

Payne RB, Payne LL, Woods JL (1998) Song learning in brood parasite indigobirds *Vidua chalybeata*: song mimicry of the host species. *Animal Behaviour,* **55**, 1537-1553.

Payne RB, Payne LL, Woods JL, Sorenson MD (2000) Imprinting and the origin of parasite-host species associations in brood parasite indigobirds *Vidua chalybeata*. *Animal Behaviour,* **59**, 69-81.

Payne RB, Westneat DF (1988) A genetic and behavioural analysis of mate choice and song neighborhoods in indigo buntings. *Evolution,* **42**, 935-947.

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

Perkins RCL (1901) An introduction to the study of the Drepanidae, a family of birds peculiar to the Hawaiian islands. *Ibis,* **1 (eighth series)**, 562-585.

Perkins RCL (1903) Vertebrata. In: *Fauna Hawaiiensis* (ed. Sharp D), pp. 365-466. Cambridge University Press, Cambridge.

Perkins RCL (1913) Introduction. In: *Fauna Hawaiiensis* (ed. Sharp D), pp. xv-ccxxviii. Cambridge University Press, Cambridge.

Peters JL (1945) *Check-list of Birds of the World.* Harvard University Press, Cambridge.

Peterson AT, Sobnerón J, Sanchez-Cordero V (1999) Conservation of ecological niches in evolutionary time. *Science,* **285**, 1265-1267.

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology,* **12**, 844-855.

Petren K (1998) Microsatellite primers from *Geospiza fortis* and cross-species amplification in Darwin's finches. *Molecular Ecology,* **7**, 1782-1784.

Petren K, Grant BR, Grant PR (1999) A phylogeny of Darwin's finches based on microsatellite DNA length variation. *Proceedings of the Royal Society of London B,* **266**, 321-329.

Petren K, Grant PR, Grant BR, Keller LF (2005) Comparative landscape genetics and the adaptive radiation of Darwin's finches: the role of peripheral isolation. *Molecular Ecology,* **14**, 2943-2957.

Phillimore AB, Owens IPF (2006) Are subspecies useful in evolutionary and conservation biology? *Proceedings of the Royal Society of London B,* **273**, 1049-1053.

Piertney SB, Marquiss M, Summers R (1998) Characterization of tetranucleotide microsatellite markers in the Scottish crossbill (*Loxia scotica*). *Molecular Ecology,* **7**, 1261-1263.

Plana V, Gascoigne A, Forrest LL, Harris D, Pennington RT (2004) Pleistocene and pre-Pleistocene *Begonia* speciation in Africa. *Molecular Phylogenetics and Evolution,* **31**, 449-461.

Podos J (2001) Correlated evolution of morphology and vocal signal structure in darwin's finches. *Nature,* **409**, 185-188.

Podos J, Huber S, Taft B (2004) Bird song: the interface of evolution and mechanism. *Annual Review of Ecology and Systematics,* **35**, 55-87.

Pomiankowski A, Isawa Y (1998) Runaway ornament diversity caused by Fisherian sexual selection. *Proceedings of the National Academy of Sciences USA,* **95**, 5106-5111.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics,* **14**, 817-818.

Praed CW, Grant CHB (1970) *African Handbook of Birds.* Longman Group, London.

Pratt HD (2005) *The Hawaiian Honeycreepers.* Oxford University Press, Oxford.

Price T (1998) Sexual selection and natural selection in bird speciation. *Philosophical Transactions of the Royal Society of London B,* **353**, 251-260.

Price TD, Birch GL (1996) Repeated evolution of sexual color dimorphism in the passerine birds. *Auk,* **113**, 842-848.

Price TD, Bouvier MM (2002) The evolution of F1 postzygotic incompatibilities in birds. *Evolution,* **56**, 2083-2089.

Prigogine A (1987) Disjunctions of montane forest birds in the Afrotropical Region. *Bonner Zoologische Beiträge,* **38**, 195-207.

Primmer CR, Ellegren H (1998) Patterns of molecular evolution in avian microsatellites. *Molecular Biology and Evolution,* **15**, 997-1008.

Primmer CR, Painter JN, Koskinen MT, Palo JU, Merilä J (2005) Factors affecting avian cross-species microsatellite amplification. *Journal of Avian Biology,* **36**, 348-360.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics,* **155**, 945-959.

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 (2000) Comparative evolution of the mitochondrial cytochrome b gene and nuclear β-fibrinogen intron 7 in woodpeckers. *Molecular Biology and Evolution,* **17**, 1101-1111.

Pyrcz T (1992) Provisional check-list of the butterflies of São Tomé and Príncipe islands. *Lambillionea,* **92**, 48-52.

Pérez del Val J (1996) *Las Aves de Bioko.* Edilesa, Léon.

Quinn GP, Keough MJ (2002) *Experimental Design and Data Analysis for Biologists.* Cambridge University Press, Cambridge.

Rambaut A, Drummond AJ (2003) TRACER: MCMC Trace Analysis Tool. University of Oxford, Oxford.

Raymond M, Rousset F (1995) GENEPOP (version 3.4): population genetics software for exact tests and ecumenism. *Journal of Heredity,* **86**, 248-249.

Reed DH, Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution,* **55**, 1095-1103.

Rice WR (1989) Analysing tables of statisitcal tests. *Evolution,* **43**, 223-225.

Rice WR, Hostert EE (1993) Laboratory experiments on speciation: what have we learned in forty years. *Evolution,* **47**, 1637-1653.

Richard M, Thorpe RS (2001) Can microsatellites be used to infer phylogenies? Evidence from population affinities of the western Canary Island lizard (*Gallotia galloti*). *Molecular Phylogenetics and Evolution,* **20**, 351.

Richardson DS, Jury FL, Dawson DA, *et al.* (2000) Fifty Seychelles warbler (*Acrocephalus sechellensis*) microsatellite loci polymorphic in Sylviidae species and their cross-species amplification in other passerine birds. *Molecular Ecology,* **9**, 2226-2231.

Richmond JQ, Reeder TW (2002) Evidence for parallel ecological speciation in scincid lizards of the *Eumeces skiltonianus* species group (Squamata: Scincidae). *Evolution,* **56**, 1498-1513.

Rieseberg LH, Widmer A, Arntz AM, Burke JM (2002) Directional selection is the primary cause of phenotypic diversification. *Proceedings of the National Academy of Sciences USA,* **99**, 12242-12245.

Ritz LR, Glowatzki-Mullis ML, MacHugh DE, Gaillard C (2000) Phylogenetic analysis of the tribe Bovini using microsatellites. *Animal Genetics,* **31**, 178-185.

Robinson-Wolrath S, Owens IPF (2003) Large body size in an island-dwelling bird: intraspecific competition and the Dominance Hypothesis. *Journal of Evolutionary Biology,* **16**, 1106-1114.

Rodríguez F, Oliver JL, Marín A, Medina JR (1990) The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology,* **142**, 485-501.

Rohling EJ, Fenton M, Jorissen FJ, *et al.* (1998) Magnitudes of sea.level lowstands of the past 500,000 years. *Nature,* **394**, 162-165.

Ronquist F (1998) Phylogenetic approaches in coevolution and biogeography. *Zoologica Scripta,* **26**, 313-322.

Ronquist F (2004) Bayesian inference of character evolution. *Trends in Ecology and Evolution,* **19**, 475-481.

Ronquist F, Huelsenbeck JP (2003) MRBAYES3: Bayesian phylogenetic inference under mixed models. *Bioinformatics,* **19**, 475-481.

Roulin A, Dijkstra C (2003) Genetic and environmental components of variation in eumelanin and phaeomelanin sex-traits in the Barn Owl. *Heredity,* **90**, 359-364.

Rousset F (1996) Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics,* **142**, 1357-1362.

Roy MS (1997) Recent diversification in African greenbuls (Pycnonotidae: *Andropadus*) supports a montane speciation model. *Proceedings of the Royal Society of London B,* **264**, 1337-1344.

Roy MS, Da Silva J, Arctander P, García-Moreno J, Fjeldså J (1997) The role of montane regions in the speciation of South American and African birds. In: *Avian Molecular Evolution and Systematics* (ed. Mindell DP), pp. 325-343. Academic Press, San Diego.

Roy MS, Torres-Mura JC, Herel F (1998) Evolution and history of hummingbirds (Aves: Trochilidae) from the Juan Fernandez Islands (Chile). *Ibis,* **140**, 265-273.

Ruegg K, Slabbekoorn H, Clegg SB, Smith TB (2006) Divergence in mating signals correlates with ecological variation in a migratory songbird, the Swainson's thrush (*Catharus ustulatus*). *Molecular Ecology*, **15**, 3147-3156.

Rundle HD (2003) Divergent environments and population bottlenecks fail to generate premating isolation in *Drosophila pseudoobscura*. *Evolution,* **57**, 2557-2565.

Rundle HD, Chenoweth SF, Doughty P, Blows MW (2005) Divergent selection and the evolution of signal traits and mating preferences. *PloS Biology,* **3**, 1988-1995.

Rundle HD, Mooers AØ, Whitlock MC (1998) Single founder-flush events and the evolution of reproductive isolation. *Evolution,* **52**, 1850-1855.

Rundle HD, Nagel L, Boughman JW, Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science,* **287**, 306-308.

Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters,* **8**, 336-352.

Ryan PG, Bloomer P (1999) The Long-billed Lark complex: a species mosaic in Southwestern Africa. *Auk,* **116**, 194-208.

Ryan PG, Hood I, Bloomer P, Komen J, Crowe TM (1998) Barlow's Lark: a new species in the Karoo Lark *Certhilauda albescens* complex of southwest Africa. *Ibis,* **140**, 605-619.

Ryan PG, Wright D, Oatley G, *et al.* (2004) Systematics of *Serinus* canaries and the status of Cape and Yellow-crowned Canaries inferred from mtDNA and morphology. *Ostrich,* **75**, 288-294.

Saetre G-P, Moum T, Bures S, *et al.* (1997) A sexually selected character displacement in flycatchers reinforces premating isolation. *Nature,* **387**, 589-592.

Salomon M (1989) Song as a possible reproductive isolating mechanism between two parapatric forms. The case of the chiffchaff *Phylloscopus c. collybita* and *P. c. brehmii* in the western Pyrenees. *Behaviour,* **111**, 270-290.

Salvadori T (1901) Due nuove specie di ucelli dell'Isola di S. Thomé e dell'Isola del Príncipe raccolte dal Sig. Leonardo Fea. *Bollettino dei Musei di Zoologia ed Anatomia Comparata della R. Universitá di Torino,* **16**, 1-2.

Salvadori T (1902) On a new kingfisher of the genus *Corythornis*. *Ibis,* **8**, 566-569.

Salvadori T (1903) Contribuzione alla ornitologia delle isole del Golfo di Guinea II: Uccelli di San Thomé. *Memorie della Reale Accademia delle Scienze di Torino (Serie 2),* **53**, 1-45.

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

Sanmartín I, Ronquist F (2004) Southern Hemisphere biogeography inferred by event-based models: plants versus animal patterns. *Systematic Biology*, **53**, 216-243.

Sargeant DE, Gullick T, Turner DA, Sinclair JCI (1992) The rediscovery of the São Tomé Grosbeak *Neospiza concolor* in south-western São Tomé. *Bird Conservation International*, **2**, 157-159.

Sato A, Tichy H, O'hUigin C, *et al.* (2001) On the origin of Darwin's finches. *Molecular Biology and Evolution*, **18**, 299-311.

Schilthuizen M (2000) Dualism and conflicts in understanding speciation. *BioEssays*, **22**, 1134-1141.

Schilthuizen M (2001) *Frogs, Flies, and Dandelions: the Making of Species.* Oxford University Press, Oxford.

Schliewen UK, Tautz D, Paabo S (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature*, **368**, 629-632.

Schluter D (2000) *The Ecology of Adaptive Radiation.* Oxford University Press, Oxford.

Schluter D (2001) Ecology and the origin of species. *Trends in Ecology and Evolution*, **16**, 372-380.

Schluter D, McPhail JD (1993) Character displacement and replicate adaptive radiation. *Trends in Ecology and Evolution*, **8**, 197-200.

Schluter D, Price TD, Grant BR (1985) Ecological character displacement in Darwin's finches. *Science*, **227**, 1056-1059.

Schluter D, Smith JNM (1986) Natural selection on beak and body size in the song sparrow. *Evolution*, **40**, 221-231.

Schneider C, Smith TB, Larison B, Moritz C (1999) A test of alternative models of diversification in tropical rainforest birds: ecological gradients vs. rainforest refugia. *Proceedings of the National Academy of Sciences USA*, **96**, 13869-13873.

Sclater WL (1924) *Systema Avium Ethiopicarum: A Systematic List of the Birds of the Ethiopian Region. Part 1.* British Ornithologists Union, Tring.

Scott SN, Clegg SB, Blomberg SP, Kikkawa J, Owens IPF (2003) Morphological shifts in island-dwelling birds: the roles of generalist foraging and niche expansion. *Evolution*, **57**, 2147-2156.

Searcy WA, Andersson M (1986) Sexual selection and the evolution of song. *Annual Review of Ecology and Systematics*, **17**, 507-533.

Seddon N (2005) Ecological adaptation and species recognition drive vocal evolution in Neotropical suboscine birds. *Evolution*, **59**, 200-215.

Senar JC, Borras A, Cabrera J, Cabrera T, Björklund M (2006) Local differentiation in the presence of gene flow in the citril finch *Serinus citrinella*. *Biology Letters*, **2**, 85-87.

Servedio MR, Noor MA (2003) The role of reinforcement in speciation: theory and data. *Annual Review of Ecology and Systematics*, **34**, 339-364.

Shapiro B, Rambaut A, Drummond AJ (2006) Choosing appropriate substitutions models for phylogenetic analysis of protein-coding sequences. *Molecular Biology and Evolution*, **23**, 7-9.

Shaw KL (2002) Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals in Hawaiian crickets. *Proceedings of the National Academy of Sciences USA*, **99**, 16122-16127.

Shelley GE (1905) *The Birds of Africa.* R. H. Porter, London.

Shields GF, Helm-Bychowski KM (1988) Mitochondrial DNA of birds. *Current Ornithology*, **5**, 273-295.

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*, **16**, 1114-1116.

Sibley CG, Monroe Jr BL (1990) *Distribution and Taxonomy of Birds of the World.* Yale University Press, New Haven.

Sinclair I, Ryan PG (2003) *Birds of Africa South of the Sahara.* Struik Publishers, Cape Town.

Slabbekoorn H, Peet M (2003) Birds sing at a higher pitch in urban noise. *Nature,* **424**, 267.

Slabbekoorn H, Smith TB (2002) Bird song, ecology and speciation. *Philosophical Transactions of the Royal Society of London B,* **357**, 493-503.

Slade RW, Moritz C, Heideman A, Hale PT (1993) Rapid assessment of single-copy nuclear DNA variation in diverse species. *Molecular Ecology,* **2**, 359-373.

Slater PJB (1986) The cultural transmission of bird song. *Trends in Ecology and Evolution,* **1**, 94-97.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics,* **139**, 457-462.

Slikas B, Jones IB, Derrickson SR, Fleischer RC (2000) Phylogenetic relationships of Micronesian white-eyes based on mitochondrial sequence data. *Auk,* **117**, 355-365.

Smith TB (1987) Bill size polymorphism and the intraspecific niche utilization in an African finch *Pyrenestes ostrinus*. *Nature,* **329**, 717-719.

Smith TB (1993) Disruptive selection and the genetic basis of bill size polymorphism in the African finch, *Pyrenestes*. *Nature,* **363**, 618-620.

Smith TB (1997) Adaptive significance of the mega-billed form in the polymorphic black-bellied seedcracker *Pyrenestes ostrinus*. *Ibis,* **139**, 382-387.

Smith TB, Calsbeek R, Wayne RK, *et al.* (2005) Testing alternative mechanisms of evolutionary divergence in an African rain forest passerine bird. *Journal of Evolutionary Biology,* **18**, 257-268.

Smith TB, Girman D (2000) Reaching new adaptive peaks: Evolution of bill size polymorphism in an African finch. In: *Adaptive Genetic Variation in the Wild* (eds. Mousseau T, Sinervo B, Endler J), pp. 139-156. Oxford University Press, Oxford.

Smith TB, Holder K, Griman D, *et al.* (2000) Comparative avian phylogeography of Cameroon and Equatorial Guinea Mountains: implications for conservation. *Molecular Ecology,* **9**, 1505-1516.

Smith TB, Skúlason S (1996) Evolutionary significance of resource polymorphisms in fishes, amphibians, and birds. *Annual Review of Ecology and Systematics,* **27**, 111-133.

Smith TB, Wayne RW, Girman DJ, Bruford MW (1997) A role for ecotones in generating rainforest biodiversity. *Science,* **276**, 1855-1857.

Snow DW (1950) The birds of São Tomé and Principe in the Gulf of Guinea. *Ibis,* **92**, 579-589.

Snow DW (1978) *An Atlas of Speciation in African Non-Passerine Birds.* Trustees of the British Museum (Natural History), London.

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.

Sorenson MD, Quinn TW (1998) Numts: a challenge for avian systematics and population biology. *Auk,* **115**, 214-221.

Sorenson MD, Sefc KM, Payne RB (2003) Speciation by host shift in brood parasitic indigobirds. *Nature,* **424**, 928-931.

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

Stattersfield AJ, Crosby MJ, Long AJ, Wedge DC (1998) *Endemic Bird Areas of the World: Priorities for Biodiversity Conservation.* BirdLife International, Cambridge.

Stebbins GL (1950) *Variation and Evolution in Plants.* Columbia University Press, New York.

Stresemann E (1948) A small contribution to the ornithology of Fernando Po. *Ibis,* **90**, 334-335.

Sttatersfield AJ, Crosby MJ, Long AJ, Wedge DC (1998) *Endemic Bird Areas of the World: Priorities for Biodiversity Conservation.* BirdLife International, Cambridge.

Sullivan J, Abdo Z, Joyce P, Swofford DL (2005) Evaluating the performance of a successive-approximations approach to parameter optimization in maximum-likelihood phylogeny estimation. *Molecular Biology and Evolution,* **22**, 1386-1392.

Sullivan J, Holsinger E, Simon C (1996) The effect of topology on estimates of among-site rate variation. *Journal of Molecular Evolution,* **42**, 308-312.

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

Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: *Molecular Systematics* (eds. Hillis DM, Moritz C, Mable BK), pp. 407-514. Sinauer, Sunderland, MA.

Swofford DL, Waddell PJ, Huelsenbeck JP, *et al.* (2001) Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Systematic Biology,* **50**, 525-539.

Takahata N, Slatkin M (1984) Mitochondrial gene flow. *Proceedings of the National Academy of Sciences USA,* **81**, 1764-1767.

Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics,* **144**, 389-399.

Tarr CL, Fleischer RC (1993) Mitochondrial-DNA variation and evolutionary relationships in the Amakihi complex. *Auk,* **110**, 825-831.

Tarr CL, Fleischer RC (1995) Evolutionary relationships of the Hawaiian Honeycreepers (Aves, Zosteropidae). In: *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago* (eds. Wagner WL, Funk VA). Smithsonian Institution Press, Washington and London.

Tegelstrom H, Gelter HP (1990) Haldane's rule and sex biased gene flow between two hybridizing flycatcher species (*Ficedula albicollis* and *F. hypoleuca*, Aves: Muscicapidae). *Evolution,* **44**, 2012-2021.

Templeton AR (1980) The theory of speciation via the founder principle. *Genetics,* **94**, 1011-1038.

Ting CT, Tsaur SC, Wu C-I (2000) The phylogeny of closely related species as revealed by the genealogy of a speciation gene, Odysseus. *Proceedings of the National Academy of Sciences USA,* **97**, 5313-5316.

Tonnis B, Grant PR, Grant BR, Petren K (2005) Habitat selection and ecological speciation in Galápagos warbler finches (*Certhidea olivacea* and *Certhidea fusca*). *Proceedings of the Royal Society of London B,* **272**, 819-826.

Turelli M, Barton NH, Coyne JA (2001) Theory and speciation. *Trends in Ecology and Evolution,* **16**, 330-343.

Usque S (1553/1965) *Consolation for the Tribulations of Israel.* Jewish Society of America, Philadelphia.

Vaneechoutte M (1997) Bird song as a possible cultural mechanism for speciation. *Journal of Memetics - Evolutionary Models of Information Transmission,* **1**, http://cfpm.org/jom-emit/1997/vol1991/vaneechoutte_m.html.

Vences M (2005) Madagascar as a model region for the study of tempo and pattern in adaptive radiations. In: *African Biodiversity: Molecules, Organisms, Ecosystems* (eds. Huber BA, Sinclair BJ, Lampe K-H). Springer, New York.

Via S (2001) Sympatric speciation in animals: the ugly duckling grows-up. *Trends in Ecology and Evolution,* **16**, 381-390.

Via S (2002) The ecological genetics of speciation. *American Naturalist,* **159**, S1-S7.

Vines TH, Schluter D (2006) Strong assortative mating between allopatric sticklebacks as a by-product of adaptation to different environments. *Proceedings of the Royal Society of London B,* **273**, 911-916.

Voelker G (1999) Dispersal, vicariance, and clocks: historical biogeography and speciation in a cosmopolitan passerine genus (*Anthus*: Motacillidae). *Evolution,* **53**, 1536-1552.

Voelker G, Spellman GM (2003) Nuclear and mitochondrial DNA evidence of polyphyly in the avian superfamily Muscicapoidea. *Molecular Phylogenetics and Evolution,* **30**, 386-394.

Warren BH, Bermingham E, Bowie RCK, Prys-Jones R, Thébaud C (2003) Molecular phylogeography reveals island colonization history and diversification of western Indian Ocean sunbirds (*Nectarinia*: Nectariniidae). *Molecular Phylogenetics and Evolution,* **29**, 67-85.

Warren BH, Bermingham E, Prys-Jones R, Thébaud C (2005) Tracking island colonization history and phenotypic shifts in Indian Ocean bulbuls (*Hypsipetes*: Pycnonotidae). *Biological Journal of the Linnean Society,* **85**, 271-287.

Warren BH, Bermingham E, Prys-Jones R, Thébaud C (2006) Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on Indian Ocean islands. *Molecular Ecology*, **15**, 3769-3786.

Weir BS, Cockerham CC (1984) Estimating F-statisitics for the analysis of population structure. *Evolution,* **38**, 1358-1370.

Weisrock DW, Kozak KH, Larson A (2005) Phylogeographic analysis of mitochondrial gene flow and introgression in the salamander, Plethodon shermani. *Molecular Ecology,* **14**, 1457-1472.

Werner TK, Sherry TW (1987) Behavioural specialization in *Pinaroloxias inornata*, the "Darwin's finch" of Cocos Island, Costa Rica. *Proceedings of the National Academy of Sciences USA,* **84**, 5506-5510.

Whaling CS, Solis MM, Doupe AJ, Soha JA, Marler P (1997) Acoustic and neural bases for innate recognition of song. *Proceedings of the National Academy of Sciences USA,* **94**, 12694-12698.

White CMN (1965) *A Revised Check List of African Non-Passerine Birds.* Government Printer, Lusaka.

Whittaker RJ (1998) *Island Biogeography: Ecology, Evolution, and Conservation.* Oxford University Press, Oxford.

Wiens JJ (2004) Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution,* **58**, 193-197.

Wiens JJ, Graham CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology and Systematics,* **36**, 519-539.

Wiley EO (1978) The evolutionary species concept reconsidered. *Systematic Zoology,* **27**, 17-26.

Williamson MH (1981) *Island Populations.* Oxford University Press, Oxford.

Winker K, Glenn TC, Graves GR (1999) Dinucleotide microsatellite loci in a migratory wood warbler (Parulidae: *Limnothlypis swainsonii*) and amplification among other songbirds. *Molecular Ecology,* **8**, 1553-1556.

Winnepenninckx B, Backeljau T, De Watcher R (1993) Extraction of high molecular weight DNA from mollusks. *Trends in Genetics,* **9**, 407.

Wolters HE (1982) *Die Vogelarten der Erde.* Paul Parey, Hamburg & Berlin.

Wolters HE (1983) Zur systematik einiger passeres aus Kamerun. *Bonner Zoologische Beiträge,* **34**, 279-291.

Wong BBM, Keogh JS, Jennions MD (2004) Mate recognition in a freshwater fish: geographical distance, genetic differentiation, and variation in female preference for local over foreign males. *Journal of Evolutionary Biology,* **17**, 701-708.

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

Wright S (1951) The genetical structure of populations. *Annals of Eugenics,* **15**, 323-354.

Wright TF, Wilkinson GS (2001) Population genetics and vocal dialects in an Amazon parrot. *Proceedings of the Royal Society of London B,* **268**, 609-616.

Wu C-I (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology,* **14**, 851-865.

Yang SY, Patton JL (1981) Genic variability and differentiation in Galápagos finches. *Auk,* **98**, 230-242.

Yoder AD, Irwin JA, Payseur BA (2001) Failure of the ILD to determine data compatibility for slow loris phylogeny. *Systematic Biology,* **50**, 408-424.

Yuri T, Mindell DP (2002) Molecular phylogenetic analysis of Fringillidae, "New World nine-primaried oscines" (Aves: Passeriformes). *Molecular Phylogenetics and Evolution,* **23**, 229-243.

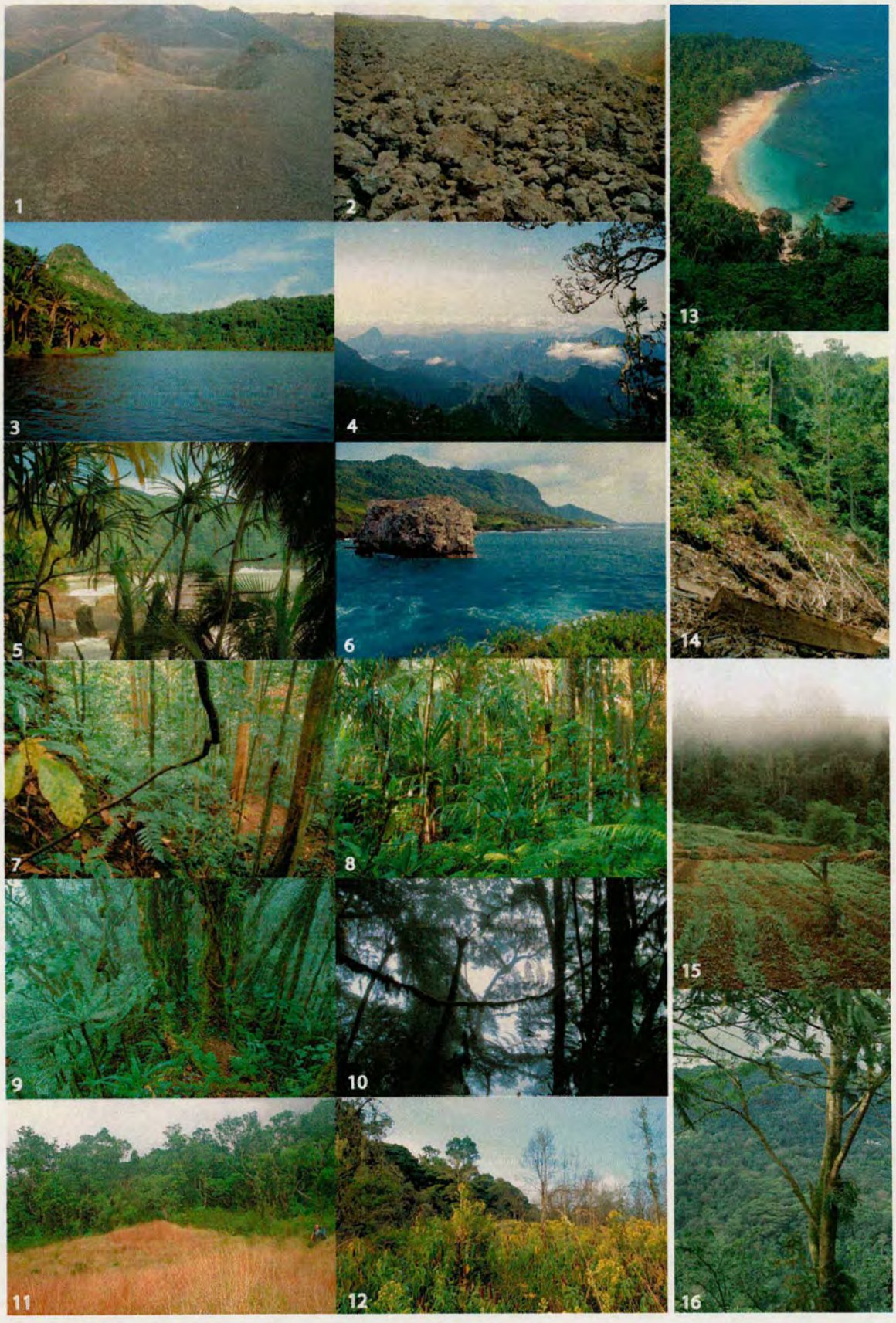
Zharkikh A (1994) Estimation of evolutionary distances between nucleotide sequences. *Journal of Molecular Evolution,* **39**, 315-329.

Zink RM (1985) Genetic population structure and song dialects in birds. *Behavioral and Brain Sciences,* **8**, 118-119.

Zink RM, Blackwell-Rago RC, Ronquist F (2000) The shifting roles of dispersal and vicariance in biogeography. *Proceedings of the Royal Society of London B,* **267**, 497-503.

Thesis appendices

APPENDIX 1.1. An overview of the Gulf of Guinea island system (legend on next page).



APPENDIX 1.1. Legend. The islands of the Gulf of Guinea system are of volcanic origin. Currently only Mount Cameroon remains active (1–2: crater and lava flow of 1999 eruption), but evidence of extended periods of volcanism can be seen on all islands, e.g., the crater lake of Annobón (3) and in the steep relief of the islands (4: São Tomé). The three oceanic islands are separated by seas of more than 1800m in depth, and therefore their biota were never physically connected (5: Príncipe; 6: Annobón). The habitat of the islands is rainforest, which is stratified with altitude: 0–800 m, lowland rainforest (7: São Tomé; 8: Príncipe); 800–1,400 m, montane forest; 1,400–2,500, mist forest (9–10: São Tomé) that is replaced above 2,500 m by sub-alpine meadows on Mt. Cameroon (11) and Pico Basilé on Bioko (12). Human impact affected mostly the lowland rainforest, which was replaced by sugar cane plantations, coconut groves (13: Príncipe) and oil palm plantations. This impact was much heavier in the northern, drier, parts of the islands. On São Tomé, fragments of what could have been the original forest persisted in river gullies, but these are fast disappearing for charcoal production (14). In the last decade, on São Tomé, the increase in horticulture has led to localised encroachment of the montane rainforest (15). Cocoa and coffee plantations are grown under 'shade-forests' that resemble the original forest in structure (16). As such, many endemic birds have adapted to these plantations. Besides their value for biodiversity, shade-forest plantations are vital for capturing and retaining water and prevent erosion.

APPENDIX 2.1. Taxa and sampling localities used in the study of the diversification of the Gulf of Guinea white-eyes (Zosteropidae).

Locality	Region	Taxon	Code	Sample	Collection
Mt Cameroon	Cameroon	S. melanocephalus	MEL001	Blood	M. Melo
Mt Cameroon	Cameroon	S. melanocephalus	MEL002	Blood	M. Melo
Mt Cameroon	Cameroon	Z. stenocricotus	STC001	Blood	M. Melo
Mt Cameroon	Cameroon	Z. stenocricotus	STC002	Blood	M. Melo
Bioko	Gulf of Guinea	S. brunneus	BRU001	Blood	M. Melo
Bioko	Gulf of Guinea	S. brunneus	BRU003	Blood	M. Melo
Bioko	Gulf of Guinea	Z. stenocricotus	STB001	Blood	M. Melo
Bioko	Gulf of Guinea	Z. stenocricotus	STB002	Blood	M. Melo
Principe	Gulf of Guinea	S. leucophaeus	LEU001	Blood	M. Melo
Principe	Gulf of Guinea	S. leucophaeus	LEU002	Blood	M. Melo
Principe	Gulf of Guinea	Z. ficedulinus ficedulinus	FIP001	Blood	M. Melo
Principe	Gulf of Guinea	Z. ficedulinus ficedulinus	FIP002	Blood	M. Melo
Sao Tome	Gulf of Guinea	S. lugubris	LUG033	Blood	M. Melo
Sao Tome	Gulf of Guinea	S. lugubris	LUG042	Blood	M. Melo
Sao Tome	Gulf of Guinea	S. lugubris	LUGMM2	Blood	M. Melo
Sao Tome	Gulf of Guinea	Z. ficedulinus feae	FIS002	Blood	M. Melo
Sao Tome	Gulf of Guinea	Z. ficedulinus feae	FIS003	Blood	M. Melo
Annobón	Gulf of Guinea	Z. griseovirescens	A001	Blood	M. Melo
Annobón	Gulf of Guinea	Z. griseovirescens	A002	Blood	M. Melo
Karisia, N Aberdares	Kenya	Z. poliogaster kikuensis	O8425	Blood	ZMUC
Mbololo, Teita	Kenya	Z. poliogaster silvanus	O8580	Blood	ZMUC
Mt Kenya	Kenya	Z. poliogaster kikuensis	RB2	Blood	R. Bowie
Mt Kulal, E Turkana	Kenya	Z. poliogaster kulalensis	O8629	Blood	ZMUC
Uhafiwa	Tanzania	Z. senegalensis stierlingi	O8255	Blood	ZMUC
Anjouan	Comoros	Z. maderaspatanus anjuanensis	BW253	Blood	B. Warren
Grande Comore	Comoros	Z. maderaspetanus kirki	BW146	Blood	B. Warren
Grande Comore	Comoros	Z. mouroniensis	BW140	Blood	B. Warren
Mayotte	Comoros	Z. maderaspetanus mayottensis	BW67	Blood	B. Warren
Moheli	Comoros	Z. maderaspetanus comorensis	BW127	Blood	B. Warren
La Reunion	Mascarenes	Z. borbonicus borbonicus	BWM54	Blood	B. Warren
La Reunion	Mascarenes	Z. olivaceus olivaceus	BWM49	Blood	B. Warren
Mauritius	Mascarenes	Z. olivaceus chloronotos	BWM28	Blood	B. Warren
Mauritius	Mascarenes	Z. borbonicus mauritianus	BWM24	Blood	B. Warren
Aldabra	Seychelles	Z. maderaspatanus aldabrensis	BW301	Blood	B. Warren
Conception	Granitic Seychelles	Z. modestus	BW344	Blood	B. Warren
Southern Antsiranana	Madagascar	Z. maderaspetanus maderaspetanus	393446	Tissue	FMNH
Southern Antsiranana	Madagascar	Z. maderaspetanus maderaspetanus	393447	Tissue	FMNH
Socotra	Gulf of Aden	Z. abyssinicus socotranus	BW292	Blood	B. Warren
Brisbane	Australia	Z. lateralis familiaris	SCB77	Blood	S. Clegg
Mt Trusmadi, Sabah	Borneo	Z. atricapillus atricapillus	B36434	Tissue	LSUMZ
captivity	(Japan)	Z. japonicus	B20880	Tissue	LSUMZ
Baidaihe, Hebei	China	Z. erythropleurus	O2653	Blood	ZMUC
captivity	(China)	Z. erythropleurus	O2776	Blood	ZMUC
Nepal	Nepal	Z. palpebrosus palpebrosus	RF1	Blood	R. Fleischer
Dinapique, Isabella	Philippines	Z. nigrorum aureiloris	O2663	Blood	ZMUC
Mt Pulong, Luzon	Philippines	Stachyris whiteheadi	O2410	Blood	ZMUC
Mt Pulong, Luzon	Philippines	Z. montanus whiteheadi	O2655	Blood	ZMUC
Mt Pulong, Luzon	Philippines	Z. montanus whiteheadi	O2662	Blood	ZMUC

FMNH: Field Museum of Natural History – LSUMZ: Louisiana State University Museum of Natural Science – ZMUC: Zoological Museum University of Copenhagen.

APPENDIX 2.2. Primers and PCR conditions for amplification and sequencing of mtDNA and nDNA regions in white-eyes (Zosteropidae).

Gene	Product	Primer[1]	Primer sequence 5' - 3'	PCR-mix				PCR-cycle						
				Mg mM	Primer µM	dNTP mM	Taq U	[PREMIT]	[DENAT	ANNEAL	EXT]	CYCLES	FINAL	SOURCE
MITOCHONDRIAL SEQUENCES														
ATPase6	661	L9245	CCTGAACCTGACCATGAAC	2.5	0.5	0.2	1.25	95° - 3'	95° - 45''	54° - 30''	72° - 45''	35	72° - 5'	1
		H9929	CTATGTGGGTAAGAGTGTGCTTGGTG											
ND3	351	L10755	GACTTCCAATCTTTAAAATCTGG	2.0	0.5	0.2	1.25	95° - 3'	95° - 45''	54° - 20''	72° - 45''	30	72° - 5'	2
		H11151	GATTTGTTGAGCCGAAATCAAC											
Cyt b	290	L14990	CCATCCAACATCTCAGCATGATGAAA	2.5	0.5	0.2	1.5	95° - 3'	94° - 45''	55° - 30''	72° - 45''	30	72° - 5'	3
		H15298	CCCTCAGAATGATATTTGTCCTCA											
NUCLEAR SEQUENCES														
b-Fibrinogen V	na	F: Fib5	CGCCATACAGAGTATACTGTGACAT	DID NOT AMPLIFY										3
		R: Fib6	GCCATCCTGGCGATTCTGAA											
b-Fibrinogen VII	1000	F: Fib7U	GGAGAAAACAGGACAATGACAATTCAC	3.0	0.5	0.2	1.25	94° - 2'	94° - 45''	50-54° - 1'	74° - 2'	30	74° - 15'	5
		R: Fib7L	TCCCCAGTAGTATCTGCCATTAGGGTT											
G3PDH XI	330	F: G3P13	TCCACCTTTGATGCGGGTGCTGGCAT	2.0	0.4	0.1	0.5	94° - 5'	94° - 40''	60° - 40''	72° - 1'	40	72° - 5'	6
		R: G3P14	AAGTCCACAACACGGTTGCTGTA											
MPP IV	310	F: MPP.4	TACATYTACTTYAAYACCTGGACCACCTG	2.0	0.5	0.2	1.25	94° - 7'	94° - 20''	54° - 20''	72° - 1'	40	72° - 7'	7
		R: MPP.5	AGATGGAGAGNAGGTTGGAGCCACA											
PEPCK9 IX	1000	F: GTP1601	ACGAGGCCTTTAACTGGCAGCA	2.0	0.5	0.2	1.25	94° - 7'	94° - 20''	54° - 20''	72° - 1'	40	72° - 7'	7
		R: GTP1793	CTTGGCTGTCTTTCCGGAACC											
PEPCK9 IX	672	F: PEPCK9	GGAGCAGCCATGAGATCTGAAGC	2.5	0.5	0.2	1.25	94° - 7'	94° - 20''	54° - 20''	72° - 1'	40	72° - 7'	7
		R: PEPCK9	GTGCCATGCTAAGCCAGTGGG											

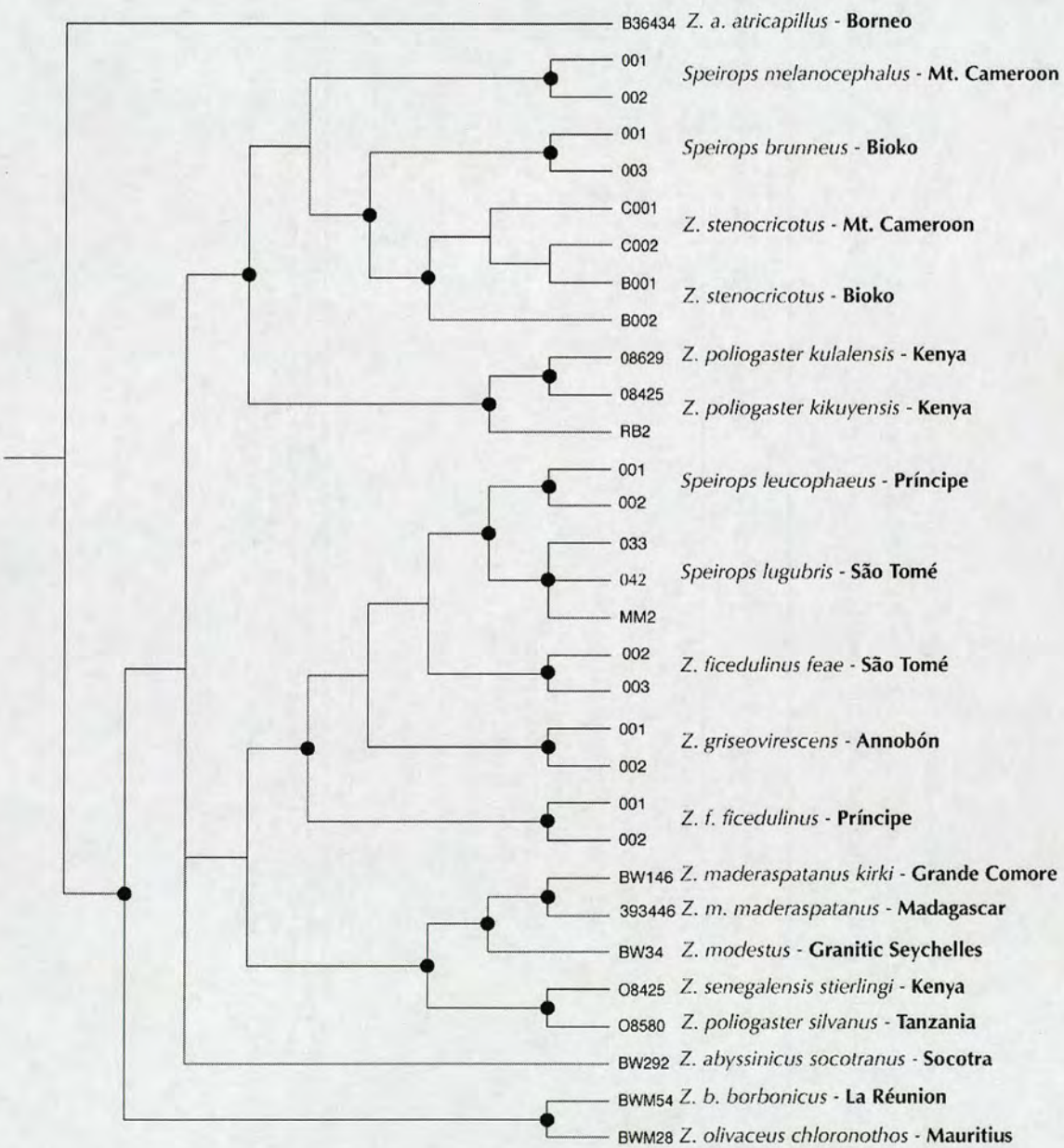
[1]:Primer name given as the position of the 3' end base on the chicken mtDNA genome (Desjardins & Morais 1990) – L: light strand, H: heavy strand; [2]:Primer name as in source – F: forward, R: reverse.
PREMIT: Initial denaturation; DENAT: Denaturation; ANNEAL: Annealing; EXT: Extension; FINAL: Final extension.
SOURCE: 1 – Eldredge Bermingham Lab: http://striweb.si.edu/bermlab.htm; 2 – Chesser (1999); 3 – Palumbi (1996); 4 – Rauri Bowie *pers. comm.*; 5 – Prychitko & Moore (1997); 6 – Fjeldså *et al.* (2003); 7 – McCracken & Sorenson (2005).

APPENDIX 2.3. Primers and PCR conditions for microsatellite loci that successfully cross-amplified in Gulf of Guinea Zosteropidae.
The first 6 loci were the ones chosen for the inference of the relationships among Gulf of Guinea taxa.
Allele numbers are based on a sample of 77 individuals (8 taxa, 9 populations) for the first 6 loci, and on a sample of 12 individuals (5 taxa) for the last 6 loci.
[Primer ZL35 did not amplify. Primer Ase48 amplified a polymorphic product but with too many stutter bands.]

Locus		Primer sequence 5' - 3' (fluoro-labelled sequences are underlined)	PCR-mix				PCR-cycle						ALLELES	SOURCE
			Mg mM	Primer µM	dNTP mM	Taq U	[PREMIT] 3'	[DENAT 30''	ANNEAL 30''	EXT] 30''	CYCLES	[FINAL] 10'		
ZL12	F -	6-FAM-CCGGGGAATATTTGGCTTCGG	1.0						55.2				13	1
	R -	GAGGAGCCAGAGGAGCCGG												
ZL18	F -	6-FAM-CTGCAGAGGAGCTCAGGTAAC	1.0						55.2				3	1
	R -	CTCGGTGCTGCCAGAACTCAG												
ZL22	F -	6-FAM-ATTAATATCAACACTTTCTATGCATC	1.5						61.3				5	1
	R -	GTCAGTCATAGGTTGGATTTCATG		0.5	0.2	1.25	94	94		72	30	72		
ZL44	F -	6-FAM-CTGTCCCTGCCCTCTCATC	1.5						61.3				5	2
	R -	ACCATGGCAGAGGCACCAA												
Ase19	F -	6-FAM-TAGGGTCCCAGGGAGGAAG	1.5						55.2				3	3
	R -	TCTGCCCATTAGGGAAAAGTC												
Ase64	F -	6-FAM-CCACCTTTCATACTGGGGAG	2.0						63.7				12	3
	R -	TTCAGCCAGTCAGTGTAGCC												
ZL38	F -	CCTCAAGGTTAACCACTATAGAC	1.5										2	1
	R -	GTAGTAGTATCTTCTGCATCAAGG												
ZL46	F -	GTCAGTGCTGTGCTTTGAT	1.5										1	1
	R -	AACCTGAAATTACACTTCT												
ZL50	F -	AAGGTGCCGAGGTCCTGT	1.0										2	1
	R -	TTTGCATGAGTGCATGCTGG		0.5	0.2	1.25	94	94	55.2	72	30	72		
ZL54	F -	CACGACTTCTCAAGCAGAC	1.5										1	2
	R -	GAGCCTTGCACAAACGGAC												
Ase18	F -	ATCCAGTCTTCGCAAAAGCC	1.5										1	3
	R -	TGCCCCAGAGGGAAGAAG												
Ase28	F -	AAATCCTTGGGAAAGTTTGTGA	1.5										1	4
	R -	TTGTATGTATGGGGAATGAAGGT												

PREMIT: Initial denaturation; DENAT: Denaturation; ANNEAL: Annealing; EXT: Extension; FINAL: Final extension. Temperatures in ° Celsius.
SOURCE: 1 - Degnan *et al.* (1999); 2 - Frentiu *et al.* (2003); 3 - Richardson *et al.* (2000); 4 - Non-published primer developed by 3, available at: D. Dawnson BIRMARKER Homepage, Sheffield Molecular Genetics Facility.

APPENDIX 2.4. Relationships between the subset of taxa used for the Bayesian inference of the area of origin and colonisation sequence of the Gulf of Guinea oceanic islands, as implemented in BAYESMULTISTATE (Pagel *et al.* 2004). Presented is the 50% majority consensus tree of 500 trees obtained from the analyses of the combined data set (ATP6, ND3, cyt *b*) using a single chain with a cooling algorithm to explore the parameter space, as implemented in BAYESPHYLOGENIES (Pagel & Meade 2004; trees sampled at 10,000 steps apart, after the chain had reached stationarity). Black dots indicate node support (posterior probability > 0.95). The same relationships were recovered as those inferred with the complete data set under a MC^3 search implemented in MRBAYES (Fig. 2.5).



APPENDIX 3.1. Primers and PCR conditions of mitochondrial markers used to infer the phylogenetic relationships of the São Tomé grosbeak *Neospiza concolor*, the Príncipe seedeater *Serinus rufobrunneus* and the thick-billed seedeater *S. burtoni*.

Rauri Bowie (see 'Collaborations') sequenced also the cytochrome *b* and the 5th intron of the β-fibrinogen gene, a nuclear intron, to infer the relationships within the *Serinus* genus.

Gene	Product	Primer name	Primer sequence 5'-3'	Source	PCR-mix				PCR-cycle					
					Primer µM	Mg mM	dNTP mM	Taq U	[PREMIT]	[DENAT	ANNEAL	EXT]	CYCLES	FINAL
12S-1	591	L1267	F: AAAGCATGACACTGAAGATG	1	0.5	2.5	0.2	1.25	94° - 2'	94° - 45''	56° - 45''	72° - 1'	30	72° - 5'
		H1858	R: TCGATTACAGAACAGGCTCCTCTAG	1										
12S-2	400	L1753	F: AAACTGGGATTAGATACCCCACTAT	1	0.5	2.5	0.2	1.25	94° - 2'	94° - 45''	56° - 45''	72° - 1'	30	72° - 5'
		H2157	R: GAGGAGGGTGACGGGCGG	2										
ND2-1	547	L5219	F: CCCATACCCCGAAAATGATG	1	0.5	2.5	0.2	1.25	94° - 2'	94° - 45''	56° - 45''	72° - 1'	30	72° - 5'
		H5766	R: GGATGAGAAGGCTAGGATTTTKCG	1										
tRNAs[1]	545	L6136	F: TTCTACCTTCGTCTTGCNTAYTG	2	0.5	2.5	0.2	1.25	94° - 2'	94° - 45''	55° - 45''	72° - 1'	30	72° - 5'
		H6681	H: GGTATAGGGTGCCGATGTCTTTGTG	1										
ATP8[2]	507	L8929	F: GGMCARTGCTCAGAAATCTGYGG	3	0.5	2.5	0.2	1.25	94° - 2'	94° - 45''	56° - 45''	72° - 1'	30	72° - 5'
		H9436	R: GGATTAGGGCTCATTTGTGTC	2										
ATP6	559	L9370	F: CGATGAATTACTAACCGMCTHTC	2	0.5	2.5	0.2	1.25	94° - 2'	94° - 45''	56° - 45''	72° - 1'	30	72° - 5'
		H9929	R: CTATGTGGTAAGAGTGTGCTTGGTG	2										

Primer names for mitochondrial markers given as the position of the 3' end base on the chicken mtDNA genome (Desjardins & Morais 1990) - L: light strand, H: heavy strand; Primer names for intron 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 - Yuri & Mindell (2002); 3 - Eldredge Bermingham Lab: http://striweb.si.edu/bermlab.htm.
1: tRNAs - fragment including 150 bp of the 3' end of the ND2 and the following transfer RNAs: tryptophane, alanine, aspargine, cysteine, and tyrosine. 2: Includes also 85 bp of the 3'end of CO2 and tRNA-Lysine.

APPENDIX 4.1. Taxa and populations of the *Turdus olivaceus* species complex and *Turdus* outgroups, analysed in this study.

Taxon	Region	Locality	Sample	Collection	Code
Turdus philomelos	Germany	Schelswig-Holstein	Blood	LSUMZ	B13468
T. hortulorum	Russia	Primorskiy Kray		UWBM	51161
T. ruficollis	Russia	Avt. Rep. Gorno-Altay		UWBM	46282
T. iliacus	Germany	Schelswig-Holstein		LSUMZ	B13478
T. grayi	Honduras	Dpto. Copan		MBM	6620
T. plebejus	Nicaragua	Matagalpa		MBM	4322
T. infuscatus	Guatemala	Quetzaltenango		MBM	13588
T. libonyanus	South Africa	KwaZulu Natal	Blood	UWBM	52923
T. bewsheri comorensis	Comoro Is.	Grande Comore	Blood	B. Warren	GC159
T. bewsheri moheliensis	Comoro Is.	Mohéli	Blood	B. Warren	GC193
T. olivaceofuscus olivaceofuscus	Gulf of Guinea	São Tomé	Blood	M. Melo	MD1
T. olivaceofuscus xanthorhynchus	Gulf of Guinea	Príncipe	Blood	M. Melo	P2-036
T. pelios saturatus	Ghana	Buipe		FMNH	396636
T. pelios centralis	D. R. Congo	Lwiro		FMNH	429758
T. olivaceus olivaceus	South Africa	Cape Town	Blood	FitzPatrick	#31
T. olivaceus pondoensis	South Africa	Himeville	Blood	FitzPatrick	#65
T. olivaceus transvaalensis	South Africa	Dullstroom	Blood	FitzPatrick	#1
T. olivaceus smithi	South Africa	Kimberley	Blood	MBM	5877
T. olivaceus milanjensis	Mozambique	Mount Namuli	Blood	M. Melo	BD84702
T. olivaceus helleri 1	Kenya	Taita Hills	Blood	L. Lens	09
T. olivaceus helleri 2	Kenya	Taita Hills	Blood	L. Lens	20
T. olivaceus roelhi 1	Tanzania	W. Usambara Mts.		ZMUC	JK01-201099
T. olivaceus roelhi 2	Tanzania	W. Usambara Mts.		FMNH	356763
T. olivaceus nyikae 1	Tanzania	Udzungwa Scarp Forest		ZMUC	O7887
T. olivaceus nyikae 2	Tanzania	Udzungwa Scarp Forest		ZMUC	O7865
T. olivaceus baraka	Uganda	Kabale		FMNH	385051
T. olivaceus abyssinicus	Kenya	Aberdares	Blood	R. Bowie	BB0165

LSUMZ: Louisiana State University Museum of Natural Science - ZMUC: Zoological Museum of the University of Copenhagen - FMNH: The Field Museum of Natural History - RMCA: Royal Museum for Central Africa - MBM: Barrick Museum at the University of Nevada Las Vegas - UWBM: University of Washington Burke Museum - FitzPatrick: Percy FitzPatrick Institute, University of Cape Town.

APPENDIX 4.2. Primers and PCR conditions used to infer phylogenetic relationships among *Turdus* species.

Gene	Product	Primer name	Primer sequence 5'-3'	Source	PCR-mix				PCR-cycle					
					Primer µM	Mg mM	dNTP mM	Taq U	[PREMIT]	[DENAT	ANNEAL	EXT]	CYCLES	FINAL
ND2	1108	L5204	F: GCTAACAAAGCTATCGGGCCCAT	1	0.5	2.5	0.2	1.25	94° - 2'	94° - 30''	50/52° - 30''	72° - 1'	30	72° - 5'
		H6312	R: CTTATTTAAGGCTTTGAAGGCC	1										
ND2	Seq. primers	ND2-L1	F: CTGGCTTTCTCATCCATCTC	2										
		ND2-H1	R: AGGTGGGAGATGGATGAG	2										
ND3	351	L10755	F: GACTTCCAATCTTTAAAATCTGG	3	0.5	2.5	0.2	1	94° - 2'	94° - 30''	50/52° - 30''	72° - 1'	30	72° - 5'
		H11151	R: GATTTGTTGAGCCGAAATCAAC	3										

Numbers in primer names refer to the position of the 3' end base on the chicken mtDNA genome (Desjardins & Morais 1990) - L: light strand, H: heavy strand; Primer names for intron as in source - F: forward, R: reverse. PREMIT: Initial denaturation; DENAT: Denaturation; ANNEAL: Annealing; EXT: Extension; FINAL: Final extension.
Source: 1 - Cicero & Johnson (2001); 2 - Bowie *et al.* (2005); 3 - Chesser (1999).

APPENDIX 4.3. Datasets used for the morphological analyses of the Gulf of Guinea thrush. Field measurements obtained by M. Melo. Museum measurements obtained by Nigel Collar.

FIELD				AMERICAN MUSEUM OF NATURAL HISTORY			
Population	**Code**	**Sex**	**Locality**	**Population**	**Code**	**Sex**	**Locality**
Príncipe	P2-036	M	Ribeira Porco	Príncipe	266082	M	Southwest
São Tomé	ST1-153	M	Alto Douro	Príncipe	266083	M	Southwest
São Tomé	ST1-304	M	Alto Douro	Príncipe	266085	F	Southwest
São Tomé	ST2-247	M	Alto Douro	São Tomé	265653	M	na
São Tomé	ST2-248	M	Alto Douro	São Tomé	264938	M	na
São Tomé	ST2-255	M	Alto Douro	São Tomé	264932	M	na
São Tomé	ST3-025	M	Contador dam	São Tomé	264944	F	na
São Tomé	ST3-036	M	Contador dam	São Tomé	264945	F	na
São Tomé	ST3-040	M	Contador dam	São Tomé	268516	F	na
São Tomé	ST3-051	M	Contador dam	São Tomé	264946	F	na
São Tomé	ST3-059	M	Contador dam				
São Tomé	ST3-063	M	Contador dam				
São Tomé	ST1-325	M	Quija				
São Tomé	ST2-133	M	Quija				
São Tomé	ST2-257	F	Alto Douro				
São Tomé	ST2-077	F	Lagoa Amélia				
São Tomé	ST2-109	F	Lagoa Amélia				
São Tomé	ST2-222	F	Lagoa Amélia				
São Tomé	ST2-223	F	Lagoa Amélia				
São Tomé	ST2-224	F	Lagoa Amélia				
São Tomé	ST2-227	F	Lagoa Amélia				
São Tomé	ST1-234	F	Quija				
São Tomé	ST2-136	F	Quija				
São Tomé	ST3-177	F	Quija				
São Tomé	ST3-224	F	Quija				

APPENDIX 5.1. Samples used in the study of the Gulf of Guinea kingfishers (Alcedinidae), with sampling locality. MHNH = Muséum National d'Histoire Naturelle, Paris.

Taxon	Locality	Sample type	Collection	Code
Alcedo atthis	France	Liver	MNHN CG 1995-24	22-53
Alcedo cristata 1	Malawi	Blood	C. Spottiswoode	AF100
Alcedo cristata 2	Malawi	Blood	C. Spottiswoode	AF101
Alcedo leucogaster C1	Cameroon	Blood	MNHN unvouchered	3-02
Alcedo leucogaster C2	Cameroon	Blood	MNHN unvouchered	01-14
Alcedo leucogaster B1	Bioko	Blood	M. Melo	B215
Alcedo leucogaster B2	Bioko	Blood	M. Melo	B219
Alcedo [nais] 1	Príncipe	Blood	M. Melo	P1-138
Alcedo [nais] 2	Príncipe	Blood	M. Melo	P1-220
Alcedo quadribrachys	Cameroon	Blood	MNHN unvouchered	03-06
Alcedo [thomensis] 1	São Tomé	Blood	M. Melo	ST1-039
Alcedo]thomensis] 2	São Tomé	Blood	M. Melo	ST1-353
Ceyx erithacus	Laos	Blood	MNHN unvouchered	4-1G
Ispidina lecontei	Cameroon	Blood	M. Melo	C21
Ispidina picta 1	Cameroon	Blood	MNHN unvouchered	01-25
Ispidina picta 2	Gabon	Blood	M. Melo	G32
Halcyon malimbica	Cameroon	Blood	MNHN unvouchered	02-09
Megaceryle maxima	Cameroon	Blood	M. Melo	C219

APPENDIX 5.2. Primers and PCR conditions used in the study of the Gulf of Guinea kingfishers. Myo2 is a nuclear intron; all other are mitochondrial genes.

Gene	Product	Primer name	Primer sequence 5'-3'	Source	PCR-mix				PCR-cycle					
					Primer µM	Mg mM	dNTP mM	Taq U	[PREMIT]	[DENAT	ANNEAL	EXT]	CYCLES	FINAL
ND2-1	547	L5219	F: CCCATACCCCGAAAATGATG	1	0.5	2.5	0.2	1.25	94° - 2'	94° - 45''	56° - 45''	72° - 1'	30	72° - 5'
		H5766	R: GGATGAGAAGGCTAGGATTTTKCG	1										
ND2-2	555	L5758	F: GGCTGAATRGGMCTNAAYCARAC	1	0.5	2.5	0.2	1.25	94° - 2'	94° - 45''	56° - 45''	72° - 1'	30	72° - 5'
		H6313	R: CTCTTATTTAAGGCTTTGAAGGC	1										
ND3	351	L10755	F: GACTTCCAATCTTTAAAATCTGG	2	0.5	2.5	0.2	1	95° - 3'	95° - 45''	56° - 30''	72° - 1'	30	72° - 5'
		H11151	R: GATTTGTTGAGCCGAAATCAAC	2										
ATP8	507	L8929	F: GGMCARTGCTCAGAAATCTGYGG	3	0.5	2.5	0.2	1.25	94° - 2'	94° - 45''	56° - 45''	72° - 1'	30	72° - 5'
		H9436	R: GGATTAGGGCTCATTTGTGTC	4										
ATP6	559	L9370	F: CGATGAATTACTAACCGMCTHTC	4	0.5	2.5	0.2	1.25	94° - 2'	94° - 45''	56° - 45''	72° - 1'	30	72° - 5'
		H9929	R: CTATGTGGTAAGAGTGTGCTTGGTG	4										
Myo2	c. 700	Myo2 *or*	F: GCCACCAAGCACAAGATCCC	5	0.3	1.5	0.2	0.3	94° - 3'	94° - 30''	56° - 30''	72° - 30''	35	72° - 5'
		Myo2-Pi	F: CCTGTCAAATATCTGGAGGTATG	6										
		Myo3F	R: GCAAGGACCTTGATAATGACTT	7										

Primer names for mitochondrial markers given as the position of the 3' end base on the chicken mtDNA genome (Desjardins & Morais 1990) – L: light strand, H: heavy strand; Primer names for intron 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 – Chesser (1999); 3 – Eldredge Bermingham Lab: http://striweb.si.edu/bermlab.htm; 4 – Yuri & Mindell (2002); 5 – Heslewood *et al.* (1998); 6 – Fuchs *et al.* (in press); 7 – Slade *et al.* (1993).

APPENDIX 6.1. Primers and PCR conditions for 15 passerine microsatellite loci used in the population genetic study of the Príncipe seedeater.

Locus	Primer sequence 5'-3' (fluoro-labelled sequences are underlined)	PCR-mix				PCR-cycle						ALLELES	SOURCE
		Primer µM	Mg mM	dNTP mM	Taq U	[PREMIT] 2′	[DENAT 30″	ANNEAL 30″	EXT] 30″	CYCLES	[FINAL] 10′		
Ase42	F: NED-CATGGGTAGGTTGGGATGTC R: AGGTGAGGGTATGCAAACATG							60				7	1
Ase43	F: 6-FAM-ATTGTGTGGGATTTGCAT R: TTGCTGTGCAGTTTGCTTTT							TD 60-50				12	1
Ase48	F: 6-FAM-TTTATTTCCTGGACTGGAACAATC R: GAACATTGGGCTACTGGGC							57				17	1
Cuµ04	F: 6-FAM-AATTGCATAAATGTGATCCAC R: AAATGAAATGTGGTAGAATTCC							55				7	2
Cuµ28	F:6-FAM-GAGGCACAGAAATGTGAATT R:TAAGTAGAAGGACTTGATGGCT							60				13	2
Gf08	F: 6-FAM-TGGGAGAGCAAGGTGGGAACAG R: TGGAGTGGTGATTAAACCAGCAGG							60				3	3
Gf10	F: Fluoro-TTGAGGGTCCCATCCAACTG R: GGTGCTCTTGATACAAAGCATAA							50				3	3
LOX1	F: 6-FAM-ATGATGGTAAGTCTAATGAAAGC R: CCACACACATTCACTCTATTG	0.4	1.5	0.2	0.5	94	94	53	72	35	72	40	4
LOX3	F: 6-FAM-TTCTGTGGTGAAGTTTTCTGGAG R: CCAACCCATTCCATGACAAC							60				74	4
LOX7	F: 6-FAM-AACCTAAGCACATTTATTCAGC R: AACAAATAACATAGGTCAGAAGC							55				76	4
LOX8	F: 6-FAM-TTGTGAAGGTTTGGGACATAAG R: AGTTGAGGCCATTAAAAAGATTC							57				71	4
Lswµ14	F: 6-FAM-GTTATGCTCCAACAAAATAGATA R: AGGTTTTRAAGGATAGATTTATA							49				7	5
Lswµ18	F: 6-FAM-TTGCTGAAAGAAGTACTAAGA R: CTGKTTGCAGGATATGTATAC							55				12	5
Pdoµ4	F: 6-FAM-CGATAAGCTTGGATCAGGACTAC R: CTTGGGAAGAGAATGAGTCAGGA							53				9	6
WBSW7	F: 6-FAM-TCTGGAGTTCTGGGACCTGT R: CTCACTCAACAGCAGGACCA							54				4	7

PREMIT: Initial denaturation; DENAT: Denaturation; ANNEAL: Annealing; EXT: Extension; FINAL: Final extension. TD: touch-down. Temperatures in ° Celsius.
ALLELES: number of alleles in a sample of 217 birds (from São Tomé, Príncipe and Boné).
SOURCE: 1 - Richardson *et al.* (2000); 2 - Gibbs *et al.* (1999); 3 - Petren (1998); 4 - Piertney *et al.* (1998); 5 - Winker *et al.* (1999); 6 - Neumann & Wetton (1996); 7 - McRae & Amos 1999)

Appendix 6.2. Cross-amplification success in nine *Serinus* species of the 15 microsatellite loci that were polymorphic in the Príncipe seedeater. The number of individuals tested is given below each species name abbreviation; number of alleles amplified is given; numbers in brackets indicate number of successful amplifications when lower than the total of individuals tested. Alb: White-throated canary *S. albogularis*; Bur: Thick-billed seedeater *S. burtoni*; Can: Cape canary *S. canicollis*; Fla: Yellow canary *S. flaviventris*; Moz: Yellow-fronted canary *S. mozambicus*; Oid: African citril *S. citrinelloides*; Pec: Lemon-breasted canary *S. citrinipectus*; Sul: Brimstone canary *S. sulphuratus*; Tot: Cape siskin *S. totta*.

Locus	Alb 5	Bur 5	Can 1	Fla 13	Moz 2	Oid 2	Pec 3	Sul 1	Tot 1
Ase42	4	1	2	11	2	2	3	1	1
Ase43	2	1	1	3	1	1	2	1	1
Ase48	7	6	0	14	4	3	4	2	2
Cuμ04	6	3	1	10	3	4	2	1	1
Cuμ28	4	3	1	9	0	3	6	2	2
Gf08	1	1	1	2	1	1	1	1	1
Gf10[z]	2	1	1	2	1	1	3	1	1
LOX1	8	8	2	16	4	4	5	2	2
LOX3	9	8	2	20	1 (1)	4	2	2	2
LOX7	3 (3)	8	2	14 (10)	0	4	4 (2)	2	2
LOX8	6 (3)	6	1	18	2 (1)	2 (1)	5	1	1
Lswμ14	4	3	1	7 (8)	0	0	1 (1)	1	0
Lswμ18	2	0	1	7	2 (1)	3	6	0	2
Pdoμ4	5	2	1	8	1	2	2 (1)	2	2
WBSW7	1	5	1	5 (12)	1 (1)	2	3	2	2

Extra appendices

Molecular Ecology Notes (2006) doi: 10.1111/j.1471-8286.2006.01510.x

PRIMER NOTE

Identification of 15 polymorphic microsatellite loci in the Príncipe seedeater (*Serinus rufobrunneus*) and assessment of their utility in nine other *Serinus* species (Fringillidae, Aves)

MARTIM MELO* and BENGT HANSSON*†
*Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, West Mains Road, EH9 3JT Edinburgh, UK,
†Department of Animal Ecology, Ecology Building, Lund University, SE-223 62 Lund, Sweden.

Abstract

We tested 74 passerine microsatellite loci for cross-amplification in the Príncipe seedeater (*Serinus rufobrunneus*), and identified 15 loci that were both polymorphic and easy to score. In a sample of 113 individuals, the number of alleles ranged between three and 71. Three loci deviated from Hardy–Weinberg equilibrium after correcting for multiple tests, and one locus had high estimated null allele frequency. These 15 loci were highly successful in amplifying polymorphic products also in nine other *Serinus* species.

Keywords: avian cross-species amplification, canary, Fringillidae, microsatellite, seedeater, *Serinus*

Received 24 May 2006; revision accepted 20 June 2006

The Gulf of Guinea islands (West Africa) constitute a spectacular centre of endemism, and the bird endemism in particular appears to have no parallel worldwide (Jones & Tye 2006). The Príncipe seedeater (*Serinus rufobrunneus*) is an endemic songbird with populations on São Tomé and Príncipe, located 150 km apart, and on a small islet, 3 km from Príncipe. These three populations show sufficient morphological differentiation to be classified as subspecies (Jones & Tye 2006). This species was chosen as a model to investigate population level processes that promote divergence and eventually speciation and microsatellite data were required to estimate gene flow between and within islands.

We tested the utility of 74 passerine microsatellites in the Príncipe seedeater. Because this species has sometimes been placed in a separate monotypic genus (*Poliospiza*; Jones & Tye 2006), we tested the successful primers on nine other *Serinus* species in order to identify loci useful for a wider research community.

Blood samples were collected nondestructively from mist-netted wild individuals and preserved in ethanol. DNA was extracted using the DNeasy Tissue extraction kit (QIAGEN). Initially, we tested the primers in three Príncipe seedeaters, one from each island. The (polymerase chain reaction) PCR mix (Bioline PCR kit) contained 4 pmol of each primer, $1 \times NH_4$-buffer, 15 nmol $MgCl_2$, 2 nmol dNTP, 0.5 U Bio*Taq* DNA polymerase and 10–25 ng template in 10 μL reaction volume. PCR conditions (MJ Research PTC-220 DNA Engine Dyad) were as follows: 94 °C for 2 min, then 35 cycles at 94 °C for 30 s/T_a for 30 s/72 °C for 30 s, followed by 72 °C for 10 min; where T_a is the locus-specific annealing temperature (Table 1). The fluorescent-labelled PCR products were separated and alleles were detected in an ABI PRISM 3730 capillary sequencer (Applied Biosystems).

Of the 74 primer pairs tested in the three test individuals, we deemed that 19 were potentially useful since they had apparently polymorphic products of expected size (some had additional unspecific products). The remaining loci did not amplify (10 loci), amplified complex and sometimes weak products (12), were monomorphic (29) or had weak but potentially polymorphic products of expected size and additional strong unspecific products (4). The 19 apparently polymorphic loci were further tested in 96 individuals from São Tomé, and it became evident that one locus (Lswμ7) was monomorphic and that it was difficult to achieve scoreable products for three loci (Cuμ02, Gf02 and LOX2). Thus we identified 15 useful loci and these were scored in another 17 individuals from São Tomé and tested on 33 individuals from nine other *Serinus* species. Results for all primers and species have been deposited in the BIRDMARKER database (http://www.shef.ac.uk/misc/groups/molecol/deborah-dawson-birdmarkers.html).

Correspondence: Martim Melo, Fax: +44 (0) 131 6506564; E-mail: martim.melo@ed.ac.uk

Table 1 Characteristics of 15 passerine microsatellite loci that were polymorphic in the Príncipe seedeater (*Serinus rufobrunneus*)

Locus	EMBL Accession no.	Primer sequence 5′–3′ (fluoro-labelled sequences are underlined)	T_a (°C)	n	A	H_O	H_E	HW P-value	Null alleles	Allele size range
Ase42	AJ276644	F: NED-CATGGGTAGGTTGGGATGTC R: AGGTGAGGGTATGCAAACATG	60	111	7	0.81	0.74	0.987	0.00	373–387
Ase43	AJ276645	F: 6-FAM-ATTGTGTGGGATTTGCAT R: TTGCTGTGCAGTTTGCTTTT	TD 60–50	110	12	0.69	0.83	0.001	0.07	234–262
Ase48	AJ276777	F: 6-FAM-TTTATTTCCTGGACTGGAACAATC R: GAACATTGGGCTACTGGGC	57	113	13	0.79	0.88	0.002	0.04	272–330
Cuμ04	AF122891	F: 6-FAM-AATTGCATAAATGTGATCCAC R: AAATGAAATGTGGTAGAATTCC	55	113	7	0.72	0.71	0.636	0.00	126–140
Cuμ28	AF122894	F: 6-FAM-GAGGCACAGAAATGTGAATT R: TAAGTAGAAGGACTTGATGGCT	60	112	13	0.86	0.88	0.260	0.01	176–202
Gf08	AF081932	F: 6-FAM-TGGGAGAGCAAGGTGGGAACAG R: TGGAGTGGTGATTAAACCAGCAGG	60	110	3	0.17	0.23	0.010	0.07	94–97
Gf10^Z	AF081934	F: Fluoro-TTGAGGGTCCCATCCAACTG R: GGTGCTCTTGATACAAAGCATAA	50	66^Z	3	0.32	0.34	0.500	0.02	204–210
LOX1	Y16820	F: 6-FAM-ATGATGGTAAGTCTAATGAAAGC R: CCACACACATTCACTCTATTG	53	113	35	0.96	0.95	0.691	0.00	286–446
LOX3	Y16822	F: 6-FAM-TTCTGTGGTGAAGTTTTCTGGAG R: CCAACCCATTCCATGACAAC	60	111	71	0.99	0.98	0.257	0.01	226–432
LOX7	Y16824	F: 6-FAM-AACCTAAGCACATTTATTCAGC R: AACAAATAACATAGGTCAGAAGC	55	112	59	0.98	0.98	0.695	0.00	278–436
LOX8	Y16825	F: 6-FAM-TTGTGAAGGTTTGGGACATAAG R: AGTTGAGGCCATTAAAAAGATTC	57	112	62	0.92	0.97	0.005	0.01	350–540
Lswμ14	AF129095	F: 6-FAM-GTTATGCTCCAACAAAATAGATA R: AGGTTTTRAAGGATAGATTTATA	49	92	7	0.33	0.52	0.001	0.13	200–214
Lswμ18	AF129096	F: 6-FAM-TTGCTGAAAGAAGTACTAAGA R: CTGKTTGCAGGATATGTATAC	55	111	8	0.51	0.69	0.001	0.09	208–222
Pdoμ4	X93505	F: 6-FAM-CGATAAGCTTGGATCAGGACTAC R: CTTGGGAAGAGAATGAGTCAGGA	53	113	28	0.77	0.82	0.077	0.00	198–338
WBSW7	AF130434	F: 6-FAM-TCTGGAGTTCTGGGACCTGT R: CTCACTCAACAGCAGGACCA	54	113	4	0.57	0.55	0.708	0.00	140–146

Loci sources: Ase, Richardson *et al.* (2000); Cuμ, Gibbs *et al.* (1999); Gf, Petren (1998); LOX, Piertney *et al.* (1998); Lswμ, Winker *et al.* (1999); Pdoμ, Neumann & Wetton (1996); WBSW, McRae & Amos (1999). ^{Z}, Z-linked locus. T_a, annealing temperature (TD, touchdown cycle); n, number of successful amplifications out of 113 (Gf10^Z is Z-linked, therefore only males were considered for diversity estimates); A, number of alleles; H_O and H_E, observed and expected heterozygosities; HW P-value, P-value of the Hardy–Weinberg test implemented in FSTAT (Goudet 2001), adjusted nominal level = 0.001; null alleles estimated in FREENA (Chapuis & Estoup, submitted).

The characteristics of the 15 polymorphic loci, based on the sample of 113 Príncipe seedeaters from São Tomé, are summarized in Table 1. The number of alleles ranged between three and 71, and the expected heterozygosity between 0.23 and 0.98. The highly variable LOX-loci from the Scottish crossbill (*Loxia scotica*, Piertney *et al.* 1998) were also extremely variable in the Príncipe seedeater (35–71 alleles), which may be explained by the close relationship between these genera (Yuri & Mindell 2002). Gf10 was Z-linked as all females were 'homozygous' (hemizygous).

Three loci departed from Hardy–Weinberg equilibrium after Bonferroni correction (six before correction) in tests implemented in FSTAT version 2.9 (Table 1; Goudet 2001). Estimations of null allele frequencies done with FREENA (Chapuis & Estoup, submitted) showed low levels for most loci and only Lswμ14 had frequency over 10% (Table 1). Nevertheless null alleles do not appear to induce biases in studies of population differentiation in Príncipe seedeaters (M. Melo, unpublished data). Tests of linkage equilibrium between all pairs of autosomal loci (i.e. excluding Gf10) performed in ARLEQUIN version 2.00 (Schneider *et al.* 2000) revealed a single significant deviation after Bonferroni correction (between Lswμ14 and Lswμ18).

Finally, several primers successfully amplified polymorphic loci in nine other *Serinus* species, and in a sample of 13 *Serinus flaviventris* all 15 loci were in fact polymorphic (Table 2). Senar *et al.* (2006) found another six primers that amplified polymorphic products in *Serinus citrinella*, and together these two panels of microsatellites constitute a valuable resource for future research in the genus *Serinus*.

Table 2 Cross-amplification success of the 15 polymorphic loci on nine *Serinus* species. The number of individuals tested, *n*, is given below each species name abbreviation; number of alleles amplified is given; numbers in brackets indicate number of successful amplifications when lower than the total of individuals tested. Alb, white-throated canary *S. albogularis*; Bur, thick-billed seedeater *S. burtoni*; Can, Cape canary *S. canicollis*; Fla, yellow canary *S. flaviventris*; Moz, yellow-fronted canary *S. mozambicus*; Oid, African citril *S. citrinelloides*; Pec, lemon-breasted canary *S. citrinipectus*; Sul: brimstone canary *S. sulphuratus*; Tot, Cape siskin *S. totta*

Locus	Alb	Bur	Can	Fla	Moz	Oid	Pec	Sul	Tot
n	5	5	1	13	2	2	3	1	1
Ase42	4	1	2	11	2	2	3	1	1
Ase43	2	1	1	3	1	1	2	1	1
Ase48	7	6	0	14	4	3	4	2	2
Cuμ04	6	3	1	10	3	4	2	1	1
Cuμ28	4	3	1	9	0	3	6	2	2
Gf08	1	1	1	2	1	1	1	1	1
Gf10	2	1	1	2	1	1	3	1	1
LOX1	8	8	2	16	4	4	5	2	2
LOX3	9	8	2	20	1 (1)	4	2	2	2
LOX7	3 (3)	8	2	14 (10)	0	4	4 (2)	2	2
LOX8	6 (3)	6	1	18	2 (1)	2 (1)	5	1	1
Lswμ14	4	3	1	7 (8)	0	0	1 (1)	1	0
Lswμ18	2	0	1	7	2 (1)	3	6	0	2
Pdoμ4	5	2	1	8	1	2	2 (1)	2	2
WBSW7	1	5	1	5 (12)	1 (1)	2	3	2	2

Acknowledgements

We thank S. Andersson, R. Covas, P. Jones and C. Spottiswoode for providing *Serinus* spp. samples. M. Melo is supported by a PhD scholarship from the Fundação para a Ciência e a Tecnologia (Portugal), and B. Hansson by a Marie-Curie postdoctoral fellowship and the Swedish Research Council.

References

Chapuis MP, Estoup A (submitted) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, in press.

Gibbs HL, Tabak LM, Hobson K (1999) Characterization of microsatellite DNA loci for a Neotropical migrant songbird, the Swainson's thrush (*Catharus ustulatus*). *Molecular Ecology*, **8**, 1551–1552.

Goudet J (2001) *FSTAT, A Program to Estimate and Test Gene Diversities and Fixation Indices*. Version 2.9.3. Available from http://www.unil.ch/popgen/softwares/fstat.htm.

Jones PJ, Tye A (2006) *The Birds of São Tomé and Príncipe with Annobón: Islands of the Gulf of Guinea*. BOU Checklist Series 22. British Ornithologists Union, Oxford.

McRae SB, Amos W (1999) Characterization of hypervariable microsatellites in the cooperatively breeding white-browed sparrow weaver *Plocepasser mahali*. *Molecular Ecology*, **8**, 903–904.

Neumann K, Wetton JH (1996) Highly polymorphic microsatellites in the house sparrow *Passer domesticus*. *Molecular Ecology*, **5**, 307–309.

Petren K (1998) Microsatellite primers from *Geospiza fortis* and cross-species amplification in Darwin's finches. *Molecular Ecology*, **7**, 1782–1784.

Piertney SB, Marquiss M, Summers R (1998) Characterization of tetranucleotide microsatellite markers in the Scottish crossbill (*Loxia scotica*). *Molecular Ecology*, **7**, 1261–1263.

Richardson DS, Jury FL, Dawson DA *et al.* (2000) Fifty Seychelles warbler (*Acrocephalus sechellensis*) microsatellite loci polymorphic in Sylviidae species and their cross-species amplification in other passerine birds. *Molecular Ecology*, **9**, 2226–2231.

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN: A Software for Population Genetics Data Analysis*. Version 2.100. Genetics and Biometry Laboratory. Department of Anthropology. University of Geneva, Geneva.

Senar JC, Borras A, Cabrera J, Cabrera T, Björklund M (2006) Local differentiation in the presence of gene flow in the citril finch *Serinus citrinella*. *Biology Letters*, **2**, 85–87.

Winker K, Glenn TC, Graves GR (1999) Dinucleotide microsatellite loci in a migratory wood warbler (Parulidae: *Limnothlypis swainsonii*) and amplification among other songbirds. *Molecular Ecology*, **8**, 1553–1556.

Yuri T, Mindell DP (2002) Molecular phylogenetic analysis of Fringillidae, 'New World nine-primaried oscines' (Aves: Passeriformes). *Molecular Phylogenetics and Evolution*, **23**, 229–243.

Journal of Biogeography (J. Biogeogr.) (2006)

ORIGINAL ARTICLE

Freshwater paths across the ocean: molecular phylogeny of the frog *Ptychadena newtoni* gives insights into amphibian colonization of oceanic islands

G. John Measey[1]*, Miguel Vences[2], Robert C. Drewes[3], Ylenia Chiari[4], Martim Melo[5] and Bernard Bourles[6]

[1]*Laboratoire d'Ecologie des Sols Tropicaux, Institut de Recherche pour le Développement, 32 Avenue Henri Varagnat, 93143 Bondy Cedex, France,* [2]*Zoological Institute, Technical University of Braunschweig, Spielmannstrasse 8, 38106 Braunschweig, Germany,* [3]*Department of Herpetology, California Academy of Sciences, 875 Howard Street, San Francisco, CA 94103, USA,* [4]*Zoological Museum of Amsterdam, Mauritskade 61, 1092 AD Amsterdam, The Netherlands,* [5]*Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh EH9 3JT, UK and* [6]*Institut de Recherche pour le Développement, Technopole Pointe du Diable, BP 70, 29280 Plouzane, France*

*Correspondence: G. John Measey, Laboratory of Animal Ecology, Department of Biology, University of Antwerp, B-2610 Antwerp, Belgium.
E-mail: john@measey.com

ABSTRACT

Aim Amphibians are a model group for studies of the biogeographical origins of salt-intolerant taxa on oceanic islands. We used the Gulf of Guinea islands to explore the biogeographical origins of island endemism of one species of frog, and used this to gain insights into potential colonization mechanisms.

Location São Tomé and Príncipe, two of the four major islands in the Gulf of Guinea, West Africa, are truly oceanic and have an exceptionally high biodiversity.

Methods Mitochondrial DNA is used to test the endemic status of a frog from São Tomé and compare it with congeneric taxa from tropical Africa. Existing data on surface currents, surface salinity, atmospheric circulation and bird migration in the Gulf of Guinea are summarized to address hypotheses concerning colonization mechanisms.

Results The endemic status of *Ptychadena newtoni* (Bocage) is supported here by mitochondrial DNA sequences, and analysis of this and other molecular data indicates that an East African species close to *Ptychadena mascareniensis* (Duméril and Bibron) is its nearest relative. We refute the possibility that this population was anthropogenically introduced, in favour of a natural dispersal mechanism.

Main conclusions With six endemic frogs and one caecilian, the Gulf of Guinea islands harbour a diverse amphibian fauna. Five of these species appear to have their closest relatives in East Africa. Insufficient evidence exists for transportation by storms, birds or rafts alone. However, we propose a synergy of rafting, favourable surface currents and a reduction in salinity of surface waters. Catastrophic events, or wet periods in climatic history, could allow freshwater paths to open far enough to enable continental flora and fauna to reach these and other isolated oceanic islands.

Keywords
Amphibia, Anura, Cameroon line, colonization, dispersal, Gulf of Guinea, Gymnophiona, island biogeography, phylogeny.

INTRODUCTION

Amphibians are usually described as poor dispersers (Blaustein *et al.*, 1994) with limited osmotic tolerance (Balinsky, 1981; Duellman & Trueb, 1986). In most cases their populations show strong phylogeographical structuring (Avise, 2000), which indicates low dispersal. Therefore amphibian distribution patterns are usually hypothesized to be the result of vicariance (e.g. Feller & Hedges, 1998; Biju & Bossuyt, 2003). However, recent evidence for transoceanic dispersal (Hedges *et al.*, 1992; Evans *et al.*, 2003; Vences *et al.*, 2003, 2004) and long-distance dispersal over land (Smith & Green, 2005) suggests that the long-standing biogeographical debate on vicariance vs. dispersal should be

reconsidered, not only for amphibians (see McGlone, 2005; de Queiroz, 2005).

Intolerance of salt water explains why amphibians are normally absent from oceanic islands that have had no past connections to other land masses (Darwin, 1859). Some volcanic islands and archipelagos in the Lesser Antilles do harbour endemic frogs (Kaiser *et al.*, 1994) but, as with other Caribbean islands, they are located in an area with a complex tectonic history that is still disputed (e.g. Crother & Guyer, 1996) and that may have had past sub-aerial connections (e.g. MacPhee *et al.*, 2000). Similar controversies surround islands in the Sunda, Philippine and Pacific areas (see Vences *et al.*, 2003), many of which harbour endemic amphibians.

Vences *et al.* (2003) discovered that the two species of frog on Mayotte, an island of the Comoro archipelago in the Indian Ocean, are endemic. They provided one of the best documented examples of endemic amphibians on fully volcanic, oceanic islands, nevertheless their claim that these were the only such examples was erroneous: another indisputable example of dispersal of amphibians over a marine barrier is provided by the Gulf of Guinea islands of São Tomé and Príncipe. These volcanic islands, surrounded by deep waters of > 1800 m, harbour seven endemic amphibians: six frogs (Amphibia: Anura) and one caecilian (Amphibia: Gymnophiona). The presence of a caecilian on São Tomé is particularly interesting because they are largely subterranean amphibians that are considered highly unlikely to disperse over major marine barriers (Nussbaum, 1984; Gower *et al.*, 2002).

Although Gulf of Guinea island amphibians have been known since the mid-19th century, the significance of the islands in demonstrating amphibian overseas dispersal (Fahr, 1993) has been overlooked in several relevant accounts of amphibian biogeography (e.g. Savage, 1973; Duellman & Trueb, 1986; Hedges *et al.*, 1992; Vences *et al.*, 2003). Rafting, favoured by storms and sea currents, has been proposed as a possible mechanism of colonization of islands by amphibians (Myers, 1953; Savage, 1973; Kaiser *et al.*, 1994; Krause *et al.*, 1997; Censky *et al.*, 1998; Evans *et al.*, 2003), despite the osmotic problems of organisms with limited tolerance of salinity (Balinsky, 1981). Other aerial dispersal mechanisms depend on the occurrence of storms (especially tornadoes), which can transport smaller ontogenic stages (Simons, 1996; Elsom, 1988) or perhaps involve passive dispersal on volant vertebrates such as aquatic birds (cf. Darwin, 1859).

Here we highlight the importance of the Gulf of Guinea islands as an area for studies of overseas dispersal, especially in amphibians. We review geographical patterns in the Gulf of Guinea that could lead to dispersal events. We provide molecular confirmation of the endemicity to São Tomé of one of its amphibians [*Ptychadena newtoni* (Bocage)], rejecting the possibility that it was introduced anthropogenically, and we investigate phylogenetic relationships with other species of *Ptychadena* Boulenger. Lastly, we discuss our results in the light of three specific hypotheses for amphibian overseas dispersal: rafting, storms and bird carriage.

LOCATION AND METHODS

Description of study area – the Gulf of Guinea islands

Four islands, Bioko (aka Fernando Pó), Príncipe, São Tomé and Annobón (aka Pagalu), are the offshore part of the Cameroon line in the Gulf of Guinea (Burke, 2001). The swells on the Cameroon line are the result of past volcanic activity (Fitton, 1987; Burke, 2001), but currently only Mount Cameroon is active, having last erupted in June 2000. Bioko is separated from the West African mainland by only 32 km and sea depths of < 60 m, which suggests linkage to the mainland during the last glaciation (Lee *et al.*, 1994). This is consistent with the continental character of its biota (Mertens, 1964, 1965; Jones, 1994); hence it will not be considered further here. The other three islands have been isolated since their origins (Jones, 1994). The islands furthest away from Mt Cameroon are also most distant from the mainland, with minimum distances of at least 220 km and sea depths of > 1800 m (Table 1; Fig. 1). Dates for the origins of the three oceanic islands have been given as 31, 13 and 4.8 Ma (Príncipe, São Tomé and Annobón, respectively; Lee *et al.*, 1994), but the oldest lava flow-eruption ages give information only on the minimum age when each island was sub-aerial. Furthermore, volcanic activity persisted in all islands until as recently as the last 0.1 Ma in São Tomé (Fitton, 1987) or < 0.4 Ma (Caldeira & Munhá, 2002), although the impact of volcanic activity on the biota of these islands is unknown (but cf. Jesus *et al.*, 2005c). Therefore when and how the land surface area (and hence colonizable habitats) changed through time is only hinted at by the ages available.

Patterns of biodiversity

Isolation led to the evolution of a highly original flora, characterized by the co-occurrence of neoendemics and of Afromontane palaeoendemics (Figueiredo, 1994). A recent biogeographical study on the begonias suggested that São Tomé functioned as an important pre-Pleistocene refuge for these and possibly other plants (Plana *et al.*, 2004). The forests of São Tomé and Annobón have the highest fern diversity and density in Africa (Exell, 1944). The islands also form an important centre of endemism for many faunal groups, which gives them a unique and exceptional biota of global conservation significance (Jones, 1994; Juste & Fa, 1994).

The Gulf of Guinea islands centre of endemism is particularly striking in birds, reptiles and amphibians (Table 1). The three oceanic islands, with an area of *c.* 1000 km^2, hold 29 endemic bird species, with up to four endemic genera on São Tomé: *Amaurocichla* Sharpe, *Dreptes* Reichenow, *Thomasophantes* Hartlaub and *Neospiza* Salvadori; and another on Príncipe, *Horizorhinus* Oberholser. In comparison, the 13 main islands of the Galápagos archipelago, with an area totalling *c.* 8000 km^2, have 22 endemic species. The number of endemic birds makes up one-third of the endemics of the large Guinean Forests hotspot (Bakkar *et al.*, 1999).

Table 1 Physical geography and endemicity of plants and vertebrates* of the historically isolated islands in the Gulf of Guinea

	Príncipe	São Tomé	Annobón
Surface area (km^2)	139	857	17
Highest peak (m)	948	2024	654
Distance from closest island (km)	146	146	180
Distance from mainland Africa (km)	220	255	340
Plants	314 [26] (11)	601 [81] (14)†	208 [14] (6)
Mammals	5 [1] (0)	10 [3] (0)	2 [0] (0)
Birds	33 [6] (5)†	50 [15] (5)†††	11 [2] (1)
Fish*	23	14	N/A
Reptiles	13 [3] (6)	14 [1] (6)	1 [0] (1)
Amphibians*	3 [2] (2)	5 [4] (2)	0 [0] (0)
Amphibian species	*Hyperolius molleri, Leptopelis palmatus, Phrynobatrachus dispar*	*Schistometopum thomense, Hyperolius thomensis, Hyperolius molleri, Phrynobatrachus* sp. *nov., Ptychadena newtoni*	None

*Only resident birds were considered. Fish refer only to freshwater species; the difference in numbers of fish species from São Tomé and Príncipe may result solely from the proportion of time spent sampling on each island. Reptiles do not include marine turtles. The previously recognized endemic frog genus *Nesionixalus* has recently been synonymized with *Hyperolius* (Drewes & Wilkinson, 2004). Numbers of species are given, those in square brackets refer to single-island endemics; in round brackets to endemics from two or three of the islands.
†Indicates presence of an endemic genus. Data from the Gulf of Guinea Islands' Biodiversity Network (http://www.ggcg.st), Feiler (1988), Fahr (1993), Feiler *et al.* (1993), Haft (1993), Nill (1993), Figueiredo (1994), Jesus *et al.* (2003), Jones & Tye (2006), T. Iwamoto (personal communication).

Each of the three oceanic islands is classified as an Endemic Bird Area (Stattersfield *et al.*, 1998), with São Tomé and Príncipe being the only small oceanic islands in the world's top 25% of Endemic Bird Areas.

As expected for oceanic islands, most endemic mammals are bats (Juste & Ibañez, 1994), but São Tomé has an endemic shrew [*Crocidura thomensis* (Bocage)] and Príncipe has an endemic subspecies of *Crocidura poensis* (Fraser). The shrews have been said to represent a 'zoogeographical mystery' (Dutton & Haft, 1996) because their metabolism and surface-to-volume ratio necessitates that they eat constantly, making them unlikely dispersers (Heim de Balsac & Hutterer, 1982). Most fish are primarily marine (with two introduced) occurring in estuarine habitats, and all known freshwater fish are secondarily adapted (T. Iwamoto, personal communication).

The reptiles and amphibians of the Gulf of Guinea islands pose some of the most difficult questions with respect to colonization. Many reptiles and nearly all amphibians are considered to have low tolerance of sea water (Balinsky, 1981). Jesus *et al.* (2005a,b,c) studied the molecular phylogenetics of the Gulf of Guinea islands' geckos and skinks, finding that although one species [*Hemidactylus mabouia* (Moreau de Jonnès)] appears to have been introduced, neither geckos nor skinks are monophyletic, suggesting multiple independent colonizations of the islands for both groups.

Fahr (1993) assumed accidental transport of eggs by waterbirds as the most likely origin for the endemic frogs of São Tomé, but the Gulf of Guinea islands are not part of any bird migration route (Jones & Tye, 2006), and the only regular visitors are Palaearctic coastal waders and seabirds [common sandpiper, *Actitis hypoleucos* (L.), whimbrel *Numenius phaeopus* (L.), and several tern species, *Sterna* spp.]. Palaearctic landbird vagrants recorded on the islands are mostly small passerines, and Afrotropical migrants do not use the islands at all (Jones & Tye, 2006). Birds that are associated with wetland habitats, and thus are the best candidates for introducing eggs of some amphibians, include the mainland ancestor of the endemic forest ibises (*Bostrychia* sp.) and aquatic species such as herons and gallinules. The ibises have their closest relatives in West Africa (Cameroon and Gabon) (Chapin, 1923; Brown *et al.*, 1982). The common moorhen [*Gallinulla chloropus* (L.)], a near-cosmopolitan species, and the green-backed heron [*Butorides striatus* (L.)], a pantropical species, have established populations on São Tomé and Príncipe in marshes and the lower reaches of rivers (Jones & Tye, 2006). As populations of these species never differentiated from mainland populations, the probability of occasional movements between the islands and the nearby mainland (e.g. Cameroon or Gabon) is high.

Drewes (2002) commented on the bizarre occurrence of a high proportion of subterranean taxa amongst the herpetofaunal endemics: the caecilian *Schistometopum thomense* (Bocage) (Fig. 1, inset); three scolecophidian snakes, *Rhinotyphlops newtoni* (Bocage), *Rhinotyphlops feae* (Boulenger) and *Typhlops elegans* Peters; and a legless skink, *Feylinia polylepis* (Bocage). Like many other burrowing lower vertebrates, each of these species can be found within or beneath rotting logs on São Tomé and Príncipe (G.J.M. and R.D., personal observations), and this hints at the possible mechanism of their dispersal.

Geographical patterns around the Gulf of Guinea

Surface currents

Surface currents in the Atlantic have been determined recently as the result of a combined analysis of historical ship drifts,

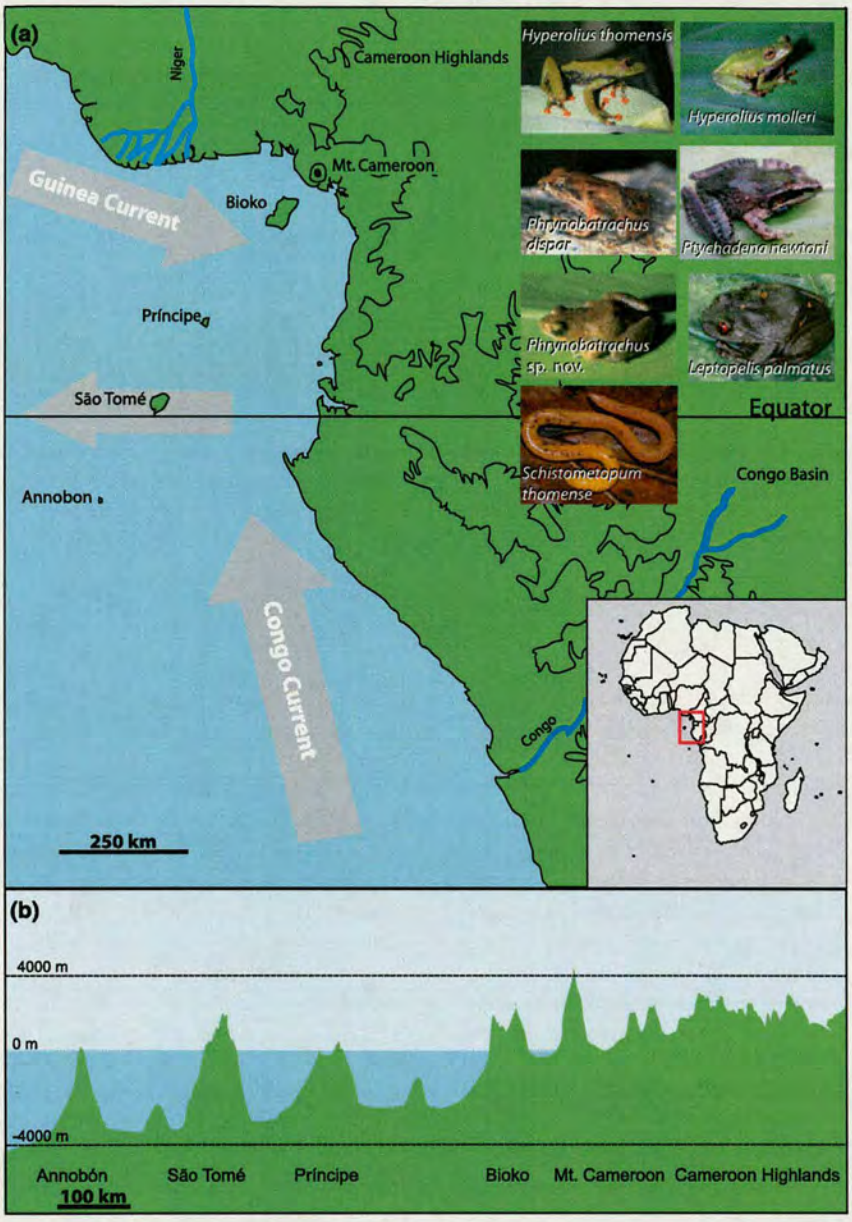


Figure 1 The Gulf of Guinea islands, lying in the Cameroon line, consist of four principal islands. (a) Arrows indicate the major sea-surface currents in the area, demonstrating the potential for objects exiting the Niger or Congo (the only rivers shown, with 500 m contours) to be taken directly to São Tomé and Príncipe. All amphibians from these islands (inset) are considered endemic, including *Ptychadena newtoni*, the closest known relatives of which appear to be East African. (b) Bioko (aka Fernando Pó) is part of the continental shelf, while Príncipe, São Tomé and Annobón (aka Pagalu) are truly oceanic islands.

hydrographical data, and surface-drifting buoy trajectories (Richardson & Walsh, 1986; Arnault, 1987; Stramma & Schott, 1999; Lumpkin & Garzoli, 2005). In summary, two perpendicular currents feature (northwards and eastwards), following the shoreline into the Gulf of Guinea and towards the Bay of Biafra (Fig. 1), and these converge to form a third westward current that carries water out along the line of the equator (see also Dupont *et al.*, 2000).

The eastward 'Guinea Current' carries surface water into the Gulf of Guinea along the southern coast of West Africa (see also Feiler, 1988; Haft, 1993), as a continuation of the North Equatorial Countercurrent (Fig. 1). This current extends to the Bay of Biafra where it becomes more diffuse, turning back westward around the equator. Moving northwards along the south-west African coast, the 'Congo Current' feeds the equatorial branch of the South Equatorial Current (Stramma

& Schott, 1999), clearly visible in ship drift data, and the mean annual Ekman surface currents. This current carries coastal waters offshore, including discharge from the Congo River (Wauthy, 1983). These opposing currents produce a swirling effect, followed by intensification westward of the South Equatorial Current.

Salinity of surface waters

Due to the difference in density, freshwater floats on underlying oceanic salted waters. Hence large rivers in spate can have a profound effect on the surface salinity of seawater in the Gulf of Guinea, and are known as tropical surface waters. These are characterized by high temperatures and low salinity, and overlie a density-discontinuous layer at the thermocline, promoting a strong and nutrient-rich upwelling effect (Binet & Marchal, 1993). The salinity of the subsurface layers reaches a maximum around the Gulf of Guinea islands, further strengthening the stratifying effect. Dessier & Donguy (1994) demonstrated large seasonal variations corresponding to the rainy seasons in the catchments of the two major river systems, the Niger and the Congo (the latter is the world's second largest river in terms of discharge volume, after the Amazon). Figure 2a shows how surface fresh waters are confined to the mouths of the Niger and Congo during the dry season. At the peak of the rainy season, the rivers' combined freshwater discharges, together with high precipitation into the Bay of Biafra, affect surface waters in the entire Gulf of Guinea (Fig. 2b). This forms a characteristic 'freshwater path' leaving the mouth of the Congo, also known as the Congo plume.

Atmospheric circulation

Modern atmospheric circulation in the Gulf of Guinea has been recently and succinctly summarized by Dupont *et al.* (2000), therefore only a brief description is given here. Light winds generally follow and are responsible for surface currents (see above). In addition to these, and of more importance in possible colonization routes from the mainland, are the seasonal Inclined Meterological Equator (IME) and Mid-Tropospheric Easterly (MTE) winds. In January and February, when the IME is at its southernmost position, the north-east trade winds blow off the Liberian coast into the Gulf of Guinea. In the boreal summer, the African Easterly winds (including MTE) dominate weather systems in West Africa (Berry & Thorncroft, 2005). Storms generated from the MTE move east to west, providing a potential colonization route from Gabon, the nearest land to the east of the Gulf of Guinea islands (Fig. 1). The MTE winds were presumed to be responsible for carrying pollen to sites in the Gulf of Guinea (Dupont *et al.*, 2000).

Study species

Frogs of the genus *Ptychadena* Boulenger, or ridged frogs, are widely distributed in Nilotic Egypt and sub-Saharan Africa (excluding south-western South Africa) and currently consist of 47 species (Frost, 2004). They are medium-sized frogs, typically 50–60 mm snout-vent length, and are often abundant in wetland areas and irrigated agricultural landscapes (Channing, 2001). Most often they are seen as an arcing flash disappearing with a plop into standing water. *Ptychadena newtoni* (Bocage) was first described as a São Tomé endemic, although it was synonymized with *Ptychadena oxyrhynchus* (Smith) by Andersson (1937). Later, Guibé & Lamotte (1957) placed it in the *Ptychadena mascareniensis* (Duméril and Bibron) group (Frost, 2004). However, both Perret (1976) and Loumont (1992) considered it a distinct species, and hence an island endemic, based on morphological characteristics including its rough dorsal skin, reduced dorsal folds, and a reported body size of 76 mm, the largest within the genus.

Molecular analyses of *Ptychadena newtoni*

A fragment of the mitochondrial 16S rRNA gene was amplified by PCR and sequenced from 30 specimens of the genus *Ptychadena*, in order to complement the data set of Vences *et al.* (2004), following methods outlined therein. These included tissue samples from four specimens of the presumed endemic *P. newtoni*, collected during a recent expedition to the

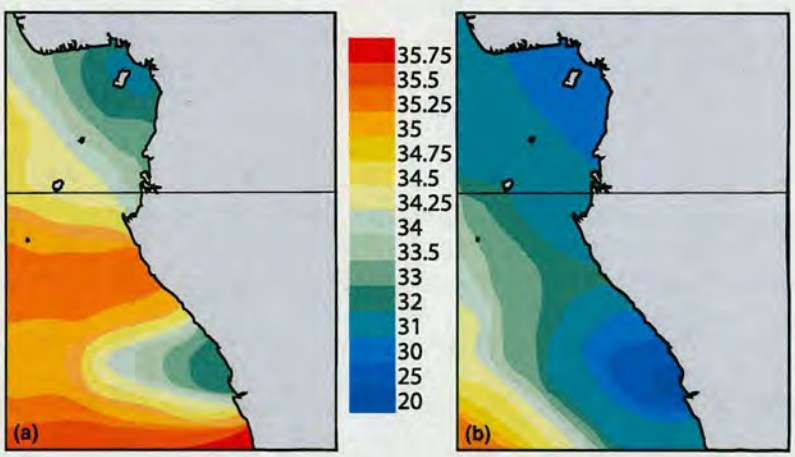


Figure 2 Sea-surface salinity (in Practical Salinity Units (psu), which is roughly equivalent to parts per thousand of salt) for the Gulf of Guinea in (a) July and (b) February, demonstrating the dramatic changes in surface salinity depending on direct rainfall and in the drainages of the Niger and Congo. Values below 30 psu characterize brackish waters and are seen extending far into the ocean, even in years of average rainfall. Data from the Mercator project (http://www.mercator-ocean.fr).

Gulf of Guinea islands (see Drewes & Wilkinson, 2004, Fig. 1 inset). For analysis we excluded haplotypes that were identical or very similar to each other (< 1% uncorrected pairwise sequence divergence); among these was the haplotype of *Ptychadena* aff. *mascareniensis* B from Cameroon, which had only slight differences from the included sequence of the same taxon from Benin.

In order to increase phylogenetic resolution within the *P. mascareniensis* complex, we sequenced fragments of the genes for cytochrome *b* (cyt *b*) and cytochrome oxidase subunit I (cox1) from representative individuals of the major lineages within the complex, and from two additional *Ptychadena* species, to obtain a data set of 1648 base pairs, using primers and PCR protocols as described by Bossuyt & Milinkovitch (2000), Chiari *et al.* (2004) and Vences *et al.* (2005). PCR conditions were 94°C for 90 s, then 10 cycles of 94°C for 30 s, 45°C (increasing 0.8°C each cycle) for 45 s, 72°C for 60 s, then 25 cycles of 94°C for 30 s, 53°C for 45 s, 72°C for 60 s, and a final extension of 10 min at 72°C for cyt *b*; and 94°C for 90 s, 35 cycles of 94°C for 30 s, 50°C for 45 s, 72°C for 60 s, final extension of 10 min at 72°C for cox1.

Fragments obtained from the amplified genes were edited and aligned manually using SEQUENCE NAVIGATOR software (Applied Biosystems Inc., Foster City, CA, USA). No gaps were present in the cox1 and cyt *b* alignments. From the 16S rDNA alignment, we excluded some areas of high variability, which needed inclusion of gaps to account for indels, and we excluded all positions with single gaps in one or more of the sequences. Maximum parsimony and maximum likelihood phylogenetic analyses were performed using PAUP* (Swofford, 2002). The Akaike information criterion, as implemented in MODELTEST (Posada & Crandall, 1998), was used to find the best-fitting substitution model for our data. To test alternative phylogenetic hypotheses, we used (a) Shimodaira–Hasegawa tests (Shimodaira & Hasegawa, 1999) which, according to Shimodaira (2002), are safe to use for this purpose but may have conservative bias; and (b) approximately unbiased tests (Shimodaira, 2002). Both statistics were calculated using the program CONSEL(Shimodaira & Hasegawa, 2001). New DNA sequences were deposited in GenBank under accession numbers DQ525918–65 (Table 2). The alignments and maximum likelihood trees were deposited in TreeBASE (http://www.treebase.org) with accession numbers M2757–8 and S1533.

RESULTS

Sequences of 16S rDNA of a total of 105 individual *Ptychadena* were available for analysis. After removal of identical or very closely related (putatively conspecific) haplotypes, and of gapped positions in the alignment, the data set consisted of 522 characters in 30 taxa. Of these, 330 were constant and 144 were parsimony-informative. MODELTEST selected a general time-reversible substitution model plus invariant sites plus gamma (GTR + I + G) with empirical base frequencies and substitution rates, a proportion of invariable sites of 0.4904 and a gamma-distribution shape parameter of 0.5131 as best fitting the data. Heuristic searches using this substitution model under the maximum likelihood optimality criterion, with 10 random addition sequences of taxa, found two trees of equal likelihood score ($-\ln L = 3739.7$). A strict consensus of these is shown in Fig. 3.

Sequences of cyt *b* and cox1 from a total of eight individuals (plus one outgroup, *Rana nigromaculata* Hallowell) were combined with 16S rDNA sequences of the same individuals for analyses. Of a total of 1648 characters in the combined data set, 1014 were constant and 442 were parsimony-informative. A general time-reversible substitution model (GTR + I + G) with a proportion of invariable sites of 0.4475 and a gamma-distribution shape parameter of 0.5793 was selected by MODELTEST as best fitting the data.

In the analysis based on 16S rDNA sequences, *P. newtoni* is found with high bootstrap support (80–93%) in a clade containing *P. mascareniensis* and several cryptic species so far considered to be *P. mascareniensis*. Interestingly, its sister species is the form named *P.* aff. *mascareniensis* A by Vences *et al.* (2004), a species occurring in eastern Africa (Tanzania, Kenya, Uganda and Egypt); however, the bootstrap support for this grouping was low (54–57%; Fig. 3). All individuals of *P. newtoni* had identical haplotypes. Their uncorrected pairwise sequence divergences to *P.* aff. *mascareniensis* A were 5.2–5.6%. The general topology of the tree recovered agrees with that published by Vences *et al.* (2004), which had a smaller taxon sampling but included more haplotypes of some taxa. Differences were found, for instance, in the basal relationships among major lineages in the *P. mascareniensis* complex, but the competing hypotheses were not supported by relevant bootstrap values in either analysis, and therefore are not discussed further.

The phylogenetic position of *P. newtoni* was strongly confirmed by the combined analysis of three genes in a reduced set of taxa. Its sister-group relationship with *P.* aff. *mascareniensis* A from eastern Africa was supported by 99% bootstrap values in both maximum parsimony and maximum likelihood analyses (Fig. 4). Separate analyses of the cyt *b* and cox1 data sets resulted in the same relationships (not shown). Both Shimodaira–Hasegawa and approximately unbiased tests significantly rejected ($P < 0.05$) nine alternative phylogenetic hypotheses in which *P. newtoni* was placed in all possible positions within and basal to the *P. mascareniensis/pumilio* clade, leaving the remainder of the topology unchanged.

DISCUSSION

Ptychadena newtoni endemicity

While our data do not resolve the taxonomy of the *P. mascareniensis* clade (which evidently contains many distinct as well as cryptic species), they do strongly suggest that *P. newtoni* is a distinct species endemic to São Tomé. The previously identified morphological characters, including its particularly large size, add credence to our molecular results. Moreover, this may be an example of island gigantism

Table 2 DNA sequences of the 16S rRNA gene obtained in addition to those published by Vences *et al.* (2004), and of the coxl and cytochrome *b* genes obtained for the reduced set of taxa

Species	Origin	Locality	Voucher	Accession 16S	Accession coxl	Accession cyt *b*
Ptychadena aff. *aequiplicata*	Cameroon	Dja Reserve	CAS 199182	DQ525919	–	–
Ptychadena anchietae	Kenya	Kararacha Pond	CAS 214837	DQ525920	–	–
Ptychadena anchietae	Somalia	Karin, Bari Region	CAS 227562	DQ525921	–	–
Ptychadena anchietae	Somalia	Karin, Bari Region	CAS 227507	DQ525922	–	–
Ptychadena anchietae	South Africa	Mtunzini	No voucher	AF215404*	DQ525952	DQ525961
Ptychadena mahnerti	Kenya	Mt. Kenya	SL 171	DQ525918	–	–
Ptychadena mascareniensis	Madagascar	Nahampoana	ZSM 190/2002	AY517587*	DQ525951	DQ525960
Ptychadena aff. *mascareniensis* A	Uganda	Lake Victoria,	MVZ 234085	DQ525923	–	–
Ptychadena aff. *mascareniensis* A	Kenya	Makuru	MVZ 223624	DQ525924	–	–
Ptychadena aff. *mascareniensis* A	Kenya	Mt. Kenya	MVZ 234087	DQ525925	–	–
Ptychadena aff. *mascareniensis* A	Kenya	Mt. Kenya	MVZ 234086	DQ525926	–	–
Ptychadena aff. *mascareniensis* A	Kenya	Taita Hills	CAS 191517	DQ525927	–	–
Ptychadena aff. *mascareniensis* A	Kenya	Taita Hills	CAS 191518	DQ525928	–	–
Ptychadena aff. *mascareniensis* A	Tanzania	Kibebe farm	AC 2087	DQ525929	DQ525950	DQ525959
Ptychadena aff. *mascareniensis* C	Guinea	No precise locality	gu03.2	DQ525930	DQ525948	DQ525957
Ptychadena aff. *mascareniensis* D	Uganda	Kampala	MVZ 234084	DQ525931	–	–
Ptychadena aff. *mascareniensis* E	Central African Republic	Dzanga-Sangha Reserve	MOR DS 52	DQ525932	DQ525947	DQ525956
Ptychadena newtoni	São Tomé	São Tomé	CAS 219249	DQ525933	DQ525946	DQ525955
Ptychadena newtoni	São Tomé	São Tomé	CAS 219250	DQ525934	–	–
Ptychadena newtoni	São Tomé	São Tomé	CAS 219251	DQ525935	–	–
Ptychadena newtoni	São Tomé	São Tomé	CAS 219252	DQ525936	–	–
Ptychadena newtoni	São Tomé	São Tomé	CAS 219253	DQ525937	–	–
Ptychadena newtoni	São Tomé	São Tomé	CAS 219263	DQ525938	–	–
Ptychadena oxyrhynchus	Malawi	no precise locality	359 (6 specimens)	DQ525939	–	–
Ptychadena oxyrhynchus	South Africa	Mtunzini	No voucher	–	DQ525954	DQ525965
Ptychadena oxyrhynchus	South Africa	Kwambonambi	No voucher	AF215403*	DQ525953	DQ525962
Ptychadena aff. *porosissima* A	Tanzania	Tatanda	AC 2034	DQ525940	–	–
Ptychadena porosissima	Tanzania	Mumba	AC 2122	DQ525941	–	DQ525963
Ptychadena cf. *pumilio*	Guinea	Mont Béro	MOR Gu 212	DQ525942	DQ525949	DQ525958
Ptychadena taenioscelis	Kakamega	Kenya	NMKA 3955–1	DQ525943	–	–
Ptychadena aff. *uzungwensis*	Tanzania	Njombe	AC 1970	DQ525945	–	DQ525964
Ptychadena sp.	Tanzania	Mikumi	AC 1976	DQ525944	–	–

AC, Working collection of Alan Channing; CAS, California Academy of Sciences; MVZ, Museum of Vertebrate Zoology, Berkeley, CA, USA; MOR, Mark–Oliver Rödel collection; NMKA, National Museums of Kenya; SL, Stefan Lötters collection; ZSM, Zoologische Staatssammlung München.
*Sequences from Vences *et al.* (2004) used in combined analysis of the three genes. Letters following *Ptychadena* aff *mascareniensis* refer to those allocated in Vences *et al.* (2004).

(Carlquist, 1965), exemplified by two of the other amphibian endemics, *Leptopelis palmatus* (Peters) (see Drewes & Stoelting, 2004) and *Hyperolius thomensis* Bocage (Drewes & Wilkinson, 2004). Interestingly, *P. newtoni* shows strong genetic divergences from those taxa of which it has been considered a synonym in the past: *P. oxyrhynchus* and *P. mascareniensis*. It is remarkable that this species is deeply nested in the *P. mascareniensis* complex (Figs 3 & 4), although Loumont (1992) noted similarities to *Ptychadena anchietae* (Bocage) and *P. oxyrhynchus*. The affinity with East African samples may reflect the paucity of collections from central Africa, specifically the Congo Basin.

This confirmation of the endemicity of *P. newtoni* suggests rejection of the hypothesis that this species was introduced by human settlers. We continue by considering the implication of our results with respect to three competing hypotheses for amphibian overseas dispersal: carriage by birds, storms and rafting.

Carriage by birds

Darwin (1859) proposed carriage by birds as the most plausible explanation for the distribution of land snails on many isolated oceanic islands, and this idea is still cited as the only explanation for their distribution (Gittenberger *et al.*, 2006). Despite speculation (e.g. Fahr, 1993), we know of no record in either ornithological or herpetological literature that has reported finding amphibian eggs associated with the feet, legs or plumage of birds. However, it remains possible that eggs have been transported in this fashion.

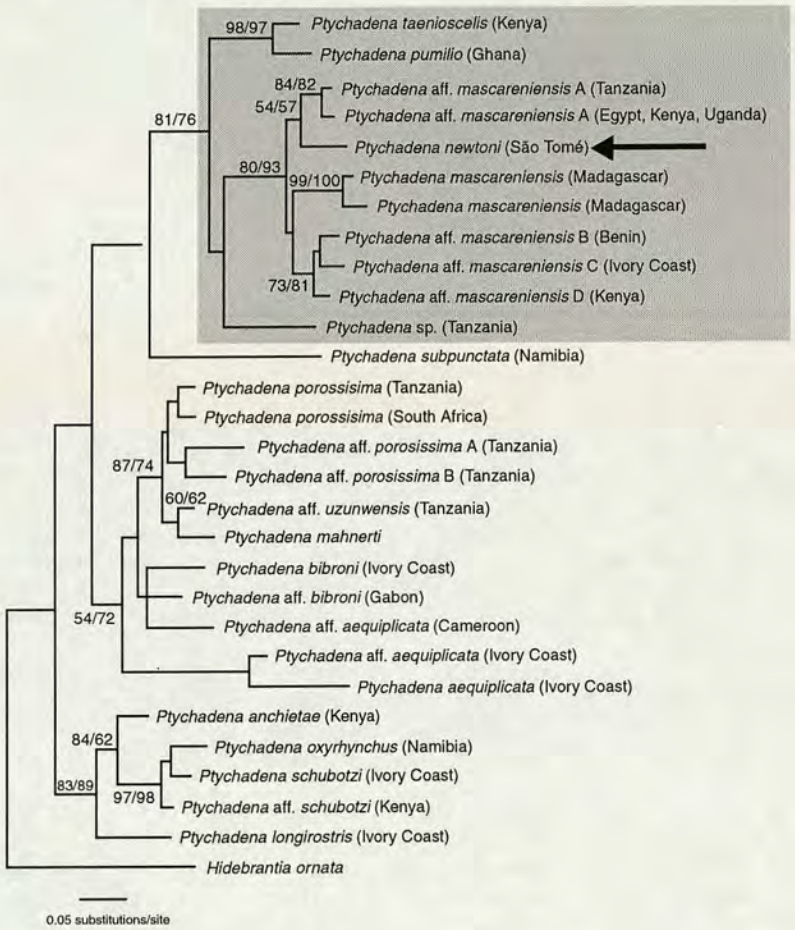


Figure 3 Maximum likelihood phylogram of species in the genus *Ptychadena* based on analysis of 522 base pairs of the mitochondrial 16S rRNA gene. Only one sequence per species or strongly divergent haplotypes were selected from the data set of Vences *et al.* (2004) and complemented with new sequences as listed (Table 2). The endemic species *Ptychadena newtoni* from São Tomé is indicated with an arrow. Numbers at nodes are support values (percentage) from maximum likelihood (100 replicates) and maximum parsimony analyses (2000 replicates), using full heuristic searches with 10 random addition sequence replicates. Values are only shown for nodes where both methods yielded values > 50%.

Although no information is available for the breeding behaviour of *P. newtoni*, other frogs in the genus deposit eggs in rafts on the surface of, and in, shallow water (see Channing, 2001), and so could become attached to the feet or plumage of birds. Frog eggs would have to avoid desiccation while being transported to São Tomé, either by direct flight or through a storm-induced medium (see below). As no east-to-west (African) migration routes are known for candidate bird species, aquatic birds would probably arrive from adjacent mainland Africa, possibly being blown to the islands in a storm. In addition, as remarked by Fahr (1993), carriage by birds does not explain the presence of the caecilian amphibian *S. thomense*, which is known to be viviparous, with juveniles measuring > 100 mm at birth (Haft & Franzen, 1996). It is also unlikely that birds would transport eggs of the tree frog *H. thomensis*, which lays eggs in tree holes on São Tomé (Drewes & Stoelting, 2004). The puddle frogs *Phrynobatrachus dispar* (Peters) and *Phrynobatrachus* sp. *nov.* lay eggs in small accumulations of water, such as those held by fallen leaves in the forest (G.J.M., personal observation), and hence are also unlikely to be dispersed by birds. Even if incidents of frogs' eggs on birds were verified, it is not parsimonious with our phylogenetic results for *P. newtoni*, and we consider it an unlikely mechanism by which any amphibians colonized the Gulf of Guinea islands.

Storms

There are several reports of movement of amphibians by particular weather conditions (see Elsom, 1988; Simons, 1996) when tornadoes suck up a portion of a water body (a waterspout) containing small stages (eggs, larvae and metamorphs) of amphibians. These may be carried a short distance, normally no more than a few tens of kilometres, after which the contents are deposited in what is sometimes referred to as a 'remarkable shower' (Simons, 1996). The more violent the tornado, the greater the distance they are likely to travel, and the less likely the contents (in the case of live amphibians) are to survive, instead being torn to pieces and subjected to low temperatures. Other incidents, where more violent storms have been implicated in colonization events of anolis lizards (e.g. Calsbeek & Smith, 2003), refer to adult individuals reaching islands by being blown into sea water, and are therefore unlikely to explain amphibian colonization. Again, storms are also unlikely to account for a live-bearing and subterranean caecilian. In the unlikely event that storms were responsible for

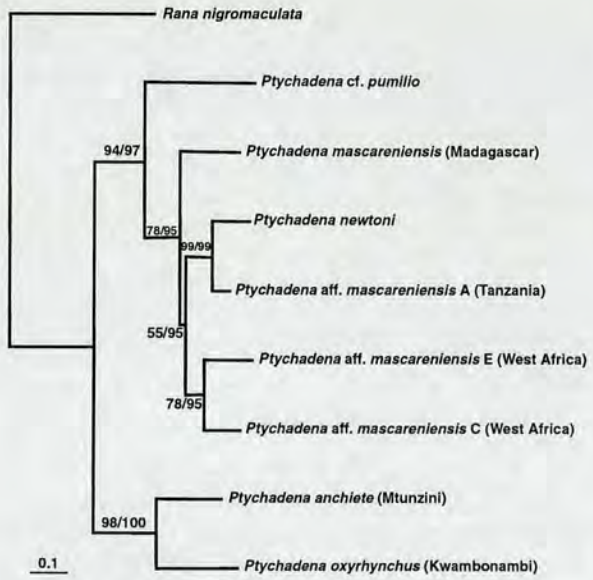


Figure 4 Maximum likelihood phylogram of selected species in the *Ptychadena mascareniensis* complex and *Ptychadena oxyrhynchus*, based on 1648 base pairs of DNA sequences of the 16S rRNA, cyt *b* and coxl genes. Numbers at nodes are bootstrap support values (percentage) from maximum likelihood and maximum parsimony analyses (2000 replicates each).

the movement of amphibians, colonizers would probably be swept in storms from coastal areas to the north or east (see above), a scenario not parsimonious with the results presented here.

Rafting

The rafting hypothesis is congruent for data we have regarding relevant surface currents, occurrence of rafts, and the results of our molecular analysis. Colonizers may have been washed by rainwater into rivers inside the Congo catchment and thence onto rafts in the Congo River. That such conglomerations regularly leave the Congo River is well documented (Renner, 2004 and references therein). Present-day sea currents, wind movements and historical data (see above) all suggest that a raft leaving the mouth of the Congo River may arrive on the shores of the Gulf of Guinea islands. However, this would involve movement of *c.* 1000 km through the sea, during which time high salinity levels would be capable of killing any potential amphibian colonizers. The consequences of high-salinity rafting have been witnessed in the form of thousands of amphibians transported on rafts from the Rio de la Plata [predominantly the caecilian *Chthonerpeton indistinctum* (Reinhardt and Lütken)] washing up dead on the beaches of Uruguay (R. de Sa, personal communication).

A number of amphibian species appear to tolerate elevated salinity by increasing the osmotic concentration of their body fluids to maintain a favourable but relatively small gradient for the osmotic influx of water (Balinsky, 1981). This is apparently possible through their ability to accumulate high levels of urea in body fluids, associated with both urea retention and increased urea synthesis (Shoemaker *et al.*, 1992). While tolerance to 100% seawater may be restricted to very few amphibians (Balinsky, 1981; Schoener & Schoener, 1984), other data show a remarkably high survival for some species in up to 65% seawater (Romspert, 1976), with unpublished data suggesting the possibility of survival in 70% seawater for several months (A. P. Romspert, personal communication). While these data are not available for a wide range of amphibian taxa, they suggest that reduced salinity of surface waters may provide for amphibian passage. Indeed, terrestrial anurans not known for their saline tolerance (*Rana temporaria* L. and *Bufo bufo* L.) are reported to have extensive gene flow by swimming between islands separated by brackish water in the northern Baltic Sea (Seppä & Laurila, 1999). Although there are no explicit data regarding salinity tolerance of *P. newtoni* (or any *Ptychadena*), we consider it unlikely that it can tolerate 100% seawater. Thus the major obstacle to colonization through rafting is the 1000 km or so of sea from the mouth of the Congo River to São Tomé.

Existing data and the simplicity of their distribution, age and number suggest that the Gulf of Guinea islands have the potential to be a model system for studying transoceanic dispersal. We propose a novel mechanism to explain our data, which may be of use in other examples of overseas dispersal by salt-intolerant organisms.

Rafting on freshwater paths

While other workers have mentioned rafting (e.g. Hedges *et al.*, 1992) or a combination of rafting and oceanic currents (e.g. Kaiser *et al.*, 1994), none has highlighted the potential of a combined mechanism with three components: a floating raft, a favourable surface current, and a reduction in salinity of surface waters.

A transporting medium

Floating conglomerations of tree trunks (rotting and living), freshwater aquatic plants, and even soil regularly raft down large rivers in times of spate. In Brazil, floating islands or 'camalotes' have been found to contain numerous terrestrial herpetofauna, including snakes, lizards, frogs and caecilians (Iherring, 1911; Archaval *et al.*, 1979; Schiesari *et al.*, 2003). Rafts have been implicated in colonization events of many herpetofauna, including amphibians (Myers, 1953; Evans *et al.*, 2003). The phenomena of rafts and floating islands have recently been extensively reviewed, demonstrating the long-distance oceanic travel of some conglomerations carrying a diverse array of taxa (VanDuzer, 2004; Thiel & Gutow, 2005). For example, rafts are the likely dispersal mechanism used by many angiosperms that are carried into the tropical Atlantic from the Congo River, and which potentially can traverse the Atlantic in as little as 2 weeks (Renner, 2004).

The presence of rotting logs in a raft is of particular significance, because burrowing taxa such as the caecilian

S. thomense and blind snakes readily inhabit such microhabitats in the forests of São Tomé. Scolecophidian snakes have already been observed on rafting materials (Thiel & Gutow, 2005), and this provides a simple explanation for the presence of burrowing taxa on the Gulf of Guinea islands, a presence that cannot be explained credibly by storms or bird carriage.

Surface currents

There is little doubt that surface currents in the Gulf of Guinea would carry the discharge of two of the world's largest rivers (Congo and Niger), and any raft they may contain, toward Príncipe, São Tomé and Annobón, via the South Equatorial Current (Fig. 1). Although these currents show seasonal changes (Richardson & Walsh, 1986), they continue to demonstrate the same major effect. More important are the effects of large climatic changes that may intensify or reduce the currents (Binet & Marchal, 1993), resulting in long interglacial periods when the currents were favourable to the scenario we propose.

Surface salinity

The data we present (above; Fig. 2) show the strong seasonal changes in surface salinity for a 'normal year', which are directly related to rainfall in the Congo Basin. This relationship further suggests that, in historical periods when rainfall was higher in tropical Africa, freshwater path values as low as 25 psu may have reached the islands. Historical fluctuations in sea surface salinities have been inferred from foraminiferans in sediment cores (Stott *et al.*, 2004). Similar evidence suggests that major discharge pulses of freshwater exited the Congo several times over the past 13 kyr (Marret *et al.*, 2001), and that the Congo plume has extended hundreds of kilometres west from its current position for at least 150 kyr (Dupont *et al.*, 2000). We therefore consider it probable that such freshwater paths have periodically drastically reduced surface salinity as far as the Gulf of Guinea islands.

In summary, our proposed mechanism of dispersal involves conglomerations of vegetation leaving the mouth of the Congo River (the largest of the rivers in the region), and being rafted by surface currents within water of reduced salinity to the Gulf of Guinea islands. The present-day catchment for the Congo River reaches as far east as the Ruwenzori Mountains, potentially bringing rafts from eastern Africa. In the mid-Miocene (15 Ma) the Congo watershed extended further east, with more rivers draining west towards the Atlantic (Goudie, 2005). This was interrupted 12–13 Ma (late Miocene) by the uprising of the present Western Rift Valley, which reduced the Congo catchment and shifted the watershed further west (Goudie, 2005). Hence the results of our molecular analyses of *P. newtoni*, which suggest an East African rather than West African ancestry, could be accounted for by an actual East African nearest relative (from the historical Miocene drainage), or a parent taxon from the poorly sampled current Congo basin. Due to the lack of fossil data for divergences in *Ptychadena* and related taxa, and of other suitable calibration points, our data set does not allow us to calculate divergence times using molecular-clock approaches. However, assuming that rates in this amphibian group concur with the generally observed rates of 0.3–1% pairwise 16S divergence per million years (Vences *et al.*, 2005), this places the divergence of *P. newtoni* from *P.* aff. *mascareniensis* A at 5.6–18.6 Ma; thus the dispersal of *P. newtoni* to the islands prior to the uplift of the Western Rift Valley is plausible.

Other Gulf of Guinea island taxa

These three components not only favour transport of amphibians, but may partly explain the presence of the observed high biodiversity of animals and plants. For example, a raft may have enough invertebrate food to explain the presence of the São Tomé and Príncipe endemic shrews (Thiel & Gutow, 2005). The phylogenetic affinities of skinks suggest that those from São Tomé are most closely related to East African species, while those from Príncipe are most closely related to specimens from Cameroon (Jesus *et al.*, 2005b). Furthermore, the phylogenetic relationships among geckos and skinks suggest multiple independent colonization events of the islands (Jesus *et al.*, 2005a,b) consistent with the arrival of several rafts, as might be expected with the proposed mechanism. Other São Tomé fauna have also been reported to have East African relatives, for example the endemic snail genus *Bocageia* Girard has its nearest relatives in the Comoros and Mt Ruwenzori (Gascoigne, 1994).

The proposed mechanism of dispersal does not rule out rafts originating from other Gulf of Guinea rivers, most notably the Niger, because prevailing oceanic currents would also favour movements towards the Gulf of Guinea islands (Fig. 1). However, the dominance of the Congo River suggests that it might be expected to be the origin for other Gulf of Guinea island amphibians. Drewes & Wilkinson (2004) affirmed the endemic status of *Hyperolius thomensis* and *H. molleri* Bedriaga, finding that they were most closely related to *Hyperolius cinnamomeoventris* Bocage collected in Uganda, although the number of taxa sampled was small. Additional support comes from a recently reported preliminary phylogeny of the genus *Phrynobatrachus* Günther (Zimkus, 2005). This showed that another of the islands' endemic frogs, *P. dispar*, is most closely related to a group of East African species: *Phrynobatrachus parvulus* (Boulenger); *Phrynobatrachus keniensis* Barbour and Loveridge; *Phrynobatrachus* cf. *minutus* (Boulenger); *Phrynobatrachus* cf *rungwensis* (Loveridge); and *Phrynobatrachus rungwensis* (Loveridge). A morphological and molecular study currently under way by J. Uyeda and co-workers (personal communication) indicates that the *Phrynobatrachus* of Príncipe and those of São Tomé are distinct species, and that phylogenetic analysis placed both in a clade with East African species.

The caecilian *S. thomense* has only one congener, *Schistometopum gregorii* (Boulenger), that has a wholly East African

distribution (Nussbaum & Pfrender, 1998). Additionally, Nussbaum & Pfrender (1998) mention a single caecilian specimen housed in the Royal Museum of Central Africa, Tervuren, which is undoubtedly a member of the genus *Schistometopum* (verified by G.J.M.) and which was most probably collected in eastern Congo.

We acknowledge that there are unexplained barriers to this particular mechanism, such as how floating rafts overcame cataracts and waterfalls existing in the present day Congo River, although many possibilities exist over the last 15 Ma, including the drainage of a central African lake (Goudie, 2005). However, the proposed colonization of the Gulf of Guinea islands using rafts on freshwater paths should give pause for thought to proponents of a vicariance-dominated explanation for amphibian distributions. A marked east–west divide of the amphibian fauna of the African Intertropical Montane Region exists whereby less than half the genera are shared (Poynton, 1999). Remarkably, all five Gulf of Guinea island genera (Table 1) are shared between this east–west divide, and those investigated to date appear to have eastern origins. Rafts leaving the Congo may not only be responsible for colonizing the islands in the Gulf of Guinea. We speculate that phylogeographical examination of more taxa from the African Intertropical Montane Region may reveal other ancient east-to-west dispersal events congruent with our proposed mechanism.

We further suggest that this mechanism may have served in colonization events of other oceanic islands. Available data on surface sea currents suggest that rafts from north-western Madagascar would be taken through a vortex to the Comoro Islands, and during the monsoon rainfall in April the surface salinity drops due to output from the Betsiboka River, providing a freshwater path for rafts from Madagascar to the Comoros (data from the Naval Research Laboratory, http://www7320.nrlssc.navy.mil/global_ncom/mad.html). The two frogs from Mayotte (Vences *et al.*, 2003) and many of the Comorian reptiles (Raselimanana & Vences, 2003) show affinities with the lowland fauna of north-western Madagascar.

The Gulf of Guinea islands are particularly suitable as a model region for transoceanic dispersal, and deserve more attention from researchers. More of the endemic flora and fauna of these islands need phylogenetic investigation, which may give better insights into the ancient colonization of remote oceanic islands. In summary, we urge reconsideration of the generally accepted idea that marine environments are always effective barriers to salt-intolerant taxa, because there is now evidence to suggest that they can, at least occasionally, provide freshwater paths for organisms on floating rafts to drift to and colonize remote oceanic islands.

ACKNOWLEDGEMENTS

We would like to thank the following for their help and useful discussion: Dan Barfod, Marcos Di-Bernardo, Angus Gascoigne, Andrew Goudie, Stan Hillman, Tomio Iwamoto, Steve Jebson, Joachim Kosuch, Stefan Lötters, Bob McDowall, Arie van der Meijden, Danny Meirte, Alain Morliere, Eric Opigez, Doug Parker, Rafael de Sá, Paul Simons and Ricka Stoetling. Alan Channing and Mark-Oliver Rödel kindly provided tissues used in the molecular analysis. Y.C. was supported by an Ambassadorial Scholarship of the Rotary Foundation. M.M. was funded by Fundação para a Ciência e a Tecnologia, Portugal. R.C.D. and G.J.M. would like to thank Ned Seligman and Quentino Quade of STePUP for help and assistance in São Tomé. Ricka Stoelting helped collect tissues of *P. newtoni*.

REFERENCES

Andersson, L.G. (1937) Reptiles and Batrachians collected in the Gambia by Gustav Svensson and Birger Rudebeck (Swedish Expedition, 1931). *Arkiv för Zoology*, **29A**, 1–28.

Archaval, F., Gonzalez, J.G., Meneghel, M. & Melgarejo, A.R. (1979) Lista comentada del material recogido en costas Uruguayas, transportado por camalotes desde el Rio Parana. *Acta Zoologica Lilloana*, **35**, 195–200.

Arnault, S. (1987) Tropical Atlantic geostrophic currents and ship drifts. *Journal of Geophysical Research*, **92**, 5076–5088.

Avise, J.C. (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge, MA, USA.

Bakkar, M.I., Bailey, B., Omland, M., Myers, N., Hannah, L., Mittermeier, C.G. & Mittermeier, R.A. (1999) Guinean forests. *Hotspots: Earth's biologically richest and most endangered terrestrial ecoregions* (ed. by R.A. Mittermeier, N. Myers, P.R. Robles Gil and C.G. Mittermeier), pp. 238–253. CEMEX, Mexico.

Balinsky, J.B. (1981) Adaptation of nitrogen metabolism to hyperosmotic environment in Amphibia. *Journal of Experimental Zoology*, **215**, 335–350.

Berry, G.J. & Thorncroft, C. (2005) Case study of an intense African easterly wave. *Monthly Weather Review*, **133**, 752–766.

Biju, S.D. & Bossuyt, F. (2003) New frog family from India reveals an ancient biogeographical link with the Seychelles. *Nature*, **425**, 711–714.

Binet, D. & Marchal, E. (1993) The large marine ecosystem of shelf areas in the Gulf of Guinea: long term variability introduced by climatic changes. *Large marine ecosystems: stocks, mitigation and sustainability* (ed. by K. Sherman, L.M. Alexander and B.D. Gold), pp. 104–118. American Association for the Advancement of Science, Washington DC.

Blaustein, A.R., Wake, D.B. & Sousa, W.P. (1994) Amphibian declines: judging stability, persistence, and susceptibility of populations to local and global extinctions. *Conservation Biology*, **8**, 60–71.

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

Brown, L.H., Urban, E.K. & Newman, K. (1982) *The birds of Africa*, Vol. 1. Academic Press, London.

Burke, K. (2001) Origin of the Cameroon line of volcano-capped swells. *Journal of Geology*, **109**, 349–362.

Caldeira, R. & Munhá, J. (2002) Petrology of ultramafic nodules from São Tomé Island, Cameroon Volcanic Line (oceanic sector). *Journal of African Earth Sciences*, **34,** 231–246.

Calsbeek, R. & Smith, T.B. (2003) Ocean currents mediate evolution in island lizards. *Nature*, **426,** 552–555.

Carlquist, S. (1965) *Island life. A natural history of the islands of the world.* Natural History Press, New York.

Censky, E.J., Hodge, K. & Dudly, J. (1998) Over-water dispersal of lizards due to hurricanes. *Nature*, **395,** 556.

Channing, A. (2001) *Amphibians of central and southern Africa.* Protea, Pretoria.

Chapin, J.P. (1923) The olive ibis of Dubus and its representative on São Thomé. *American Museum Novitates*, **84,** 1–9.

Chiari, Y., Vences, M., Vieites, D.R., Rabemananjara, F., Bora, P., Ramilijaona Ravoahangimalala, O. & Meyer, A. (2004) New evidence for parallel evolution of colour patterns in Malagasy poison frogs (*Mantella*). *Molecular Ecology*, **13,** 3763–3774.

Crother, B.I. & Guyer, C. (1996) Caribbean historical biogeography: was the dispersal-vicariance debate eliminated by an extraterrestrial bolide? *Herpetologica*, **52,** 440–465.

Darwin, C. (1859) *On the origin of species.* John Murray, London.

Dessier, A. & Donguy, J.R. (1994) The sea surface salinity in the tropical Atlantic between 10°S and 30°N – seasonal and interannual variations (1977–1989). *Deep Sea Research*, **41,** 81–100.

Drewes, R.C. (2002) Islands at the center of the world. *California Wild*, **55,** 8–19.

Drewes, R.C. & Stoelting, R.E. (2004) The California Academy of Sciences Gulf of Guinea Expedition (2001) II. Additions and corrections to our knowledge of the endemic amphibians of São Tomé and Príncipe. *Proceedings of the California Academy of Sciences*, **55,** 573–587.

Drewes, R.C. & Wilkinson, J.S. (2004) The California Academy of Sciences Gulf of Guinea expedition (2001) I. The taxonomic status of the genus *Nesionixalus* Perret, 1976 (Anura: Hyperoliidae), treefrogs of São Tomé and Príncipe, with comments on the genus *Hyperolius*. *Proceedings of the California Academy of Sciences*, **55,** 393–405.

Duellman, W. & Trueb, L. (1986) *Biology of amphibians.* McGraw-Hill, New York.

Dupont, L.M., Jahns, S., Marret, F. & Ning, S. (2000) Vegetation change in equatorial West Africa: time-slices for the last 150 ka. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **155,** 95–122.

Dutton, J. & Haft, J. (1996) Distribution, ecology and status of an endemic shrew, *Crocidura thomensis*, from São Tomé. *Oryx*, **30,** 195–201.

Elsom, D. (1988) Catch a falling frog. *New Scientist*, **1615,** 129–131.

Evans, B.J., Brown, R.M., McGuire, J.A., Supriatna, J., Andayani, N., Diesmos, A., Melnick, D.J. & Cannatella, D.C. (2003) Phylogenetics of fanged frogs: testing biogeographical hypotheses at the interface of the Asian and Australian faunal zones. *Systematic Biology*, **52,** 1–29.

Exell, A.W. (1944) *Catalogue of the vascular plants of S. Tomé (with Príncipe and Annobon).* British Museum (Natural History), London.

Fahr, J. (1993) Ein Beitrag zur Biologie der Amphibien der Insel São Tomé (Golf von Guinea) (Amphibia). *Faunistische Abhandlungen Staatliches Museum für Tierkunde Dresden*, **19,** 75–84.

Feiler, A. (1988) Die Säugetiere der Inseln im Golf von Guinea und ihre Beziehungen zur Säugetierfauna des westafrikanischen Festlandes. *Zoologische Abhandlungen Staatliches Museum für Tierkunde Dresden*, **44,** 83–88.

Feiler, A., Haft, J. & Widmann, P. (1993) Beobachtungen und Untersuchungen an Säugetieren der Insel São Tomé (Golf von Guinea) (Mammalia). *Faunistische Abhandlungen Staatliches Museum für Tierkunde Dresden*, **19,** 21–35.

Feller, A.E. & Hedges, B.S. (1998) Molecular evidence for the early history of living amphibians. *Molecular Phylogenetics and Evolution*, **9,** 509–516.

Figueiredo, E. (1994) Diversity and endemism of angiosperms in the Gulf of Guinea islands. *Biodiversity and Conservation*, **3,** 785–793.

Fitton, J.G. (1987) The Cameroon Line, West Africa: a comparison between oceanic and continental alkaline volcanism. *Alkaline igneous rocks* (ed. by J.G. Fitton and B.G.J. Upton), pp. 273–291. Geological Society of London.

Frost, D.R. (2004) *Amphibian species of the World 3.0, an online reference.* American Museum of Natural History, New York, USA. http://research.amnh.org/herpetology/amphibia/index.html.

Gascoigne, A. (1994) The biogeography of land snails in the islands of the Gulf of Guinea. *Biodiversity and Conservation*, **3,** 794–807.

Gittenberger, E., Groenenberg, D.S.J., Kokshoorn, B. & Preece, R.C. (2006) Molecular trails from hitch-hiking snails. *Nature*, **439,** 409.

Goudie, A.S. (2005) The drainage of Africa since the Cretaceous. *Geomorphology*, **67,** 437–456.

Gower, D.J., Kupfer, A., Oommen, O.V., Himstedt, W., Nussbaum, R.A., Loader, S.P., Presswell, B., Müller, H., Krishna, S.B., Boistel, R. & Wilkinson, M. (2002) A molecular phylogeny of ichthyophiid caecilians (Amphibia: Gymnophiona: Ichthyophiidae): out of India or out of South East Asia? *Proceedings of the Royal Society of London Series B, Biological Sciences*, **269,** 1563–1569.

Guibé, J. & Lamotte, M. (1957) Revision systematique des *Ptychadena* (Batraciens Anoures Randies) d'Afrique Occidentale. *Bulletin de l'Institut fondamental d'Afrique noire, Série A*, **19,** 937–1003.

Haft, J. (1993) Ein Beitrag zur Biologie der Echsen der Insel São Tomé (Golf von Guinea), mit näherer Betrachtung der Systematik von *Leptosiaphos africana* (GRAY) (Reptilia: Sauria: Geckonidae et Scincidae). *Faunistische Abhandlungen Staatliches Museum für Tierkunde Dresden*, **19,** 59–70.

Haft, J. & Franzen, M. (1996) Freilandbeobachtungen, Verhalten und Nachzucht der São Tomé-Blindwühle *Schistometopum thomense* (Bocage, 1873). *Herpetofauna*, **18,** 5–11.

Hedges, S.B., Hass, C.A. & Maxsom, L.R. (1992) Caribbean biogeography: molecular evidence for dispersal in West Indian terrestrial vertebrates. *Proceedings of the National Academy of Sciences USA*, **89,** 1909–1913.

Heim de Balsac, H. & Hutterer, R. (1982) Les Soricidae (Mammifères Insectivores) des îles du Golfe de Guinée; faites nouveaux et problèmes biogeographiques. *Bonner Zoologische Beiträge*, **33,** 133–150.

Iherring, R. (1911) Cobras e amphibios das ilhotas de Aguapé. *Revista do Museu Paulista*, **8,** 454–461.

Jesus, J., Brehm, A. & Harris, D.J. (2003) The herpetofauna of Annobón island, Gulf of Guinea, West Africa. *Herpetological Bulletin*, **86,** 375–394.

Jesus, J., Brehm, A. & Harris, D.J. (2005a) Phylogenetic relationships of *Hemidactylus* geckos from the Gulf of Guinea islands: patterns of natural colonizations and anthropogenic introductions estimated from mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics and Evolution*, **34,** 480–485.

Jesus, J., Brehm, A. & Harris, D.J. (2005b) Relationships of scincid lizards (*Mabuya* spp.) from the islands of the Gulf of Guinea based on mtDNA sequence data. *Amphibia–Reptilia*, **26,** 467–473.

Jesus, J., Harris, D.J. & Brehm, A. (2005c) Phylogeography of *Mabuya maculilabris* (Reptilia) from São Tomé Island (Gulf of Guinea) inferred from mtDNA sequences. *Molecular Phylogenetics and Evolution*, **37,** 503–510.

Jones, P.J. (1994) Biodiversity in the Gulf of Guinea: an overview. *Biodiversity and Conservation*, **3,** 772–785.

Jones, P.J. & Tye, A. (2006) *The birds of São Tomé and Príncipe with Annobón: islands of the Gulf of Guinea*. British Ornithologists' Union, Oxford, UK.

Juste, B.J. & Fa, J.E. (1994) Biodiversity and conservation in the Gulf of Guinea islands: faunal composition and origins. *Biodiversity and Conservation*, **3,** 837–850.

Juste, B.J. & Ibañez, C. (1994) Bats of the Gulf of Guinea: faunal composition and origins. *Biodiversity and Conservation*, **3,** 837–850.

Kaiser, H., Sharbel, T.F. & Green, D.M. (1994) Systematics and biogeography of eastern Caribbean *Eleutherodactylus* (Anura: Leptodactylidae): evidence from allozymes. *Amphibia–Reptilia*, **15,** 375–394.

Krause, D.W., Hartman, J.H. & Wells, N.A. (1997) Late Cretaceous vertebrates from Madagascar. Implications for biotic changes in deep time. *Natural change and human impact in Madagascar* (ed. by S.M. Goodman and B.D. Patterson), pp. 3–43. Smithsonian Institution Press, Washington, DC, USA.

Lee, D.-C., Halliday, A.N., Fitton, J.G. & 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.

Loumont, C. (1992) The amphibians of São Tomé and Príncipe: systematic revision, mating calls and karyotypes. *Alytes*, **10,** 37–62.

Lumpkin, R. & Garzoli, S.L. (2005) Near-surface circulation in the Tropical Atlantic Ocean, Part II: time-mean currents and seasonal variability. *Deep Sea Research*, **52,** 495–518.

MacPhee, R.D.E., Singer, R. & Diamond, M. (2000) Late Cenozoic land mammals from Grenada, Lesser Antilles Island-Arc. *American Museum Novitates*, **3302,** 1–20.

Marret, F., Scourse, J.D., Versteegh, G., Jansen, J.H.F. & Schneider, R. (2001) Integrated marine and terrestrial evidence for abrupt Congo River palaeodischarge fluctuations during the last deglaciation. *Journal of Quaternary Science*, **16,** 761–766.

McGlone, M.S. (2005) Goodbye Gondwana. *Journal of Biogeography*, **32,** 739–740.

Mertens, R. (1964) Die Reptilien von Fernando Poo. *Bonner Zoologische Beiträge*, **15,** 211–238.

Mertens, R. (1965) Die Amphibien von Fernando Poo. *Bonner Zoologische Beiträge*, **16,** 14–29.

Myers, G.S. (1953) Ability of amphibians to cross sea barriers, with special reference to Pacific zoogeography. *Proceedings of the Seventh Pacific Science Congress*, pp. 19–26 Whitcombe and Tombs Ltd, Auckland and Christchurch, New Zealand.

Nill, T. (1993) Die Schlangen der Insel São Tomé. *Faunistische Abhandlungen Staatliches Museum für Tierkunde Dresden*, **19,** 71–73.

Nussbaum, R.A. (1984) Amphibians of the Seychelles. *Biogeography and ecology of the Seychelles Islands* (ed. by D.R. Stoddart), pp. 379–415. W. Junk, The Hague.

Nussbaum, R.A. & Pfrender, M.E. (1998) Revision of the African caecilian genus *Schistometopum* Parker (Amphibia: Gymnophiona: Caeciliidae). *Miscellaneous Publications Museum of Zoology University of Michigan*, **I–IV,** 1–32.

Perret, J.-L. (1976) Revision des amphibians africaines et principalement des types, conserves au Musée Bocage de Lisbonne. *Arquivos do Museu Bocage, ser. 2*, **6,** 15–34.

Plana, V., Gascoigne, A., Forrest, L.L., Harris, D. & Pennington, R.T. (2004) Pleistocene and pre-Pleistocene *Begonia* speciation in Africa. *Molecular Phylogenetics and Evolution*, **31,** 449–461.

Posada, D. & Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14,** 817–818.

Poynton, J.C. (1999) Distribution of amphibians in sub-Saharan Africa, Madagascar, and Seychelles. *Patterns of distribution of amphibians: a global perspective* (ed. by W. Duellman), pp. 483–539. Johns Hopkins University Press, Baltimore, MD, USA.

de Queiroz, A. (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology & Evolution*, **20,** 68–73.

Raselimanana, A.P. & Vences, M. (2003) Introduced reptiles and amphibians. *The natural history of Madagascar* (ed. by S.M. Goodman and J.P. Benstead), University of Chicago Press, Chicago, IL, USA.

Renner, S. (2004) Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Science*, **165**(Suppl. 4), S23–S33.

Richardson, P.L. & Walsh, D. (1986) Mapping the climatological seasonal variations of surface currents in the tropical Atlantic using ship drifts. *Journal of Geophysical Research*, **91,** 10537–10550.

Romspert, A.P. (1976) Osmoregulation of the African clawed frog, *Xenopus laevis*, in hypersaline media. *Comparative Biochemistry and Physiology A*, **54,** 207–210.

Savage, J. (1973) The geographic distribution of frogs: patterns and predictions. *Evolutionary biology of the Anurans: contemporary research on major problems* (ed. by J.L. Vial), pp. 351–445. University of Missouri Press, Columbia, MO, USA.

Schiesari, L., Zuanon, J., Azevedo-Ramos, C., Garcia, M., Gordo, M., Messias, M. & Vieira, E.M. (2003) Macrophyte rafts as dispersal vectors for fishes and amphibians in the Lower Solimões River, Central Amazon. *Journal of Tropical Ecology*, **19,** 333–336.

Schoener, A. & Schoener, T.W. (1984) Experiments on dispersal: short-term floatation of insular anoles, with a review of similar abilities in other terrestrial animals. *Oecologia*, **63,** 289–294.

Seppä, P. & Laurila, A. (1999) Genetic structure of island populations of the anurans *Rana temporaria* and *Bufo bufo*. *Heredity*, **82,** 309–317.

Shimodaira, H. (2002) An approximately unbiased test of phylogenetic selection. *Systematic Biology*, **51,** 492–508.

Shimodaira, H. & Hasegawa, M. (1999) Multiple comparison of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*, **16,** 1114–1116.

Shimodaira, H. & Hasegawa, M. (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics*, **17,** 1246–1247.

Shoemaker, V.H., Hillman, S.S., Hillyard, S.D., Jackson, D.C., McClanahan, L.L., Withers, P.C. & Wygoda, M.L. (1992) Exchange of water, ions, and respiratory gases in terrestrial amphibians. *Environmental physiology of the amphibians* (ed. by M.E. Feder and W.W. Burggren), pp. 125–150. University of Chicago Press, Chicago, IL, USA.

Simons, P. (1996) *Weird weather*. Time Warner, London.

Smith, M.A. & Green, D.M. (2005) Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations? *Ecography*, **28,** 110–128.

Stattersfield, A.J., Crosby, M.J., Long, A.J. & Wege, D.C. (1998) *Endemic bird areas of the world: priorities for biodiversity conservation*. Burlington Press, Cambridge, UK.

Stott, L., Cannariato, K., Thunell, R., Haug, G.H., Koutavas, A. & Lund, S. (2004) Decline of surface temperature and salinity in the western tropical Pacific Ocean in the Holocene epoch. *Nature*, **431,** 56–59.

Stramma, L. & Schott, F. (1999) The mean flow field of the tropical Atlantic Ocean. *Deep Sea Research*, **46,** 279–303.

Swofford, D.L. (2002) PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4.0. Sinauer Associates, Sunderland, MA, USA.

Thiel, M. & Gutow, L. (2005) The ecology of rafting in the marine environment. II. The rafting of organisms and community. *Oceanography and Marine Biology: An Annual Review*, **43,** 279–418.

VanDuzer, C. (2004) *Floating islands: a global bibliography (with an edition and translation of G.C. Munz's Exercitatio academica de insulis natantibus)*. Cantor Press, Los Altos Hills, CA, USA.

Vences, M., Vieites, D.R., Glaw, F., Brinkmann, H., Kosuch, J., Veith, M. & Meyer, A. (2003) Multiple overseas dispersal in amphibians. *Proceedings of the Royal Society of London Series B, Biological Sciences*, **270,** 2435–2442.

Vences, M., Kosuch, J., Rödel, M.-O., Lötters, S., Channing, A., Glaw, F. & Böhme, W. (2004) Phylogeography of *Ptychadena mascareniensis* suggests transoceanic dispersal in a widespread Africa–Malagasy frog lineage. *Journal of Biogeography*, **31,** 593–601.

Vences, M., Thomas, M., Bonnett, R.M. & Vieites, D.R. (2005) Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philosophical Transactions of the Royal Society London, Series B*, **360,** 1859–1868.

Wauthy, B. (1983) Introduction à la climatologie du Golfe de Guinée. *Océanographie Tropicale*, **18,** 103–138.

Zimkus, B.M. (2005) Preliminary phylogeny of *Phrynobatrachus* (Anura: Petropedetidae) inferred from mitochondrial 12S and 16S rRNA sequences (abstract). *5th World Congress of Herpetology*, Stellenbosch, South Africa, 154pp.

BIOSKETCHES

John Measey is an amphibian ecologist and has worked throughout Africa, including São Tomé. Current interests include the ecology of subterranean herpetofauna, with a focus on their population biology and dispersal abilities.

Miguel Vences is professor of evolutionary biology and has long been interested in the biogeography of amphibians and other vertebrates of Madagascar and the Comoros islands.

Robert C. Drewes is curator of herpetology at the California Academy of Sciences and has been studying African herpetofauna for over 35 years.

Ylenia Chiari is a postdoctoral researcher interested in conservation genetics of amphibians and reptiles. Over the past 3 years she has worked on conservation of Malagasy amphibians.

Martim Melo is currently assessing the status and phylogeny of the Gulf of Guinea islands' avian taxa, although his interests encompass speciation and island biogeography in general.

Bernard Bourles is a physical oceanographer at Institut de Recherche pour le Développement and has been working on the Equatorial Atlantic circulation for more than 15 years. He currently works in the Gulf of Guinea with the African Monsoon Multidisciplinary Analysis program.

Editor: Robert McDowall

Bird speciation in the Gulf of Guinea, West Africa

Martim Melo University of Edinburgh

The Gulf of Guinea island system is a spectacular centre of endemism mysteriously forgotten by most biologists. It comprises three oceanic islands (Príncipe, São Tomé, and Annobón), one land-bridge island (Bioko) and one ecological island (Mt. Cameroon), all part of the Cameroon line of volcanoes (Fig. 1). Birds are one of the groups for which endemism levels are truly impressive. The three oceanic islands with a total area of about 1000km2 support at least 29 endemic species, with up to five endemic genera. For comparison, the Galapagos archipelago with 13 islands totalising *ca.* 8,000km2 holds 22 endemic species. Each of the oceanic islands is listed by BirdLife International as an independent Endemic Bird Area (EBA), whereas Bioko and Mount Cameroon are part of the Cameroon mountains EBA that, with another 29 endemic species, is the third most important in Africa. São Tomé and Príncipe are in the world top 25% EBAs, a unique feat for small oceanic islands.

Despite these levels of endemism and the presence of some peculiar species – for example, the world's largest sunbird and largest weaver and the smallest ibis – researchers have overlooked this region probably because no spectacular bird radiation occurred here. The moderate distances to mainland allowed colonisation by a wider array of species than is normally the case for very isolated islands. Interspecific competition may have been important from early on, preventing radiations from occurring. This has resulted in endemism being spread across different families rather than being concentrated in species-rich genera, which has the advantage of offering more phylogenetically independent replicates to test evolutionary hypotheses.

Below: The four 'aberrant' white-eyes of the Gulf of Guinea, making the genus *Speirops*: (left) facing the 'typical' white-eyes of the region, from the genus *Zosterops*. A molecular phylogeny showed that aberrant patterns evolved twice in the history of the group, *Speirops* being an invalid genus

An expedition to the region took place from November 2003 to February 2004, with the support of a Genetics Society field grant. This was the second expedition to continue investigations into the processes promoting bird speciation in the Gulf of Guinea. Its objectives were:

1 to complete sampling for phylogenetic and phylogeographic analyses aimed at describing the diversification patterns of different groups (thrushes, sunbirds, white-eyes, seedeaters);

2 to collect phenotypic, genetic and behavioural data at the population level for the Príncipe seedeater (*Serinus rufobrunneus*), the model species chosen to investigate the factors driving the divergence of small allopatric populations.

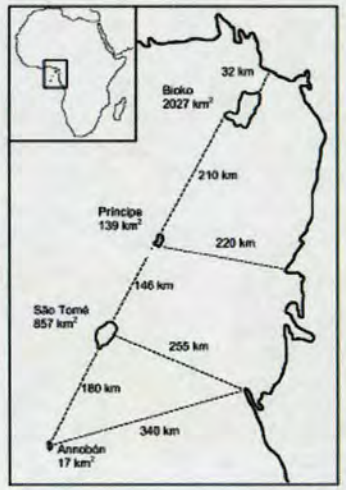
Regional diversification patterns

Taxon sampling for the phylogenetic analyses was successfully carried in the four islands, Mt. Cameroon, and mainland. The highlight was the capture of the extremely rare São Tomé grosbeak (*Neospiza concolor*), a passerine of uncertain affinities last captured 113 years ago, and only observed five times since then. Having handled the bird, the hypothesis that it is a true finch closely related to the Príncipe seedeater with which it co-occurs seems the most plausible.

Direct sequencing of mitochondrial and nuclear DNA sequences is currently being carried out to build genealogies for the different bird groups. The results will allow us to infer the speciation modes, geographic origin of taxa, colonisation routes and timing of colonisation events, allowing the identification of any major regional factor(s) responsible for the high levels of endemism. This will establish the basis for any subsequent testing of process-oriented hypotheses.

Preliminary results are available for the white-eyes (*Zosteropidae*), passerines known for their ability to disperse to islands and speciate once there. The Gulf of Guinea is no exception, with 8 endemic taxa present. In the four instances where two species co-occur (São Tomé, Príncipe, Bioko, Mount Cameroon), one of the species differentiated dramatically from the typical white-eye phenotype. These 'aberrant' white-eyes have been placed in a genus of their own: *Speirops*. Curiously the molecular phylogeny shows that the aberrant white-eyes do not form a monophyletic group, so that *Speirops* must be an invalid genus. The Gulf of Guinea white-eyes are separated in a 'mainland' clade (which includes Bioko) and an 'oceanic island' clade. This suggests that inter-island dispersal was more common than mainland-island dispersal in the history of the group. The two mainland *Speirops* species are basal to the tree, probably being relicts from the last glacial period, whereas the two youngest species are the two oceanic islands *Speirops*. This contrasting pattern suggests that the evolution of the aberrant pattern was not necessarily due to the same causes. It is nevertheless puzzling why such a pattern has evolved only in the Gulf of Guinea.

Fig. 1. The Gulf of Guinea island system



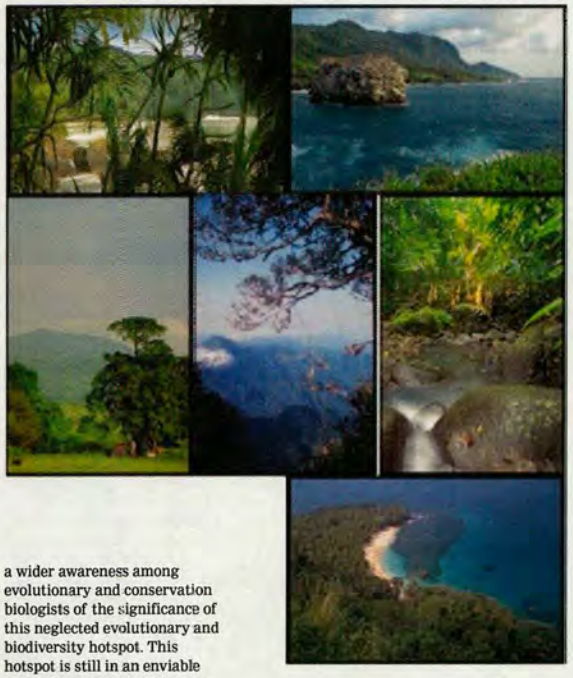
What drives the divergence of allopatric populations?

Príncipe seedeater individuals were sampled on the three islands where they occur and in two different habitats (cocoa plantations and primary forest). Populations in each sampling site will be characterised for a) neutral

genetic diversity, b) adaptive morphometric traits, c) male songs, and d) plumage colour. Divergence in these traits between all populations will be measured in order to determine the relative importance of isolation (drift) versus ecology (selection) in the divergence of allopatric populations. The degree of gene flow between and within islands will be inferred from microsatellite variation data. The target of sampling 20 birds per site was achieved in three sites out of six. In the next field season sampling will be completed and extended to more plantation-forest replicates.

For the analysis of morphological divergence individuals still need to be sexed following a genetic based protocol. The pooled morphological data suggest that isolation is more important in promoting divergence in this species than habitat-type: there are significant between-island but not between-habitat differences. Song repertoire also differed between islands, with no single song type shared between populations. The relevance of this song divergence for the evolution of reproductive isolation will be tested by means of playback experiments.

By gaining a better understanding of the phylogenetic relationships and of the factors promoting divergence of birds in this system we hope to contribute to a wider awareness among evolutionary and conservation biologists of the significance of this neglected evolutionary and biodiversity hotspot. This hotspot is still in an enviable state of conservation, though this may change with the increasing development pressures it faces since the discovery of deep-sea oil reserves.



Príncipe, Annobón
Bioko, São Tomé, Annobón
Príncipe

This research is part of a PhD project primarily funded by the Fundação para a Ciência e a Tecnologia, Portugal. Fieldwork was carried with the support from the following institutions: Annobón and Bioko: Universidad Nacional de Guinea Ecuatorial. Cameroon: Cameroon Biodiversity Project; Limbe Botanic and Zoological Gardens. São Tomé and Príncipe, and Gabon: ECOFAC

INVESTIGATING HYBRIDISATION BETWEEN **ANADROMOUS** AND **FRESHWATER THREE SPINED STICKLEBACKS**

Felicity Jones . University of Edinburgh

Three-spine sticklebacks *(Gasterosteus aculeatus)* have become a model for the study of fish speciation due to the work of Schluter and colleagues on species pairs in four Canadian lakes. Less widely appreciated is the fact that throughout their distribution in estuarine environments divergent 'freshwater' and 'anadromous' morphs of the three-spine stickleback reproduce sympatrically. It is likely that there is incipient speciation in nearly every major river system, giving replication of the process on a very large scale (McKinnon 2002).

Migratory anadromous and freshwater-resident sticklebacks typically differ in morphometric traits such as size, armour plating and spine length, and show differences in mating preferences, life history and allele frequencies. Nevertheless, intermediate morphs are common in most sympatric populations, and their genetic status and fate has not been studied. However, circumstantial evidence suggests that both natural and sexual selection act on hybrid morphs to maintain reproductive isolation (McKinnon 2002). Within-river population processes such as migration, microhabitat preference, natal homing, and population density are likely to have a major effect on hybrid survival and thus the evolution of reproductive isolation.

At the start of 2003 a field grant from the Genetics Society was used to launch a year-long field study of the River Tyne, Dunbar (Scotland), where anadromous and freshwater sticklebacks breed sympatrically. Specific aims of this research included:

- investigating the effect of factors such as population size and the scale within-stream movement on the extent of hybridisation between the two morphs,
- determining the nature of selection on hybrids by investigating temporal changes in morphology and genotypes,
- a QTL association study investigating whether the divergence between marine and freshwater forms involve selection of the same genes in different places.

Photo: © Felicity Jones

The field work involved monthly collection of morphological data from each of eight sites within the river. In addition, over this year-long period approximately 3500 fish were tagged, sampled for genetic analysis, and released back into the river for mark-recapture purposes.

Preliminary results suggest that pre-zygotic isolation is not strong between the morphs and that hybridisation and introgression occurs readily. Sticklebacks in the River Tyne show strong site fidelity with little or no movement between sites. Mark-recapture results indicate little difference between mortality rates of freshwater, hybrid and anadromous young of the year. Interestingly, weirs within the river appear to act as significant barriers to the upstream migration of anadromous sticklebacks having large repercussions on the extent of introgression and gene flow within the river. Further analyses will investigate the maintenance of the two genetically and morphologically distinct stickleback types by looking at the factors affecting hybrid survival.

I would like to take this opportunity to thank the Genetics Society for the Field Work Grant from which many interesting results have emerged. Research carried out in this report has received additional funding from the Royal Society.

McKinnon, J. and Rundle, H. (2002) Speciation in nature: the threespine stickleback model systems. Trends in Ecology and Evolution, 17(10) 480-488.

SÃO TOMÉ E PRÍNCIPE

Explosão de Vida

Um biólogo português em busca de uma espécie quase invisível. Um arquipélago darwiniano, onde a evolução seguiu um caminho especial. Um conjunto de espécies que só foi descoberto há um século. Uma aventura na densa selva são-tomense. E um esforço tremendo de conservação.

Texto e fotografias de Martim de Melo

A cidade de Santo António é a capital e a única povoação do Príncipe. Ostenta a designação honrosa de capital mais pequena do mundo, mas a biodiversidade das suas florestas não tem paralelo. Num país de escasso crescimento demográfico, praticamente sem emigração, o fomento de planos de conservação dependerá da sustentabilidade das actividades humanas.

A rosta sul do Príncipe (à esquerda) e a zona mais selvagem da ilha, onde Martim Melo (em baixo, à esquerda) realizou a maior parte das capturas. Para Bikegila, o guia do biólogo, a praia Sati Inima, junto à foz do rio Poeo (nos mapas portugueses, identificado como ribeira Porco), possibilita boas condições de trabalho: a floresta chega quase até à praia, e aqui é possível realizar as medições e as colheitas de sangue logo da escuridão florestal. Em cima, três tinhosas-pequenas, ícones dos ilhéus Tinhosa, 20km a sul do Príncipe. Os ilhéus em redor de São Tomé e Príncipe contribuem para o mais extraordinário endemismo de aves marinhas registado na região tropical atlântica.

O apoio dos são-tomenses às iniciativas de conservação é essencial para o seu sucesso. Octavio Horta da Veiga (à esquerda) é agricultor e ofereceu-se para ajudar Martim Melo no esforço de captura de aves. A maior parte dos sãotomenses desconhece que as aves que vêem diariamente nas ilhas não existem em mais nenhum lado do mundo. Apesar disso, as iniciativas de educação ambiental sucedem-se e há representações de aves em várias manifestações da cultura de São Tomé (selos, postais, notas).

local, acompanhado por Rita Covas Monteiro, bióloga, como eu, da Universidade de [illegible], e por Freda Pentyney, uma bióloga australiana. Após uma semana de chuva e sem qualquer sinal do anjolo, desmontámos o acampamento e todas as redes utilizadas para apanhar aves. Todas excepto a rede em que eu depositava mais esperanças. Quando comecei a temer que não seria possível achar o [illegible], dirigi-me cabisbaixa para a rede [illegible] um anjolo à minha espera. Um grito de alegria espalha-se pela floresta.

A captura de um espécime foi sem dúvida um momento alto, mas não era essa a única razão para a minha visita às ilhas de São Tomé e Príncipe. Mais do que a raridade, o que torna interessante o anjolo é o facto de ser uma espécie endémica, uma espécie que se formou em São Tomé e não existe em mais lado nenhum do mundo. Ora o anjolo não é uma ave especial. Muito pelo contrário, quase todas as espécies de aves que se podem observar nas florestas de São Tomé e Príncipe são endémicas. No total, estas duas pequenas ilhas do golfo da Guiné têm 28 aves endémicas. É um número extraordinário, considerando que ilhas de dimensões semelhantes costumam ter uma ou duas espécies de aves endémicas. Mesmo o famoso arquipélago das Galápagos, constituído por 13 ilhas com uma superfície oito vezes superior a São Tomé e Príncipe, fica-se pelas 22 espécies endémicas. É a compreensão desta exuberância que me traz aqui.

Durante vários meses, calcorreio as florestas de dois mundos à parte. Em São Tomé vivem o maior beija-flor e o maior tecelão do mundo lado a lado com o íbis mais pequeno do mundo. Um outro tecelão trepa as [illegible] como um pica-pau. O picanço de São Tomé é o único adaptado a florestas fechadas. Para além do anjolo, outras espécies são tão distintas que não se sabe a que família

O papa-moscas de São Tomé (à esquerda) é uma das espécies mais intrigantes do arquipélago. Localmente, chamam-lhe tomé-gagá. Nesta espécie, tanto o macho como a fêmea incubam os ovos. O ninho é "enfeitado" com peças brancas, mas a função dos ornamentos ainda não foi estabelecida. Existem dois beija-flor em São Tomé (em baixo, à esquerda), ambos endémicos. Uma das hipóteses para a origem das duas espécies sugere que elas descendem do mesmo ancestral do continente que terá colonizado a ilha em dois períodos distintos. O macho pequeno e o papagaio cinzento (ambos à direita) são espécies pouco comuns. Um estudo de biólogo em 1998 revelou que este papagaio está geneticamente mais próximo da subespécie que vive na costa norte do golfo do que da subespécie que vive na costa a leste, bastante mais perto.

pertencem. É o caso do *Amaurocichla bocagii*, uma ave tão estranha que houve quem a considerasse descendente de um grupo sul-americano entretanto extinto.

Geologicamente, as ilhas surgiram como rochas nuas no meio do oceano, em resultado de vulcanismo submarino. A diversidade descende das plantas e animais que chegaram acidentalmente a estas rochas arrastados por ventos e marés. Para muitos seres, o mar é uma barreira intransponível e para os que sobreviveram à travessia só existe futuro se outros indivíduos da mesma espécie também aqui tiverem chegado. Uma colonização bem sucedida é por isso rara.

A posição geográfica de São Tomé e Príncipe rodeados a norte e a leste por um dos maiores centros de biodiversidade do mundo, a floresta tropical africana, aumentou as probabilidades de tal evento. Esta será uma das razões para a grande diversidade de espécies que aqui se encontram. A topografia levou à formação de microclimas muito diferentes dentro de cada ilha. Nas florestas do Sul de São Tomé, a precipitação anual é de 7.000mm enquanto no Norte pode não passar dos 600mm. A diversidade topográfica e climática também justifica a riqueza biológica de São Tomé e Príncipe.

Ao contrário do que sucede na maioria das ilhas oceânicas, este património biológico ainda é [illegible]. Cerca de 90% do país está sob coberto florestal, dividido entre floresta primária, floresta secundária e plantações de cacau e café. Todas as espécies endémicas encontram-se na floresta primária mas a maior parte utiliza as florestas secundárias e muitas adaptaram-se às plantações.

A floresta primária constitui de longe a jóia das ilhas. A floresta do Sul de São Tomé foi considerada pela organização BirdLife International a segunda floresta mais importante para a conservação de aves de todo o continente africano, incluindo Madagáscar. Sobreviveu até hoje protegida pela chuva e o acesso difícil. No entanto, a conservação das florestas não pode depender apenas dos elementos. São Tomé e Príncipe é um país pequeno e com um crescimento populacional de 3%. Sem emigração é fácil compreender a dificuldade em estabelecer planos de desenvolvimento que conciliem as necessidades humanas e do meio ambiente. O governo são-tomense tem demonstrado uma visão exemplar deste problema. Acolheu na Direcção das Florestas o programa de conservação e utilização racional das florestas ECOFAC financiado pela União Europeia. Elaborou uma lei de bases do ambiente, um plano nacional para o ambiente e desenvolvimento sustentável e completa neste momento o plano nacional da biodiversidade. Finalmente, a lei das áreas protegidas foi recentemente aprovada por unanimidade pela Assembleia Nacional. Cerca de 1/3 da superfície de cada ilha foi definida como área protegida.

O maior desafio vai ser a aplicação destas políticas cuja urgência pode ser ofuscada pela recente descoberta de petróleo. Se por um lado o petróleo pode trazer os rendimentos necessários para a aplicação de planos de desenvolvimento progressistas, por outro pode levar a uma febre de desenvolvimento rápido que esquece o ambiente e aumenta a desigualdade social. Os próximos anos serão decisivos. ◻

São Tomé et Príncipe : Obo

L'archipel de São Tomé et Príncipe abrite un nombre impressionnant d'espèces d'oiseaux endémiques. Dans les deux îles, qui totalisent moins de mille kilomètres carrés, on trouve 28 endémiques. À son tour le beaucoup plus connu archipel des Galápagos abrite 22 oiseaux endémiques dans treize îles ayant une superficie de huit mille kilomètres carrés. Un des oiseaux endémiques a particulièrement fasciné les ornithologues, atteignant durant le XX^{e} siècle une aura mythique. C'est le néospize de São Tomé *Neospiza concolor*, actuellement considéré comme un des oiseaux les plus rares du monde. En 1888, le naturaliste portugais Francisco Newton explore les forêts de São Tomé jusqu'aux coins les plus reculés. Parmi les spécimens qu'il collecte se trouve un passereau brun-roux, de la taille d'une grive, avec un bec énorme rappelant celui des perroquets. En 1890, il capture deux nouveaux spécimens. Deux de ces spécimens seront détruits dans l'incendie du muséum d'histoire naturelle de Lisbonne en 1978. Le troisième se trouve au muséum d'histoire naturelle britannique. C'est l'unique spécimen connu du néospize (également appelé gros-bec) de São Tomé. Cet oiseau va confondre les taxinomistes. Il est initialement décrit comme un tisserin du genre *Amblyospiza*, mais il est ultérieurement considéré comme un fringille, probablement proche du genre *Serinus*. On le place aujourd'hui dans un genre à lui seul : *Neospiza*.

Une procession espacée mais régulière d'ornithologues visite São Tomé à la suite de Newton. Le gros-bec, lui, disparaît. En 1990, l'expédition de l'université de East Anglia fait histoire en redécouvrant à São Tomé quatre espèces qui n'avaient pas été observées depuis soixante ans, mais pas le gros-bec. Le grand jour arrive en août 1991 quand le gros-bec est redécouvert dans les forêts du sud-ouest, près de la rivière Xufe-xufe, qui deviendra une référence pour ornithologues et *birdwatchers* motivés. Pendant les dix années suivantes, c'est à nouveau le silence, à peine interrompu par deux observations non confirmées. L'observation de 1991 a bien démontré l'existence de l'espèce, mais sa rareté demeure surprenante. L'espèce est-elle vraiment rare ou son étrange bec est-il une adaptation à un fruit particulier qui se trouve hors de vue dans les niveaux supérieurs de la canopée ?

Entre décembre 2001 et février 2002, une équipe de l'université d'Edinburgh part à São Tomé pour essayer de déterminer les densités des oiseaux endémiques dépendants de la forêt primaire et accroître la connaissance de leurs nécessités écologiques. Rapporter de nouvelles données sur le néospize est donc un objectif prioritaire. Conduits par l'incontournable guide santoméen Pedro Leitão, ils explorent les forêts de la région de la rivière São Miguel, au nord du Xufe-xufe. Leurs efforts sont couronnés par plusieurs observations qui offrent pour la première fois des données sur son comportement alimentaire. Les oiseaux ont toujours été observés se nourrissant sur des arbres en fruits : *Uapaca guineensis* et *Dicranolepis thomensis*. Particulièrement intéressant est le fait que les oiseaux ont été observés à faible hauteur, entre 3 et 15 mètres.

En octobre 2002, je commence à São Tomé la saison de terrain pour mon doctorat sur la spéciation de plusieurs groupes d'oiseaux endémiques du golfe de Guinée. Les fringilles sont un de mes modèles et il me faut échantillonner le néospize. En décembre 2002,

suivant les conseils de Pedro Leitão et Lúcio Primo, j'explore les forêts du sud-est, plus accessibles mais moins visitées du fait du magnétisme du Xufe-xufe. Un gros-bec est à nouveau observé. Un individu se déplace dans l'arbre endémique *Dicranolepis thomensis* qui est en fleurs. Les efforts de capture au filet échouent. On y retourne un mois plus tard lorsque *Dicranolepis thomensis* est en fruits. Au bout d'une semaine sans trace de gros-bec, tout le matériel est rangé pour le départ. Tout, excepté le filet qui, 114 ans après, capture à nouveau un gros-bec. Avec les photographies et les mesures morphométriques, des échantillons de sang ont été pris pour les analyses génétiques. Celles-ci permettront finalement de connaître les origines de cette espèce si énigmatique. L'ayant observé en main, il me paraît que l'espèce la plus proche du gros-bec se trouve à São Tomé : le serin roux *Serinus rufobrunneus*, endémique de São Tomé et Príncipe. Si c'est le cas, le mythique oiseau devra abandonner son poste de genre unique pour joindre le genre *Serinus*. Après, il faudra voir quel est le serin du continent le plus proche de ceux de São Tomé. Comme Amadon le suggérait déjà en 1953, *Serinus burtoni* est l'hypothèse la plus probable.

Après cette redécouverte, les forêts du sud-est sont le nouveau lieu de pélerinage d'un nombre croissant de *birdwatchers*. Nik Borrow y a enregistré son chant. On espère qu'à partir de maintenant les observations de cet oiseau rarissime seront un peu plus fréquentes et que petit à petit on en apprendra plus sur ses nécessités écologiques.

Aujourd'hui, ce qu'il faut retenir est que :

– le néospize est bel et bien vivant mais ses effectifs doivent être très faibles vu qu'on l'observe dans des arbres de petite taille et qu'il n'est pas timide – il ne devrait pas passer inaperçu s'il était plus nombreux;

– la forêt du sud-est de São Tomé héberge tous les oiseaux endémiques, comme c'est le cas de la plus fameuse forêt du sud-ouest. Cela accroît significativement l'aire d'occurrence des endémiques dépendants de la forêt primaire. En même temps cela rappelle la nécessité d'assurer la conservation de la totalité des forêts du sud. Ceci est particulièrement important dans le cas des forêts du sud-est qui sont beaucoup plus accessibles que celles de l'ouest (le néospize a été observé à une heure et demie de marche de la présence humaine). La très récente approbation unanime par le parlement santoméen de la loi sur les parcs nationaux apporte un grand espoir pour la conservation du néospize et de toutes les autres espèces d'oiseaux endémiques de São Tomé et Príncipe, ainsi que des écosystèmes uniques où elles habitent. Cette loi adopte la délimitation des aires protégées proposée par Écofac-São Tomé qui, à São Tomé, se traduit par la protection de plus ou moins un tiers de l'île !

Martim MELO
Institute of Cell, Animal and Population Biology
University of Edinburgh
King's Buildings
Edinburgh EH9 3JT,
Royaume-Uni

Bibliographie :

Amadon, D. 1953. *Avian systematics and evolution in the Gulf of Guinea.* Bull. Am. Mus. Nat. Hist. 100 : 393–452.

Atkinson, P., Peet, N. & Alexander, J. 1991. *The status and conservation of the endemic bird species of São Tomé and Príncipe, West Africa.* Bird Conserv. Internat. 1 : 255–282.

Dallimer, M. & King, T. 2003.*New records of the São Tomé Grosbeak Neospiza concolor.* Bull. African Bird Club 10 : 23-25.

King, T. & Dallimer, M. 2003. *Daily activity, moult and morphometrics of the birds of São Tomé and Príncipe.* Bull. African Bird Club 10 : 84-93

Sargeant, D.E., Gullick, T., Turner, D.A. & Sinclair, J.C.I. 1992. *The rediscovery of the São Tomé Grosbeak Neospiza concolor in south-western São Tomé.* Bird Conserv. Internat. 2 : 157–159.

(photographies de l'auteur)