Molecular Systematics and Diversification of African Zosteropidae (Aves: Passeriformes)

Siobhan Claire Cox

A thesis submitted for the degree of Doctor of Philosophy

Date (Jan 2013)

University College London

Department of Genetics, Evolution and Environment

Declaration

I, Siobhan Claire Cox, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in this thesis.

Siobhan Claire Cox
Candidate

Statement of authorship

Obtaining samples: The DNA samples used in this thesis were obtained through fieldwork and collaborations with various institutes. I set up all links with institutes (six in total) that donated blood samples and DNA extracts. In organising and carrying out fieldwork to collect blood samples in east Africa, I also set up the collaboration with the National Museum of Kenya (NMK). With the help of Dr. Robert Prys-Jones I set up the collaboration with the Museum of Natural History, Stockholm (NRM) who funded and provided training for working with DNA extracted from museum material. While I collected tissue samples from museum specimens in the NHM collections, Martin Irestedt (NRM) sampled specimens from the NRM collections.

Data generation: For the sequence data that was generated for Chapter 2, I conducted the primer design, primer testing and generated the majority of the sequence data (some extractions and PCRs were carried out with the assistance of a lab technician, Mari-Wyn Burley). For Chapter 3, I generated all AFLP profiles. This included testing methodologies, designing and testing suitable primers and conducting dye ratio tests. Sequences and AFLP reads were run by the sequencing facility at UCL. For Chapter 4, all molecular work on museum specimens was performed under the supervision of Martin Irestedt at the NRM. My work there included designing and testing new primers, extraction of DNA and all PCR and clean-up reactions. Sequences were run by the sequencing facility at the NMR.

Data analysis: I assembled and aligned all sequence data and scored all AFLP profiles. I conducted all analyses including the phylogenetic and dating analyses of Chapter 2, the phylogenetic and hypothesis testing analysis of Chapter 3 and the phylogenetic and species delimitation analysis of Chapter 4.

Abstract

The hyper-diverse avian family Zosteropidae (~100 species) is renowned for its exceptional colonising ability and rapid recent diversification. The genus *Zosterops* dominates within the family, including over 70% of all recognised species, and is notorious for the phenotypic uniformity that has traditionally made designation of taxon boundaries within the genus difficult. While the last decade has seen an abundance of research focusing on colonisation abilities and speciation patterns of insular taxa, relationships between continental forms, specifically mainland African taxa, remain subject to great uncertainty.

This thesis focuses on uncovering the relationships, origin and evolutionary history of African Zosteropidae. Chapter 1 introduces the family Zosteropidae, reviews the current literature that is based predominantly on insular systems. In introducing the African Zosteropidae complex, this chapter highlights questions associated with this group and presents the aims of the thesis.

Chapter 2 focuses in on one of the most geographically complex areas within the African system, to explore the relative importance of past climatic fluctuations as a driver of diversification in *Zosterops* endemic to the isolated montane massifs of East Africa. Results provide the first molecular assessment of mainland African Zosteropidae and are used to examine alternative models of speciation. A dated molecular phylogeny demonstrates that divergence within African Zosteropidae is very recent (<5Ma) coinciding with periods of climatic instability during the Plio-Pleistocene. Furthermore, the non-monophyly of mainland taxa, specifically the polyphyletic nature of *Z. poliogaster*, leads to the rejection of a widely held assumption that the montane endemics of East Africa are relics of a previously widespread population. Instead results provide evidence for evolutionary model based on ancestrally adaptive populations.

Chapter 3 attempts to further investigate relationships within the East African *Zosterops* and explores the usefulness of amplified fragment length polymorphism (AFLPs) in revealing inter- and intra-specific relationships. Resolution of relationships across the AFLP phylogeny is generally poor which is attributed to the low information content of the AFLP matrix

generated. Bayesian hypothesis testing failed to provide support for various topological constraints tested and consequently this study was unable to confirm or reject the non-monophyly of East African montane endemics.

Chapter 4 builds upon the molecular phylogeny of Chapter 2, by substantially increasing the taxonomic sampling of African species using DNA obtained from museum specimens. The use of both archive and fresh material enabled the largest genetic assessment of western Zosteropidae to date. Extensive sampling across Sub-Saharan Africa, the Indian Ocean region, the Arabian Peninsula and the Gulf of Guinea region reveals six major clades within the African Zosteropidae complex. Results confirm the widespread non-monophyly of mainland African species rendering current taxonomic arrangements invalid. GMYC (General mixed Yule-coalescent) analysis recovers 14 distinct evolutionary lineages within the African Zosteropidae system and provides a framework for further work using model-based species delimitation approaches.

Finally, Chapter 5 draws together key findings from Chapters 2-4, and reviews how this work advances our understanding of the African Zosteropidae system. This chapter also highlights new gaps in our understanding of the western Zosteropidae and discusses several areas for future research.

Table of Contents

Chapter 2: The role of Plio-Pleistocene climatic fluctuations in the diversification of East African *Zosterops*

Chapter 3: The phylogenetic utility of Amplified Fragment Length Polymorphisms (AFLPs) in resolving relationships within East African *Zosterops*

Chapter 4: Molecular phylogeny and species status of continental African *Zosteropidae* using museum skins

List of figures

List of tables

Final word count: 47,815

Acknowledgements

Firstly, I would like to thank my supervisors Dr. Julia Day from University College London (UCL) and Dr. Robert Prys-Jones from the Natural History Museum (NHM, Tring) for giving me the opportunity to work with such as fascinating group of birds. The support, guidance and encouragement you have provided over the last four years will never be forgotten.

This thesis would not have been possible without a number of collaborators, who have kindly provided blood samples and DNA extracts from their extensive collecting trips: Dr Jan Habel (Universität Trier), Dr Ben Warren (University of Reunion), Dr Martim Melo (CIBIO) and Luca Borghesio (University of Illinois). In addition, my own field collection trip would not have been possible without the help of Phillista Malaki, Bernard Amakobe and Nicodemous Nalianya from the National Museums of Kenya (NMK) whose assistance both in and out of the field has been invaluable. A special mention should go to Nigel Cleere (NHM, Tring) who assisted on fieldwork. Your knowledge and tireless enthusiasm for birding made the trip all the more enjoyable.

I would like to show my gratitude to the various lab managers and technicians I have worked with during the course of my PhD. Their help and troubleshooting advice was welcomed on numerous occasions. Specifically, I would like to thank Steve Russell, Andy Hall and Alex Aitken who provided lab support at the NHM (London) and to Julia Llewellyn-Hughes and Claire Griffin who provide support at the Molecular Biology Sequencing Unit (NHM, London). In addition to Mary-Wyn Burley, who provided assistance and support in the molecular labs at UCL. A massive thank you should go to the lab group at the Museum of Natural History Stockholm (NRM), particularly Dr Per Erikson, Dr Martin Irestedt, and Dr Pia Eldenäs. Their expertise, guidance and financial support enabled me to overcome many of the problems I ran into when working with DNA from museum material.

I would like to offer my special thanks to Eileen Cox and Lukas Rüber who provided me with support, advice and a place to work (NHM, London). Their willingness to give their time so generously has been very much appreciated. I would also like to thank all colleagues at the NHM, Tring, specifically Mark Adams (Natural History Museum, Tring) for curatorial support.

I would like to express my very great appreciation to Dr David Pearson for his valuable and constructive suggestions during the planning and development of this work. In addition to Dr Jeff Powell (Freie Universität Berlin) and fellow PhD student Chris Barton (NHM, London) provided me with very valuable help and direction with regards to running species delimitation analysis and to Simon Ho (University of Sydney) for his advice on working with BEAST.

For funding, I would like to thank NERC, the Systematics Research Fund (SRF), the UCL Graduate School and the NHM. This work would not have been possible without the financial support given.

A massive thank you needs to go to the many colleagues past and present at both UCL and the NHM, specifically: Kai Winkelmann; Cameron Richards; Jessica Bryant; Katherine Brown; Claire Peart; Lisa de Silva; Antonia Ford; Simon Maddock; Jake Morris and Ripu Bains, your support and words of encouragement have helped supported me throughout. To my friends both in London and in Lancashire, you have helped keep me sane and made sure I had a life outside of work. The nights out, evenings in and weekends away brought me back to reality when I need it most.

To my family, particularly my parents David and Bernadette and my sisters Jennie, Nicola and Jodie, although you might not have always understood what I was doing, your love, support and encouragement have enabled me to follow my dreams. Finally, to my partner Tom who has been my rock through the up's and down's of the last three years. Your endless love, patience and sacrifice have made this work possible. Thank you for believing in me.

'Tis a lesson you should heed, try, try again.

If at first you don't succeed, try, try again.'

Edward Hickson

Dedication: I would like to dedicate this thesis to my parents David and Bernadette Cox. Thank you for your unconditional support with my studies. From the woodlouse in ‘Witch Wood’ to completing my PhD you have always nurtured my love of learning and taught me the importance of self-belief. Thank you for giving me the chance to prove and improve myself through supporting the choices I have made. I love you.

Chapter 1

Introducing the family Zosteropidae (white-eyes)

Table of Contents

1.1. Zosteropidae (white-eyes)

1.1.1. General characteristics of the family Zosteropidae

Zosteropidae are a diverse old world passerine family made up of small, gregarious, arboreal birds that have a broad distribution occupying tropical, subtropical and temperate Sub-Saharan Africa, southern and eastern Asia, Australasia and the tropical islands of the Indian Ocean, the western Pacific Ocean, and the Gulf of Guinea region (van Balen 2008) (Fig. 1.1). Morphological variation across the range of the family is slight, leading most members to be homogenous in appearance. This group exhibits remarkable uniformity in their structure and plumage colouration; which is generally greenish-olive above and pale grey below. There is a general trend for continental species to be more yellow/green, while insular taxa are more grey/brown (Fry *et al.* 2000).

As their common name implies, many species have a conspicuous ring of tiny white feathers around the eyes. The breadth of this eye-ring varies between species, being highly exaggerated in some taxa and reduced or absent in others. Some species have a white or bright yellow throat, breast or lower parts, and several have buff flanks (van Balen 2008). All species are highly sociable and form large flocks that separate on the approach of the breeding season (Moreau 1957). Members of this family are highly vocal, but tend to have weak rather simple vocalizations that are far carrying (Fry *et al.* 2000). While mainly insectivorous, they have a generalist diet eating nectar and fruits of various kinds (Moreau 1957).

1.1.2. Systematics

The presence of a brush-tipped tongue has previously placed the family next to nectarivorous groups of Nectariniidae (sunbirds) and Meliphagidae (honeyeaters), although they have also been treated as a sub-family of Promeropidae (sugarbirds) (van Balen 2008). Molecular data has placed Zosteropidae in a Sylvioid lineage (Cibois 2003) and consequently they are now placed in the super-family Sylvioidae, between Sylviidae (old world warblers) and Cisticolidae (cisticolas) (van Balen 2008). More recent molecular studies indicate a close relationship with Timaliidae (babblers), and place Zosteropidae

in a clade with *Yuhina* and *Stachyris,* both of which belong to the family Timaliidae (Cibois 2003; Moyle *et al.*2009).

Zosteropidae are highly diverse, current estimates place 98 species in 14 genera, of which 74 species belong to the genus *Zosterops.* The remarkably homogeneous genus *Zosterops* occupies the entire range of the family (Fig. 1.1) compared to other genera (*Woodfordia, Rukia, Cleptornis, Apalopteron, Tephrozosterops, Madanga, Lophozosterops, Oculocincta, Heleia, Chlorocharis, Megazosterops, Speirops*) that have much smaller distributions (van Balen 2008). Other genera are thought to be derived from 'typical' *Zosterops*, and have been described alongside (van Balen 2008).

1.1.3. Taxonomic complexities

The family Zosteropidae has long posed problems for traditional taxonomists. Recovering relationships using traditional approaches is notoriously difficult, particularly at the species level where the abundance of morphologically similar forms has complicated efforts to identify natural groupings. Despite extensive analyses by Moreau (1957) for western Zosteropidae, and Mayr (1965) and Mees (1961; 1969; 1953) for eastern Zosteropidae, the affinities of numerous taxa still remain unresolved. In some of the more phenotypically divergent groups, recent molecular studies have highlighted a large discordance between relationships obtained from morphological and molecular characters (Melo *et al.* 2011).

Broad molecular investigations have revealed that many of the 'aberrant white-eyes', currently classified in distinct genera, nest well within the genus *Zosterops*: *Speirops* (Melo *et al.* 2011), *Rukia* (Slikas *et al.* 2000), *Woodfordia* and *Chlorocharis* (Moyle *et al.* 2009). Within the Gulf of Guinea, the genus *Speirops* is recovered as non-monophyletic, with each 'aberrant' species being more closely related to 'typical' *Zosterops* than they are to each other (Melo *et al.* 2011). These molecular insights are bringing into question the utility of morphological characters in Zosteropidae (van Balen 2008), with results suggesting that phenotypic characters are evolving in a non-neutral fashion.

Figure 1.1. Global distribution of the avian family Zosteropidae Image modified from Warren *et al.* (2006). Distribution of the genus *Zosterops* indicated by the thick dark line. Distribution of other genera within indicated by numbered red areas

While molecular studies are starting to tease apart relationships, to date much of the work has been focused on a few oceanic island systems (Slikas *et al.* 2000; Warren *et al.* 2006; Phillimore *et al.* 2008; Moyle *et al.* 2009; Melo *et al.* 2011) with little investigation into continental relationships (Oatley *et al.* 2012). Much of the current taxonomy therefore (particular in the genus *Zosterops*) is based solely on morphology and ecology, which in some cases is supported by facts regarding the general behaviour and vocalisations of individual populations (van Balen 2008).

1.1.4. Colonisation abilities

A renowned feature of the family Zosteropidae is its ability to colonise islands and then speciate there (Slikas *et al.* 2000; Warren *et al.* 2006; Phillimore *et al.* 2008; Moyle *et al.* 2009; Melo *et al.* 2011). Within the family, the genus *Zosterops* appears to have among the highest dispersal capabilities of birds, with an exceptionally wide distribution owing to its high colonising potential (Moyle *et al.* 2009). Levels of diversity and endemism peak in the Australian and Oriental regions, where oceanic island concentrations are highest (Dickinson 2003; van Balen 2008). With 46% of the worlds Zosteropidae being single-island endemics, it is unsurprising that this highly speciose family has stimulated interest into the relative contribution of long distance immigration and local in-situ speciation to the diversity of island systems (Warren *et al.* 2006).

Investigation into the origin and diversification of Indian Ocean *Zosterops* (Warren *et al.* 2006) has revealed that much of the regions diversity is a result of long-distance immigration, rather than regional in-situ processes. This is particularly evident for two sympatric species occurring on the islands of Mauritius and La Réunion (Mascarenes). The non-monophyletic placement of taxa is consistent with double island colonisation rather than with-in island speciation. This relationship is also observed in the Grande Comore and Granitic Seychelles which are, or have previously been, occupied by two species that are recovered in different clades supporting a multiple colonisation model (Warren *et al.* 2006).

The predominance of long-distance dispersal as a primary driver of diversity within the family has also been documented in Melanesia (Mayr and Diamond 2001; Phillimore *et al.* 2008). The molecular phylogeny of Phillimore *et al.* (2008) revealed at least two independent colonisations of the Vanuatu archipelago and reports long periods of isolation between island populations that is consistent with very little gene flow between islands. In many island systems there is a general lack of geographical overlap between closely related species. In the majority of cases, molecular investigation has shown that speciation of insular taxa is a consequence of geographical isolation (van Balen 2008). While there are several examples where two (or more rarely three) species co-occur on the same island, the co-existence of taxa has repeatedly been attributed to multiple colonisations from mainland areas (Warren *et al.* 2006; Phillimore *et al.* 2008). In cases where islands are occupied by multiple taxa, species are generally distantly related and occupy different elevation and/or habitat distributions (Warren *et al.* 2006; Phillimore *et al.* 2008).

1.1.5. Evidence of an adaptive radiation

Although adaptive radiations of oceanic island birds have played a central role in the development of speciation theory (e.g. Darwin's finches: Grant and Grant 2008, Hawaiian Honeycreepers: Pratt 2005; Learner *et al.* 2011), adaptive radiations in birds are in fact rare events (Ricklefs and Bermingham 2007; Price 2008). However, recent molecular work investigating the relationships and colonisation sequence of Gulf of Guinea Zosteropidae (Melo *et al.* 2011) has identified two radiations whose tempo and patterns of morphological divergence are strongly supportive of an adaptive radiation rivalling those of Darwin's finches and the Hawaiian honeycreepers. Species occupying the Gulf of Guinea region currently fall into two genera ('typical' *Zosterops* and 'aberrant' *Speirops*) and exhibit a breadth of phenotypic diversity that is unmatched across the family worldwide (Moreau 1957).

The build-up of phenotypically differentiated island endemics in the Gulf of Guinea region were, for some time, attributed to multiple independent colonisations from mainland Africa. Contradicting previous hypotheses (Jones and Tye 2006: and references therein), the molecular phylogeny of Melo *et al.* (2011) places the Gulf of Guinea white-eyes in just two radiations. Furthermore, the two phenotypic groups ('typical' *Zosterops* and 'aberrant' *Speirops*) were not

recovered as independent clades, rejecting previous predictions that the two phenotypic groupings are derived from separate colonisation events (Melo *et al.* 2011). In contrast to many island archipelagos, species diversity within the Gulf of Guinea regions fits the archipelago radiation model rather than multiple colonisation models. In this system, rapid phenotypic divergence is consistent with the model of asymmetric divergence owing to resource competition in sympatry (Melo *et al.* 2011). Investigation into two other congeneric species (Vanuatu archipelago) found much lower levels of diversification with no evidence of an adaptive radiation (Clegg and Phillimore 2010). In assessing the relative role of gene flow between allopatric populations, Clegg and Phillimore (2010) found no evidence that inter-island gene flow constrains phenotypic divergence. This may lead to greater emphasis on the role of ecological divergence and diversifying selection pressures, rather than geographic context, in driving population divergence within Zosteropidae.

1.1.6. Zosteropidae: an example of a 'Great Speciator'

The recent molecular work of Moyle *et al.* (2009) characterises the family Zosteropidae as a 'Great Speciator'. Divergence time estimates revealed that the majority of divergence events within Zosteropidae have occurred within the last 2 million years, yielding diversification rate estimates of 1.93-2.63 species per million years (Moyle *et al.* 2009). This exceptionally high diversification rate is supported by a previous analysis of an Indian Ocean-centred phylogeny that recovered a substitution rate estimate of 4.66% per million years (Warren *et al.* 2006). This is significantly faster than the 2% substitution rate that was found across many bird groups over longer timescales (Weir and Schluter 2008).

Like other groups that exhibit high degrees of differentiation across broad spatial scales, the species-rich family Zosteropidae presents a paradox: while the exceptional colonisation abilities of Zosteropidae may generate more geographical opportunities for speciation, in theory they should limit differentiation by reducing the impact of barriers to gene flow (Moyle *et al.* 2009). Given the high dispersal capabilities demonstrated within Zosteropidae, dispersal events between islands may occur relatively frequently, particular during early periods of divergence. Warren *et al.* (2006) suggests that in the absence of niche partitioning, invasibility of resident communities (the ability to out-compete small colonising populations) has probably played a pivotal role in

divergence between insular taxa, with competition between congeneric species limiting gene flow between islands.

In contrast, the propensity for long-distance emigration within the family, and the large number of single-island endemic species, may suggest a rapid loss of dispersal capabilities following the establishment of island populations (Moyle *et al.* 2009; Melo *et al.* 2011). Given the morphological conservatism demonstrated across the range of the family, Moyle *et al.* (2009) suggests that rapid evolutionary shifts in dispersal ability, rather than ecological explanations, were important for the high speciation rates demonstrated by this family. The paradox presented by Zosteropidae, and other groups that exhibit high degrees of differentiation across broad spatial scales, highlights the need for further investigation into factors that influence genetic and phenotypic differentiation in highly vagile groups.

1.1.7. Patterns of genetic and phenotypic divergence

At the population level, considerable interest has been paid to the colonisation history of island populations, particularly the colonisation dynamics and evolutionary processes associated with founding populations (Estoup and Clegg 2003; Clegg *et al.* 2008; Clegg and Phillimore 2010). Bayesian analysis of the colonisation dynamics of *Zosterops lateralis lateralis* has indicated that a large number of effective founders were involved in establishing the south New Zealand, north New Zealand and Chatham Island populations (Estoup and Clegg 2003). In the context of Warren *et al.* (2006) previous hypothesis, the absence of small founding populations within this system may provide support for the idea that comparatively larger populations out-competed smaller populations during early stages of divergence. Furthermore, contrary to Moyle *et al.* (2009) predictions, a recent investigation into the population genetic structure of island populations within the Vanuatu archipelago provided no evidence for a rapid shift in dispersal ability. Instead, populations demonstrated complex gene flow dynamics consistent with high degrees of asymmetrical migration between island populations that persisted long after colonisation (Clegg and Phillimore 2010).

While geographical isolation does not seem to be supported as the primary driver of lineage divergence, the relative contribution of intra-specific competition and ecological divergence remains unclear. Under a scenario of

strong divergent selection pressures, substantial phenotypic divergence is expected (Clegg and Phillimore 2010). While phenotype is remarkably homogenous across the range of the family, divergent phenotypes have been demonstrated in numerous insular taxa (Clegg *et al.* 2002; Frentiu *et al.* 2007; Clegg *et al.* 2008; Melo *et al.* 2011), with insular taxa often found to be substantially larger than their mainland counterparts (Frentiu *et al.* 2007). Furthermore, investigation into recently founded populations has highlighted the potential for rapid differentiation in newly formed insular populations (Clegg *et al.* 2008; Melo *et al.* 2011). In *Zosterops lateralis chlorocephalus*, Clegg *et al.* (2008) found a substantial increase in body size that was estimated to have occurred in fewer than 500 generations after colonisation, consistent with strong directional selection in the early stages of divergence.

Shifts in morphology are not shown to be associated with ecological niche expansion (Scott *et al.* 2003) nor do they coincide with time or degree of genetic isolation (Clegg and Phillimore 2010). Instead, adaptive divergence or strong directional selection towards new or novel environments has been highlighted as an important factor in explaining phenotypic divergence within Zosteropidae (Clegg *et al.* 2008). The 'Dominance hypothesis' has also been suggested as a mechanism for the evolution of large size island Zosteropidae (Robinson-Wolrath and Owens 2003). Yet in the absences of empirical data comparing intra-specific competition within oceanic island and mainland areas, relationships between size, dominance and intra-specific competition are unclear (Robinson-Wolrath and Owens 2003).

1.1.8. Continental systems

Given that each of the five most species-rich avian families is primarily continental (Fringillidae -993 *spp*; Corvidae -647 *spp*; Sylviidae -552 *spp*, Tyrannidae -537 *spp*; Muscicapidae -449 *spp*: Sibley 1990), it could be argued that islands are not among the most important engines of global diversity which should be reflected in research effort. While the last decade has seen an accumulation of studies addressing the relationships, colonisation ability and evolutionary dynamics of insular Zosteropidae (Slikas *et al.* 2000; Clegg *et al.* 2002; Warren *et al.* 2006; Phillimore *et al.* 2008; Clegg *et al.* 2008; Moyle *et al.* 2009; Clegg and Phillimore 2010; Milá *et al.* 2010; Melo *et al*. 2011), genetic relationships and patterns of phenotypic divergence between mainland taxa

have been largely ignored (but see Oatley *et al.* 2012). The African Zosteropidae system illustrates the trends and complexities of both continental and island species. Molecular studies have highlighted that the evolutionary history of Zosteropidae inhabiting the island systems on either side of Africa are closely linked to that of mainland taxa (Warren *et al.* 2006; Melo *et al.* 2011). However, limited taxonomic sampling for mainland areas has hindered extensive molecular investigation.

1.2. Aims of this thesis

1.2.1. Investigating evolutionary patterns and processes in continental island systems (Chapter 2 and 3)

Within Africa the most complex geographical setting exists in East Africa, which encompasses several widely scattered but bio-geographically similar mountain ranges that belong to the eastern Afromontane region. The tops of these montane fragments are covered in cool, moist cloud forest, which is surrounded in the lowlands by dry semi-desert or acacia scrub creating an archipelago-like setting (Moreau 1957). This region has globally significant levels of biological diversity and endemism that has led to it being recognised as a world biodiversity hotspot (Mittermeier *et al.* 1999; Myer *et al.* 2000). However, the mechanisms driving the build-up of diversity within this region remain poorly understood.

Within East Africa the ranges of three *Zosterops* species come into contact, where geographically fragmented montane populations are surrounded by lowland taxa. In some cases, geographical ranges are shown to overlap but ecological ranges remain separate with no evidence of interbreeding (Moreau 1957). These populations provide an excellent system to test temporal, geographic and habitat driven hypotheses of speciation in the eastern Afromontane region. Supporting the montane speciation model (Fjeldså and Lovett 1997; Roy *et al.* 1997), the taxonomic treatment of these montane populations (subspecies of a wider species complex) suggests that they are relics of a previously widespread population (Fry *et al.* 2000). However, in the absence of a species-level molecular phylogeny, alternative models are yet to be explored.

This thesis aims to identify patterns and processes that are driving diversification within the eastern Afromontane biodiversity hotspot by addressing the relationships and evolutionary history of East Africa *Zosterops*. Specifically Chapter 2 will address the following questions:

- Are lowland forest taxa ‘ancient’ relative to montane taxa?
- What is the relative role of past climatic fluctuations in the divergence of montane endemics?
- Have stable montane areas promoted the differentiation of populations leading to aggregates of restricted endemic taxa?

Chapter 3 aims to obtain a nuclear assessment of genetic relationships within East African *Zosterops* to answer the following question:

- Is there concordance between mitochondrial and nuclear assessments of relationships between East African *Zosterops*?
- Do the montane endemics of East Africa represent a single radiation of montane forms or do they represent convergent evolution of a montane phenotype?

1.2.2. Systematics review of African Zosteropidae (Chapter 4)

Africa and its associated island systems (Gulf of Guinea and Indian Ocean) encompass 14 described *Zosterops* species and four *Speirops* species (Dickinson 2003). More than half the African species are offshore endemics with only five species restricted to mainland Africa (Dickinson 2003). While recent molecular studies have given considerable insight into the relationships of insular taxa, the systematics of mainland African forms has received considerably less attention (Moreau 1957; Clancey 1967; Oatley *et al.*2012). By generating a robust molecular phylogeny this thesis aims to uncover relationships within the African Zosteropidae system to evaluate if there is concordance between genetics and the morphological characters previously used to delineate species within mainland Africa. Specifically Chapter 4 will address the following questions:

- Do East African montane endemics represent independent taxonomic units?
- What is the relationship between the restricted highland populations of Mt Cameroon and those of East Africa?
- What are the genetic affinities of yellow-bellied races that have a wide distribution across much of sub-Saharan Africa?
- Do the grey-bellied forms of northeast Africa and southern Africa represent a wider species complex?
- Are the two belly races (yellow and white) that are restricted to the lowlands of east Africa a single species?
- Does the width of the eye-ring or markings on the forepart of the head show any taxonomic affinities?

1.3. References

Cibois A (2003) Mitochondrial DNA phylogeny of babblers (Timaliidae). *The Auk*, **120**, 35-54.

Clancey PA (1967) Taxonomy of the southern African *Zosterops*. *Ibis*, **109**, 318-327.

Clegg SM, Degnan SM, Moritz C, Kikkawa J, Estoup A, Owens IP (2002) Microevolution in island forms: the roles of drift and directional selection in morphological divergence of a passerine bird. *Evolution*, **56**, 2090–2099.

Clegg SM, Frentiu FD, Kikkawa J, Tavecchia G, Owens IPF (2008) 4000 years of phenotypic change in an island bird: heterogeneity of selection over three micro-evolutionary timescales. *Evolution*, **69**, 2393-2410.

Clegg SM, Phillimore AB (2010) The influence of gene flow and drift on genetic and phenotypic divergence in two species of *Zosterops* in Vanuatu. *Philosophical Transactions of the Royal Society B*, *Biological Sciences* **365**, 1077-1092.

Dickinson EC (ed.) (2003) *The Howard & Moore Complete Checklist of the Birds of the World, 3rd Edition*. Christopher Helm, London.

Estoup A, Clegg SM (2003) Bayesian inference on the recent island colonisation history by the bird *Zosterops lateralis lateralis. Molecular Ecology*, **12**, 657-674.

Fjeldså J, Lovett JC (1997) Geographical 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.

Frentiu FD, Clegg SM, Blows M, Owens IPF (2007) Large body size in an island-dwelling bird: a microevolutionary analysis. *Journal of Evolutionary Biology*, **20**, 639–649.

Fry CH, Keith S, Urban EK (Eds) (2000) *The birds of Africa, Vol VI*. London: Academic Press.

Grant PR, Grant BR (2008) *How and why species multiply: the radiation of Darwin's Finches*. Princeton University Press, Princeton.

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

Lerner HRL, Meyer M, James HF, Hofreiter M, Fleischer RC (2011) Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian Honeycreepers. *Current Biology*, **21**, 1838-1844.

Mayr E (1965) Relationships among Indo-Australian Zosteropidae. *Breviora*, **228**, 1-6.

Mayr E, Diamond JM (2001) *The Birds of Northern Melanesia: Speciation, Ecology, and Biogeography*. Oxford University Press, Oxford.

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

Milá B, Warren BH, Heeb P, Thébaud C (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very

small spatial scale in the Mascarene grey white-eye (Aves: *Zosterops borbonicus*). *BMC Evolutionary Biology*, **10**, 158-171.

Mittermeier RA, Myers N, Gil PR, Mittermeier CG (1999) *Hotspots: Earth's biologically richest and most endangered terrestrial ecoregions*. (Cemex, Conservation International and Agrupacion Sierra Madre, Monterrey, Mexico).

Moreau RE (1957) Variation in the western Zosteropidae (Aves). *Bulletin of the British Museum,* **4**, 318–433.

Moyle RG, Filardi CE, Smith CE, Diamond JC (2009) Explosive Pleistocene diversification and hemispheric expansion of a ‘‘great speciator’’. *Proceedings of the National Academy of Science USA*, **106**, 1863–1868.

Myers N, Mittermeier RA, Mittermeier CG, daFonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853-858

Oatley G, Voelker G, Crowe TM, Bowie RCK (2012) A multi-locus phylogeny reveals a complex pattern of diversification related to climate and habitat heterogeneity in southern African white-eyes. *Molecular Phylogenetics and Evolution*, **64**, 633-644.

Phillimore AB, Owens IPF, Black RA, Chittock J, Burke T, Clegg SM (2008) Complex patterns of genetic and phenotypic divergence in an island bird and the consequences for delimiting conservation units. *Molecular Ecology*, **17**, 2839-2853.

Pratt HD (2005) *The Hawaiian Honeycreepers*. Oxford University Press, Oxford.

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

Ricklefs RE, Bermingham E (2007) The causes of evolutionary radiations in archipelagos: passerine birds in the Lesser Antilles. *The American Naturalist*, **169**, 285–297.

Robinson-Wolrath SI, Owens IPF (2003) Large size in an island-dwelling bird: intraspecific competition and the Dominance Hypothesis. *Journal of Evolutionary Biology*, **16**, 1106–1114.

Roy MS (1997) Recent diversification in African greenbuls (Pycnonotidae: *Andropadus*) supports a montane speciation model. *Proceeding of the Royal Society of London B*, *Biological Sciences* **264**, 1337-1344.

Scott SN, Clegg SM, Blomberg SP, Kikkawa J, Owens IP (2003) Morphological shifts in island-dwelling birds: the role of generalist foraging and niche expansion. *Evolution*, **57**, 2147-2156.

Sibley CG, Monroe BL (1990) *Distribution and taxonomy of birds of the world*. Yale University Press, New Haven.

Slikas B, Jones IB, Derrickson SR, Fleischer RC (2000) Phylogenetic relationships of Micronesian white-eyes based on mitochondrial sequence data. *The Auk*, **117**, 355-365.

van Balen S (2008) *Family Zosteropidae (White-eyes).* In Handbook of the birds of the world, Vol. 13: 402–485. Del Hoyo, J., Elliott, A. & Christie, D.A. (Eds). Barcelona: Penduline Tits to Shrikes. Lynx Editions.

Warren B, Bermingham E, Prys-Jones RP, Thebauds C (2006) Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on the Indian Ocean islands. *Molecular Ecology,* **15**, 3769-3786.

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

Chapter 2

The role of Plio-Pleistocene climatic fluctuations in the diversification of East African *Zosterops*

Table of Contents

2.1. Abstract

Background: The eastern Afromontane region encompasses several widely scattered, but bio-geographically similar mountain ranges in eastern Africa and the Arabian Peninsula. This region has globally significant levels of biological diversity and endemism that have led to it being recognised as a world biodiversity hotspot. However, the mechanisms driving the build-up of diversity within this region remain poorly understood. *Zosterops poliogaster* (Montane white-eye) is a montane forest specialist and occurs throughout these regions of East Africa where individual subspecies are endemic to isolated 'sky Islands' (montane forest fragments). Endemic montane populations are ecologically segregated from neighbouring species (*Z. senegalensis* and *Z abyssinicus*) providing an excellent system to test temporal, geographic and habitat driven hypotheses of speciation in the eastern Afromontane region.

Methods: This study provides the first strongly supported phylogenetic assessment of mainland African *Zosterops*. Novel sequence data for the mitochondrial NADH dehydrogenase subunit III (ND3) and cytochrome b (Cyt b) genes (1471 bp) were generated for eleven described *Zosterops* species from African mainland and associated islands. These sequences were analysed implementing both Bayesian inference and maximum likelihood methods. Sequence data for the nuclear transforming growth factor-beta 2 gene (TGFß2) were also generated, but provided no informative sites for phylogenetic analysis. Divergence estimates were inferred using an island calibration and compared to results generated based on the avian molecular clock.

Results: Phylogenetic analyses reveal significant non-monophyly of mainland African *Zosterops* species, specifically *Z. poliogaster* and *Z. senegalensis*. Furthermore, the results reveal that many endemic montane populations are more closely related to taxa with divergent habitat types, elevation distributions and dispersal abilities than they are to populations of restricted endemics that occur in neighbouring montane forest fragments. Divergence estimates indicate that African *Zosterops* diverged very recently (<5Ma). Mean age estimates for the divergence of montane populations (*Z. poliogaster*) coincide with a period of precessional-forced climatic variability during the Plio-Pleistocene.

Discussion: This work rejects the montane speciation model, indicating that the endemic montane populations of *Z. poliogaster* are not relics of a previously widespread population. Instead results reveal that ancestral lineages were in fact adaptive, with niche divergence leading to aggregates of taxa with divergent habitat types, elevation distributions and dispersal abilities. The non-monophyly of mainland African *Zosterops* suggests that traditional morphological characters used to delineate species within Zosteropidae are not informative in an evolutionary context, with results indicating that the current taxonomic framework greatly underestimates *Zosterops* diversity within mainland Africa.

2.2. Introduction

2.2.1. Eastern Afromontane biodiversity hotspot

Understanding the historical processes that drive the divergence of contemporary fauna is a major aim of biogeography and is critical in understanding current species distribution patterns (Moritz *et al.* 2000; Wiens and Donoghue 2004). In spite of this, historical patterns of species-level diversity in some of the world's most diverse regions remains poorly understood. The eastern Afromontane system has been listed as a world biodiversity hotspot region and harbours globally significant levels of diversity and endemism (Mittermeier *et al.* 1999; Myers *et al.* 2000).

Unlike other montane systems, such as the Himalayas or the Andes, highland areas within the eastern Afromontane region are to a considerable extent geographically isolated. This isolation means that it is potentially easier to disentangle in situ speciation events from colonisation events than in other montane systems that exhibit higher degrees of connectivity. Despite being a useful system to examine spatio-temporal relationships, phylogenetic studies are limited to groups with poor dispersal abilities that often only occur within a small area of the eastern Afromontane region (Matthee *et al*. 2004; Blackburn and Measey 2009; Shepard and Burbrink 2009; Voje *et al.* 2009; Lawson 2010; Measey and Tolley 2011). Consequently, the high levels of diversity and endemism seen in more vagile groups that occur throughout the eastern Afromontane region remain poorly understood.

Despite the remarkable taxonomic diversity of African birds there is little consensus on how geological and climatic history has affected patterns of species diversity in Africa (Jetz *et al.* 2004; Fjeldså and Lovett 2007; Fjeldså and Bowie 2008). Previous work suggests that avian species richness within Africa is geographically clustered, whereby species diversity is highest in montane areas (Jetz *et al.* 2004; Fjeldså and Lovett 2007; Fjeldså and Bowie 2008; Linder *et al.* 2012). Recent studies investigating spatial variation in species richness and endemism in the Afrotropics (Jetz *et al.* 2004) and the Neotropics (Rahbek *et al.* 2007; Thomas *et al.* 2008) have demonstrated that current climate alone fails to explain the extraordinary diversity seen in tropical montane regions. Instead, these studies suggest that current models underestimate the importance of historical factors such as past climate and

small-scale niche-driven assembly processes in shaping contemporary species-richness patterns (Jetz *et al.* 2004; Rahbek *et al.* 2007; Thomas *et al.* 2008).

2.2.2. Plio-Pleistocene African climate

The Late Cenozoic African climate can be characterised by short alternating periods of extreme wetness and aridity that are superimposed on a long-term drying trend (deMenocal 1995; Trauth *et al.* 2007). Starting in the mid-Pliocene, African palaeo-climatic records indicate a vegetation shift from closed canopy to open savannah vegetation that has been ascribed to an increase in aridity and a decrease in temperature towards the present. Prior to 2.7Ma, wet phases appear every 400kyr coinciding with maxima in the components of the Earth's eccentricity cycle. However, after 2.7Ma wet phases appear every 800kyr and are correlated with significant global climatic transitions as well as peaks in orbital eccentricity (deMenocal 1995; Trauth *et al.* 2007). Compression of the Inter-Tropical Convergence Zone (ITCZ), as a result of an increase in the pole-equator thermal gradient associated with these global climatic transitions, is thought to have increased the sensitivity of Africa to the effects of precessional forcing, leading to extreme climatic variability (Trauth *et al.* 2007).

It has been widely postulated that these climatic fluctuations would have had a profound effect on the vegetation of Africa (deMenocal 1995; Plana 2004; Trauth *et al.* 2007), causing widespread shifts in Afro-tropical forests and leading to the intermittent fragmentation of the main rainforest biome (and associated fauna) into isolated refugia. This climatic instability is thought to have played an integral role in the evolutionary history of African avifauna (Moreau 1957; Mayr and O'Hara 1986; Fjeldså and Lovett 1997; Roy 1997; Fjeldså and Bowie 2008; Voelker *et al.* 2010). Yet outside the paradigm of the 'Pleistocene Refuge Hypothesis' (Crowe and Crowe 1982; Mayr and O'Hara 1986; Diamond and Hamilton 2009), few phylogenetic studies have sought to explain how historical climate has affected patterns of species-level diversity.

2.2.3. Current hypotheses of diversification

The 'Pleistocene refuge hypothesis' (Crowe and Crowe 1982; Mayr and O'Hara 1986; Diamond and Hamilton 2009) proposes that the repeated isolation

of populations during periods of climatic instability played a primary role in the mechanisms responsible for the current species richness of Africa's tropical rainforest. However, it has been argued that divergence of many lowland species predates the Plio-Pleistocene and on this basis the model has received significant criticism for its use in explaining lowland forest diversity (Fjeldså and Lovett 1997; Roy 1997; Roy *et al.* 2001; Fjeldså and Bowie 2008; Voelker *et al.* 2010). Nevertheless, its application to montane forest systems (The montane speciation model) has been widely accepted, and it has been used to explain the high levels of endemism seen in the avifauna of the montane tropics (Fjeldså and Lovett 1997; Roy 1997; Fjeldså and Bowie 2008; Measey and Tolley 2011 Voelker *et al.* 2010).

Tropical montane regions contain a heterogeneous topography, and vegetation associated with different elevations can vary in its susceptibility to climate change (Fig. 2.1.A). During periods of climatic variability many tropical montane regions remained stable despite global eco-climatic changes. These montane forest habitats or 'sky islands', separated by intervening lowland areas, may have served as historical refugia where previously widespread populations became geographically isolated as they tracked suitable habitat to higher elevations in response to climate change (Fjeldså and Lovett 1997; Roy 1997; Fjeldså and Bowie 2008; Voelker *et al.* 2010; Measey and Tolley 2011). The temporal and spatial variation found in montane regions may have provided the conditions necessary to promote rapid divergence between non-continuous populations that persisted in forested montane areas during the cool and arid climatic episodes of the Plio-Pleistocene (Fig. 2.1.B).

An alternative mechanism of climatic zonation, whereby new species originate as populations adapted to different climatic regimes along an altitudinal gradient, has been documented in several tropical systems (Moritz *et al.* 2000; Ogden and Thorpe 2002; Hall 2005; Kozak and Wiens 2007). The low seasonality seen in East Africa, compared to other more temperate regions, means habitats at different elevational zones would experience reduced overlap in their climatic regimes. The narrowing of climatic profiles between different altitudes produces strong climatic and ecological gradients, which in turn selects for organisms with narrow ecological and climatic tolerances (Moritz *et al.* 2000; Kozak and Wiens 2007). This results in divergent selection across strong environmental gradients. The 'gradient speciation model' (Fig. 2.1.C) may have

played an integral role in the diversification of African avifauna, yet its possible contribution to the high species diversity seen in the montane tropics has been largely ignored.

While the same geographic pattern of species abundance can be explained by both gradient and refuge mechanisms (Moritz *et al.* 2000), these models predict contrasting roles for natural selection. Refuge models (Pleistocene refuge hypothesis and the montane speciation model) are founded on niche conservatism; the inability of populations to adapt to new or changing environmental conditions plays the primary role in geographical isolation, with ecologically similar populations diverging in allopatry (Moritz *et al.* 2000; Wiens and Donoghue 2004; Kozak and Wiens 2007; Wiens *et al.* 2010). In contrast, under the gradient model the ability to adapt to new or changing environmental conditions drives climatic niche divergence (thus population divergence), with differing climatic distributions and/or climatic tolerances limiting gene flow between populations in either allopatry or parapatry (Moritz *et al.* 2000; Ogden and Thorpe 2002; Hall 2005; Kozak and Wiens 2007).

A noteworthy variation of these two models is the vanishing refuge model (Fig. 2.1.D) (Vanzolini and Williams 1981). This model proposes that some populations differentiate to species through directional selection towards a tolerance of less favourable habitats as refuges become too small to retain viable populations. Like the gradient model, the vanishing refuge model is based on niche divergence, yet the latter model requires severe population bottlenecks with subsequent range expansion (Vanzolini and Williams 1981; Moritz *et al.* 2000).

Figure 2.1 Hypotheses of different mechanisms that promote speciation in the montane tropics.
A: Forest cover: Tropical montane regions contain a heterogeneous topography and vegetation associated with different elevations can vary in its susceptibility to climate change. An increase in aridity and a decrease in temperature causes a vegetation shift from closed canopy to open savannah vegetation in lowland areas, resulting in the contraction of forest habitat to higher elevations.
B: Montane speciation model: Climate change causes forest habitat to contract to high elevation refugia that are separated by dry forest and savannah. The fragmentation of forest habitats causes the isolation of forest specialists promoting speciation in allopatry. This model predicts that sister taxa should have restricted distributions occurring in adjacent montane refugia.
C: Gradient speciation model: Climate change results in a narrowing of climatic profiles between different altitudes. The resulting environmental gradient promotes divergent selection between geographically adjacent but distinct habitats. This model predicts that sister taxa occur in adjacent but distinct habitats that have elevationally non-overlapping geographical distributions.
D: Vanishing refuge model: Climate change causes forest habitat contraction and the narrowing of climatic profiles along an altitudinal gradient. Refugia that become too small to retain viable populations promote directional selection towards a tolerance of less favourable habitats (dry forest and savannah). This model predicts that sister taxa should differ in their climatic tolerances.

2.2.4. Study system- East African *Zosterops*

African white-eyes (*Zosterops*) are an excellent group to test temporal, geographic and habitat driven hypotheses of speciation in the montane regions of East Africa. *Zosterops* have a wide distribution, occurring across much of Sub-Saharan Africa and occupy a broad range of habitats and elevations (Moreau 1957). According to currently accepted taxonomy (Dickinson 2003; van Balen 2008), there are three *Zosterops* species that occur within East Africa; *Z. poliogaster* (Montane white-eye) is restricted to montane forest habitats and is ecologically segregated from neighbouring species: *Z. senegalensis* (Yellow white-eye) or *Z. abyssinicus* (White-breasted white-eye) (van Balen 2008). Supporting the montane speciation model (Fjeldså and Lovett 1997; Roy 1997), the taxonomic treatment of these montane populations (subspecies of a wider species complex) suggests that they are relics of a previously widespread population (Fry 2000). However, without a species-level molecular phylogeny, alternative mechanisms of climatic zonation within this group are yet to be explored.

2.3. Aims

Using mitochondrial DNA (mtDNA) and nuclear DNA (ncDNA) data the primary aim of this work is to generate a robust molecular phylogeny for East African *Zosterops* that would enable the assessment of species validity. By combining the resulting molecular phylogeny with information on species distribution, climatic history and divergence time estimates, this study examines whether the evolutionary history of East African *Zosterops* fits predictions of the montane speciation model. Previous avian studies that have investigated the montane speciation model predict that: i) lowland forest taxa should be 'ancient' relative to montane taxa; ii) montane speciation events should coincide with periods of climatic instability during the Pleistocene; and iii) stable montane areas will have promoted the differentiation of populations leading to aggregates of restricted endemic taxa (Fjeldså Lovett 1997, Roy *et al.* 2001; Fjeldså and Bowie 2008). By testing alternative models of speciation, this study attempts to identify whether diversification leading to the current distribution of restricted montane endemics is the result of the niche conservatism (montane speciation model) or niche divergence (gradient speciation model and/or vanishing refuge model).

2.4. Materials and methods

2.4.1. Taxonomic sampling

A total of 135 individuals representing 11 described *Zosterops* species (Dickinson 2003) from across continental Africa and associated islands are included in this study (Appendix I). Within East Africa, 51 *Z. poliogaster* (Montane white-eye) tissue samples were collected from five isolated montane forests, giving an exceptional coverage of the restricted distributions of the four subspecies: *Z. p. silvanus* (Taita Hills and Mt Kasigau); *Z. p. mbuluensis* (Chyulu Hills and Ol Doinyo Orok); *Z. p. kulalensis* (Mt Kulal); and *Z. p. kikuyuensis* (Central Kenyan Highlands) (Fig. 2.2). Additional sequences for *Z. p. winifredae* (1 individual) were obtained from the National Centre for Biotechnology Information (NCBI) database, which enabled the phylogenetic assessment of six of the eight recognised *Z. poliogaster* subspecies (Dickinson 2003). It was not possible to obtain samples for three subspecies of *Z. poliogaster*, represented by the two Ethiopian subspecies, *Z. p. poliogaster* and *Z. p. kaffensis*, and the Tanzanian subspecies, *Z. p. eurycricotus.*

To check the possible affinities between *Z. poliogaster* and *Z. senegalensis* (Yellow white-eye) populations occurring at higher elevations (2 subspecies; 29 samples), four high elevation populations of the race *Z. s. jacksoni* (Kenya) and three populations of the race *Z. s. stierlingi* (Tanzania) were included. Within Kenya an additional 21 samples of *Z. abyssinicus* (White-breasted white-eye) were collected from lowland (<1000m) scrub and riverine areas. These represented two (*Z. a. flavilateralis* and *Z. a. jubaensis*) of the four described mainland African *Z. abyssinicus* subspecies that are found throughout the lowlands of northeast Africa.

In order to test species monophyly and biogeographic scenarios, 33 samples were obtained from outside of East Africa. These represent: an insular *Z. abyssinicus* race from the Island of Socotra, Gulf of Aden (*Z. a. socotranus*); three subspecies of *Z. pallidus* from South Africa (*Z. p. pallidus*, *Z. p. capensis*, *Z. p. virens*); two *Z. senegalensis* subspecies from Ghana and Cameroon (*Z. s. senegalensis*, *Z. s. stenocricotus*); a Congolese *Z. senegalensis* (DRC) form (not identified to the sub-specific level); and, in addition, representatives for the principle lineages in the Gulf of Guinea and the Indian Ocean island systems.

Figure 2.2: Distribution of African Zosteropidae samples. A: Areas within the eastern Afromontane region (red). Image modified from www.conservation.org **B:** Distribution of *Zosterops* samples from outside the eastern Afromontane Region**.** Image modified from www.mapsof.net **C:** Distribution of *Zosterops* samples within the east Afromontane region indicating sampling localities. Image modified from www.vidiani.com (Kenya) and mapsof.net (Tanzania).

Sequences for an Asian species (*Z. palpebrosus*) and an Australian species (*Z. lateralis*) were acquired from the NCBI database, in addition to sequence data for *Stachyris whiteheadi* (Chestnut-faced babbler) whose suitability as an out-group has been shown in previous studies (Warren *et al.* 2006; Melo *et al.* 2011). Voucher numbers, collection localities, and NCBI accession numbers are listed in Appendix I.

2.4.2. Molecular markers

A multi-marker approach was used in this study, which generated sequence data for both mitochondrial (mtDNA) and nuclear genes (ncDNA). The mitochondrial protein coding genes cytochrome b (Cyt b) and NADH dehydrogenase subunit III (ND3), in addition to the nuclear transforming growth factor-beta 2 gene (TGFß2), were selected for this study. All genes are widely used across avian phylogenetic studies (Prager *et al.* 2008; Nguembock *et al.* 2009; d'Horta *et al.* 2011; Yeung *et al.* 2011) and have proved useful for revealing both relatively deep and shallow level relationships within the genus *Zosterops* (Warren *et al.* 2006; Phillimore *et al.* 2008; Moyle *et al.* 2009; Melo *et al.* 2011).

2.4.3. DNA extraction

Blood samples were taken from mist-netted specimens and stored in ETOH (99%) or Queen's lysis buffer. Samples were extracted from both mediums using a DNeasy Blood and tissue kit (Qiagen). Manufacturer's protocol was followed, with a minimum incubation period of two hours at 56ºC with a final elution of 200µl.

2.4.4. Generation of mitochondrial and nuclear sequence data

Amplification of the ncDNA gene TGFß2 and the mtDNA gene ND3 was performed using published primers (Table 2.1) (Helm-Bychowski and Cracraft 1993; Chesser 1999; Primmer *et al.* 2002). In order to obtain a larger proportion of the Cyt b gene, the published primer H16065 was used alongside three newly designed primers (Table 2.1), allowing for the amplification of two overlapping fragments that together constituted the entire Cyt b gene (1123 bp). The primer-designing program Primer 3 version 0.4 was used to check primer melting temperature (T_m), GC content and the presence of palindromes

(sequence regions that may be read the same way in either direction) and hairpin loop structures (structures formed by the complementary binding of regions along a primer sequence).

For both mitochondrial and nuclear genes, PCR amplifications were performed in 15µl volumes with 2µl total genomic DNA, 9.7µl ddH_2O, 1.5µl 10X PCR buffer, 0.75µl $MgCl_2$ (50mM), 0.15µl dNTPs (2.5mM of each), 0.45µl of each primer (10mM) and 0.15µl *Taq* DNA polymerase (5 units/µl). Thermal cycling conditions for all three genes are reported in Table 2.1. Purification of amplified PCR products was performed using a 10µl volume of Microclean (5ml NaCl (5M), 0.1ml of Tris-HCL (1M), 0.02ml of EDTA (0.5M), 20g of PEG8000, 0.86ml of $MgCl_2$ and 24.8ml ddH_2O). Samples were incubated at room temperature for 10 minutes and then centrifuged at 4000 rpm for 1 hour to pellet DNA. Centrifuged products were then inverted and briefly centrifuged for 1 minute at 1000 rpm. The DNA pellet was then re-suspended in 4µl ddH_2O. Cycle sequencing reactions were carried out in 10µl volumes using 0.25µl BigDye^{TM} Terminator (PE Applied Biosystems), 2.075µl ABI sequencing buffer, 0.176µl primer (10µM) 1.5µl of the purified PCR product and 7.5µl of ddH_2O with cycle sequencing reactions following standard ABI protocols. Cycle sequencing products were purified using an ETOH/EDTA clean up and sequenced on an ABI 3730 DNA analyser (Applied Biosystems).

2.4.5. Sequence analysis

A TGFß2 dataset (582 base pair (bp)) was generated for a subset of the taxa to assess phylogenetic signal (Appendix I). This gene fragment provided no informative sites and therefore is discounted from subsequent analyses. A total of 1471 bp of sequence data was obtained for all individuals from the mitochondrial ND3 and Cyt b genes. Chromatograms of complementary fragments were checked by eye before producing contigs (sequence read resulting from the reassembly of DNA fragments) in the program Sequencher version 4.8. Sequence data were then aligned in Clustal W version 1.83 using default settings with the resulting alignment checked by eye in the program SE-AL version 2.0.

Table 2.1. Primers and thermal cycling conditions used in amplification and sequencing of genes

Gene	Primer Name	Primer sequence	Reference	PCR			
				Denature	Annealing	Extension	Cycles
ND3	L10755	5'-GACTTCCAATCTTTAAAATCTG-3'	Chesser (1999)	94°C - 30s	58°C - 40s	72°C - 40s	35
	H11151	5'-GATTTGTTGAGCCGAAATCAAC-3'	Chesser (1999)				
Cyt-b	CytbF25	5' GGCTCTCAATCTTCGTAAAAACC-3'	Cox unpublished	94°C - 30s	60°C - 60s	72°C - 90s	35
	CytbR649	5'-GGGTGGAATGGGATTTTGTC-3'	Cox unpublished				
	CytbF409	5'-GTAGGCTACGTCCTACCCTGAG-3'	Cox unpublished				
	CytbH16065	5'-GAGTCTTCAGTCTCTGGTTTACAAGAC-3	Helm-Bychowski and Cracraft (1993)				
TFGß2	TGF5	5'-GAAGCGTGCTCTAGATGCTG-3'	Primmer *et al.* (2002)	94°C -30s	56°C - 30s	72°C - 45s	35
	TGF6	5'-AGGCAGCAATTATCCTGCAC-3'	Primmer *et al.* (2002)				

Table 2.2. Gene regions and models of sequence evolution

Gene region	Cyt-b		ND3		Concatenated	
Taxa	139		139		139	
Base pairs	1125		349		1471	
Variable sites	248		116		364	
Model of sequence evolution	Best –lnL	Used	Best –lnL	Used	Best –lnL	Used
	GTR+G	GTR+G	TIM2+G	GTR+G	TPM3uf+G	GTR+G
Log likelihood (-lnL)	4431.5657	4431.5657	2045.3537	2044.9947	6530.1418	6535.8204
Gamma distribution	0.130	0.130	0.1660	0.1670	0.1500	0.1450
Base composition						
A	0.2969	0.2969	0.2971	0.2971	0.1310	0.2993
C	0.3387	0.3387	0.3549	0.3514	0.3325	0.3414
G	0.1286	0.1286	0.1159	0.1162	0.1321	0.1251
T	0.2358	0.2358	0.2321	0.2353	0.2249	0.2343
Substitution rates						
Tr						
a-g	24.1194	24.1194	47.3419	48.9830	18.1998	33.8657
c-t	17.3914	17.3914	25.1619	36.9598	18.1998	22.3048
Tv						
a-c	2.0657	2.0657	2.5289	4.0886	1.5432	2.5205
a-t	1.1459	1.1459	2.5289	3.1012	1.0000	1.6180
c-g	1.2314	1.2314	1.0000	1.7411	1.5432	1.5549
g-t	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

To guard against the possibility of having amplified nuclear copies of mitochondrial genes (numts), alignments were checked to ensure that they contained no gaps, insertions or deletions. Sequence data were translated into amino acids using the vertebrate mitochondrial translation table in MacClade version 4.08a and checked to ensure there were no stop codons. Variation in base composition for both genes was assessed using the X^2 test of homogeneity, implemented in PAUP (Swofford 2003). Substitution saturation at different codon sites, specifically the third codon position, was assessed using an entropy-based index of substitution saturation (Xia *et al.* 2003) implemented in Dambe version 4.5.56 (Xia and Xia 2001).

The best model of molecular evolution for each dataset was determined using JModeltest version 3.0 using the Akaike Information Criterion (AIC) (Table 2.2). The GTR+G model of sequence evolution was selected in all phylogenetic analyses, as it was the most complex model of DNA substitution (six substitution rate parameters) that was available in both GARLI and MrBayes. This appeared justified because: i) the GTR+G model exhibited extremely similar log likelihood (–lnL) values when compared to the models TIM2 and TPM3uf that were selected for the ND3 and concatenated mtDNA datasets retrospectively (Table 2.2); ii) All three models of evolution are extremely similar with the GTR+G (six rate parameters) being slightly more complex than TIM2+G (four rate parameters) and TPM3uf+G (three rate parameters) (Table 2.2).

2.4.6. Phylogenetic analyses

In order to test congruence of phylogenies recovered by different methods of phylogenetic inference both Maximum Likelihood (ML) and Bayesian Inference (BI) were implemented. In a 'total evidence' approach (Huelsenbeck *et al.* 1996), ML analyses (Huelsenbeck and Crandall 1997) were performed on both the concatenated and individual gene datasets using GARLI (Genetic Algorithm for Rapid Likelihood Inference, version 0.951). Six search replicates were run to obtain a measure of confidence for the searching parameters. For search replicates 1-3, ML trees were constructed with model substitution rates applied from JModeltest, while in search replicates 4-6, substitution rates were estimated. For each dataset the six search replicates produced very similar trees and -InL scores, with estimated rates performing

slightly better than fixed rates. For this reason node support for each dataset was ascertained with 1000 non-parametric bootstraps (BS) of the ML tree generated using a GTR+G model of sequence evolution and estimated substitution rates. ML trees were summarized using a 50% majority-rule consensus tree.

BI analyses were implemented on the concatenated dataset in MrBayes version 3.1.2 (Huelsenbeck and Ronquist 2001). Analyses were run on the concatenated dataset, partitioning by gene to account for any potential variation between gene regions. Although results from DAMBE indicated little saturation of the third codon position ($I_{SS} < I_{SS.C}$, P=0.00), a second Bayesian analysis was run, additionally separating the first and second codon positions from the third. Base frequencies were estimated for both analyses and evolutionary rates were allowed to vary across partitions under a dirichlet prior. Starting from a random tree, four Metropolis-coupled Markov Chain Monte Carlo (MCMC) chains (temp=0.2) were run simultaneously for 2,000,000 generations, sampling every 100 generations with a burn-in of 7500. Convergence of the MCMC runs was assessed graphically using TRACER version 1.4.1, with the final tree constructed from 12,500 post burn-in trees. Support is assessed using Bayesian posterior probabilities (BPP).

2.4.7. Estimation of divergence times

Divergence estimates were performed using a relaxed clock molecular dating method (Drummond *et al.* 2006), implemented in BEAST version 1.48 (Drummond and Rambaut 2007). A likelihood ratio test implemented in PAUP* version 4.0b10 was used to test for clock-like evolution in the Cyt b and ND3 datasets. Results failed to reject the null hypothesis of rate constancy, resulting in enforcement of the molecular clock. Since a fossil record for the Zosteropidae is lacking (Moyle *et al.* 2009), two approaches were employed to estimate divergence times. The first approach uses the date of origin of a volcanic island as a calibration for an endemic radiation. This approach has been used in several other studies (Fleischer *et al.* 1998; Warren *et al.* 2006; Moyle *et al.* 2009; Lerner *et al.* 2011; Melo *et al.* 2011) and assumes the birds have diversified *in situ*. Under this approach the maximum age of divergence between closely related taxa occupying neighbouring islands is constrained to be the age of the youngest island, representing the earliest possible date for

colonisation. Following assumptions discussed in previous studies (Fleischer *et al.* 1998; Warren *et al.* 2003), the maximum age estimate for the volcanic origin of Grande Comore at 0.5Ma (Nougier *et al.* 1986) is used to calibrate the node separating the lowland Grande Comore white-eye (*Z. maderaspatanus kirki*) from other taxa in the *maderaspatanus* clade (Warren *et al.* 2006).

The second approach uses the average pair-wise substitution rate of 2.1% for the Cyt b gene. Weir and Schluter (2008) generated this rate from 74 calibrations spanning 12 taxonomic orders and 12 million years. The calibrations used in this study were obtained from fossils and the ages of oceanic islands, mountain ranges and land bridges. Although minor but significant variations in rates were noted across lineages (Weir and Schluter 2008), in the absence of suitable internal calibration points this consensus molecular clock rate is advocated (Fritz *et al.* 2011; Voelker *et al.* 2010).

In the calibrated approach, divergence time estimates were generated from the concatenated Cyt b and ND3 dataset, while divergence estimates obtained from the consensus molecular clock rate (2.1%) were generated from the Cyt b dataset only. Both approaches used the same starting tree that was generated from BI of the concatenated dataset (partitioned by gene and codon position); however in the consensus clock rate approach, *Z. lateralis* was pruned from the tree, as no Cyt b sequences were available for this sample. For both analyses, two independent Markov Chain Monte Carlo (MCMC) analyses were run starting from a user specified tree (BI of the concatenated dataset, partitioned by gene and codon position). Chains were run for 2,000,000 generations using a constant rate Yule speciation prior (assumes a constant speciation rate per lineage) and a GTR+G model, sampling every 1,000 generations with a burn-in of 10%. Convergence of the two independent MCMC runs was assessed graphically in TRACER version 1.4.1 (Drummond and Rambaut 2007), with the posterior distribution being summarised in the program TREE ANNOTATOR version 1.4.8 (Drummond and Rambaut 2007).

2.5. Results

2.5.1. Sequence data

Given the comparatively lower rate of evolution of ncDNA genes relative to mtDNA genes the use of ncDNA sequence data can be particularly problematic in recently divergence group such as African *Zosterops*. The ncDNA TGFß2 dataset (582 bp) that was generated in this study provided limited sequence variation leading to a lack of phylogenetic signal. As a result, subsequent analyses were reliant on mtDNA that demonstrated comparatively higher sequence variability.

The concatenated dataset consists of 1471 bp, constituting the entire Cyt b (1123 bp) and ND3 (348 bp) genes. The ND3 dataset is largely complete with the exception of four samples (Appendix I) that failed to amplify. For the Cyt b dataset, sequences were obtained from two overlapping fragments. There are missing or incomplete sequences for 21 individuals, where one or both of the overlapping primers failed to amplify (Appendix I). A further 20 incomplete Cyt b sequences were obtained from the NCBI database (~310 bp: starting 83 bp from the 5' end of fragment 1) (Appendix I). Missing or incomplete sequence data were coded as missing data in all phylogenetic analyses. The concatenated data set contains 364 variable sites and has a relatively even base composition (A: 29.9%, C: 34.1%, G: 12.5%, T: 23.4%). Results for a X^2 test of homogeneity shows no significant difference in base frequencies (Cyt b P=1, ND3 P=1) between in-group taxa. Results from an entropy-based index of substitution saturation (Xia *et al.* 2003) indicate that I_{SS} values are significantly lower than $I_{SS.C}$ values (Cyt b P=0.00, ND3 P=0.00), suggesting little saturation of the third codon position.

2.5.2. Phylogenetic relationships

Both ML and BI analyses resulted in highly congruent trees, with the majority of relationships resolved (Fig. 2.3). Support for these relationships is generally good, although unsurprisingly BPP values are higher than BS (Erixon *et al.* 2003). Phylogenetic reconstructions that were generated for both individual and concatenated datasets result in highly congruent trees. In agreement with the result from DAMBE, the two partitioning strategies used in

BI analyses (partition by i) gene and ii) gene and codon position) resulted in the same topologies, indicating no detrimental influence of third codons.

Continental African *Zosterops* form two major clades (Fig. 2.3; node A, node B) that both contain oceanic island radiations from different island archipelagos. The analyses further reveal considerable non-monophyly of mainland African *Zosterops* taxa, with three of the four continental species (*Z. poliogaster, Z. senegalensis* and *Z. pallidus*) rendered non-monophyletic. In contrast to the non-monophyly of described species, there is strong support for the monophyly of individual subspecies, especially within *Z. poliogaster* and *Z. senegalensis*.

All *Z. poliogaster* subspecies sampled in this study form independent well-supported clades that are polyphyletic with respect to each other. With the exception of *Z. s. jacksoni* (Kenyan highlands), which forms a well-supported clade (BPP 1.0/ BS 91%) with the nominate subspecies *Z. s. senegalensis* (Ghana), the remaining *Z. senegalensis* subspecies included in this study are recovered as independent clades. The low support for the placement of sample ZMUC131324 as sister to the main *Z. s. jacksoni* clade could be as a result of missing data (Appendix I). Analyses indicated considerable genetic structure within *Z. s. jacksoni;* however there is no support for relationships within this clade.

The two South African *Z. pallidus* subspecies (*Z. p. capensis* and *Z. p. virens*) form a clade (BPP 0.89 / BS 70%). The position of the single sample of *Z. p. pallidus* (AP50340) is unclear within clade A4, although there is no support for its placement as sister to the other Z. *pallidus* subspecies. Two mainland *Z. abyssinicus* subspecies (*Z. a. jubaensis* and *Z. a. flavilateralis*) that have a parapatric distribution form a single clade (BPP 1 / BS 75), with no support for any division between 'subspecies'. Results indicate that the mainland *Z. abyssinicus* is distinct from insular members of *Z. a. socotranus* (Socotra), rendering *Z. abyssinicus* an unnatural grouping. Both BI and ML analyses place *Z. a. socotranus* as sister to the two major African clades A and B (BPP 0.96).

African clade A (BPP 0.86) supports the inclusion of two island radiations: the Gulf of Guinea 'Oceanic' white-eyes (GGO) (Melo *et al*. 2011); and the Indian Ocean 'Maderaspatanus' clade (IOM) (Warren *et al.* 2006). Two internal mainland African clades are also recovered. These include members of the East Africa *Z. poliogaster* species-complex, the southern African *Z. pallidus*

species-complex, *Z. abyssinicus* from the lowland of Kenya and *Z. senegalensis* from Tanzania.

There is good support for the placement of the GGO white-eyes at the base of African clade A (BPP 0.86), with the IOM clade embedded between the two mainland African sub-clades (BPP 1.0 /BS 82%). Representatives of the Montane white-eye species-complex occur in both of these continental sub-clades (Fig. 2.3. A2, A4). The mainland sub-clade A2 (BBP 0.92 / BS 69%) contains *Z. p. mbuluensis* from southern Kenya (Chyulu Hills) that is a sister to a clade containing the two lowland *Z. abyssinicus* subspecies: *Z. a. flavilateralis;* and *Z. a jubaensis*.

The second mainland sub-clade (A4) (BBP 1.0 / BS 93%) contains two *Z. poliogaster* subspecies: *Z. p. silvanus* from southern Kenya (Taita Hills); and *Z. p. winifredae* from northern Tanzania (S. Pare Mts). *Z. p. winifredae* forms a well-supported clade (BBP 1.0 / BS 99%) with the two southern African *Z. pallidus* subspecies (*Z. p. capensis* and *Z. p. virens*) and *Z. s. stierlingi* from Tanzania. Within clade A4, *Z. p. silvanus* is basal to a clade containing *Z. p. pallidus*, *Z. p. winifredae*, *Z. s. stierlingi* and two *Z. pallidus* subspecies under BI (Fig. 2.3.), or is alternatively recovered as sister to the taxon *Z. p. pallidus* (South Africa) under ML. The placement of the latter taxon is weakly supported.

The Gulf of Guinea 'mainland' white-eyes (GGM), along with a clade of Congolese *Z. senegalensis,* are supported as sister to all other taxa within Africa clade B, although their relationships with respect to each other are unresolved. Within clade B, there is good support for an internal mainland African clade (B2: BPP 1.0/ BS 95%). This clade contains two independent clades of *Z. poliogaster* subspecies: *Z. p. kulalensis* (Mt Kulal); and *Z. p. kikuyuensis* (Mt Kenya / Aberdares range) from northern Kenya. These two taxa are not however monophyletic, as the placement of a clade containing two *Z. senegalensis* subspecies (*Z. s. jacksoni* and *Z. s. senegalensis*) as sister to *Z. p. kikuyuensis* renders them paraphyletic. The Ancient Indian Ocean white-eyes (AIO) fall outside of the African radiation (Fig. 2.3, node 1), with this clade recovered as sister to the Asian taxa *Z. p. egregious* and *Z. p. palpebrosus*, although their position with respect to *Z. lateralis* (Australia) is less clear at the base of the tree.

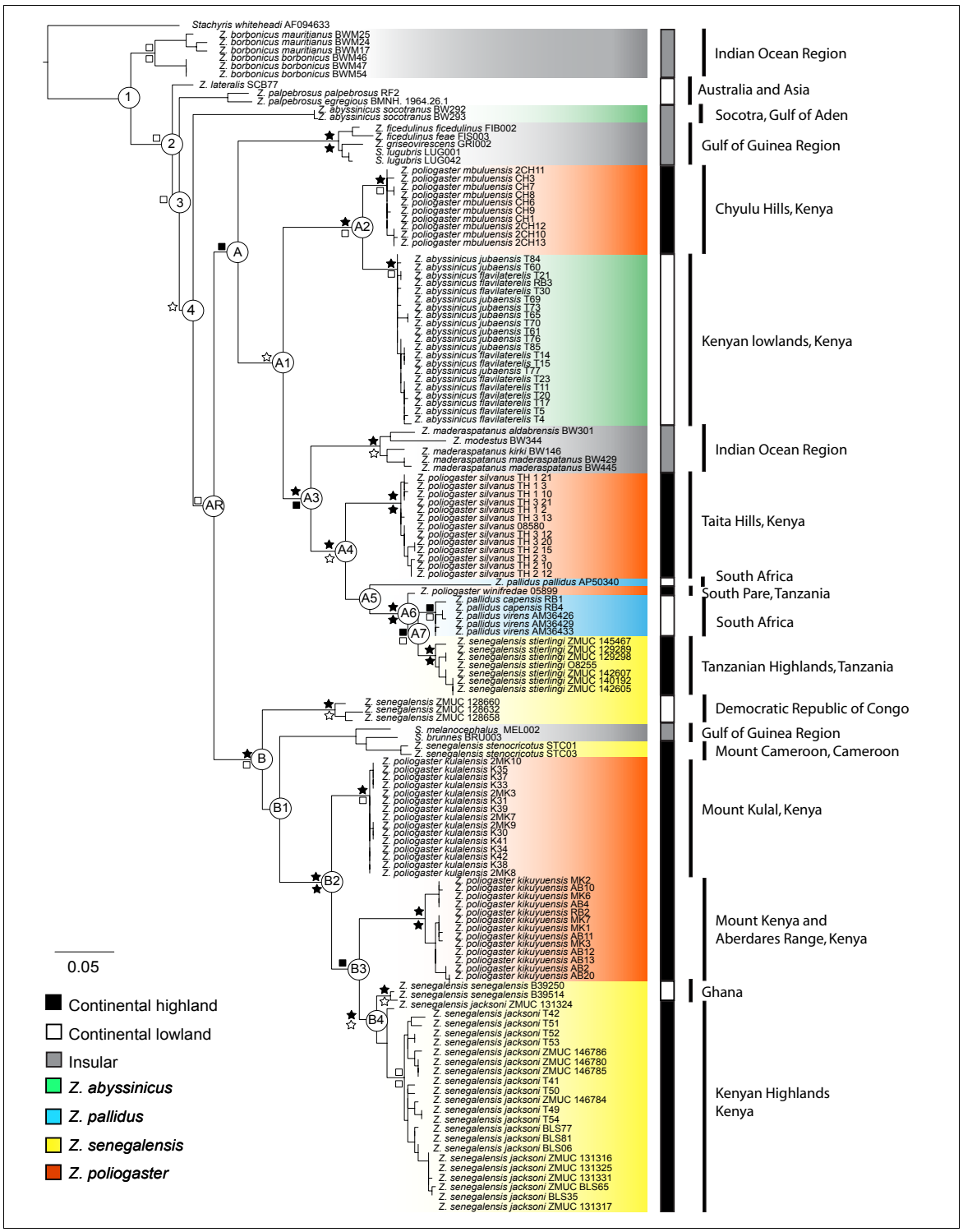


Figure 2.3: Bayesian Inference (BI) tree of African Zosteropidae generated from the concatenated mitochondrial ND3 and Cyt b datasets, partitioned by gene and codon position. Branch lengths are proportional to the degree of sequence divergence. Node support in the form of Bayesian posterior probabilities (BPP) and ML bootstrap values (1000 replicates) are displayed above and below the branches respectively. ★ Indicates nodes with >95% BPP/BS, ☆ indicates nodes with >90% BPP/BS, ■ indicates nodes with >80% BPP/BS and □ indicates nodes with > 50% BPP/BS. Nodes with < 50% BPP/BS are not shown. Key nodes are labelled 1-4, AR, A-A7 and B-B4. Taxa are labelled using full trinomial nomenclature, following the taxonomy of Dickinson (2003).

2.5.3.Temporal divergence

Irrespective of molecular dating method, African *Zosterops* are estimated as being a very recently diverged group (<5Ma). Divergence time estimates obtained from the island calibration approach are approximately 2.8 times younger than those obtained when the commonly employed 2.1% clock rate (Weir and Schluter 2008) is applied (Table 2.3). Previous estimates of the molecular rate of evolution in Zosteropidae using independent island calibrations (Warren *et al.* 2006; Moyle *et al.* 2006; Melo *et al.* 2011) have documented significantly faster rates of evolution than 2.1%.

2.5.4.Calibrated approach

For the calibrated approach, results indicate that diversification of African *Zosterops* occurred during a period of climatic instability associated with the Plio-Pleistocene (Fig 2.4; Table 2.3). Mean age estimates indicate that divergence of mainland African taxa is associated with cool and arid climatic episodes, whilst colonisation of the surrounding island system can be linked to a period of extreme wetness. However, broad confidence intervals (95% CL) on age estimates indicate that this relationship is not supported statistically.

Divergence estimates place initial diversification of African *Zosterops* (Fig. 2.4, node AR) in the Lower Pleistocene (1.54 Ma CI 95% 1.11, 2.11) with divergence in clades A and B occurring soon afterwards (1.44Ma and 1.14Ma respectively). Mean age estimates for these events coincide with a period of reduced moisture availability that is associated with a phase of heightened aridity between 1.7 and 1.1Ma. Mean age estimates for the re-colonisation of the Gulf of Guinea (node B1, 1.06 CI 95% 0.74, 1.47) and Indian Ocean system (node A3, 1.06 CI 95% 0.75, 1.46) coincide with a period of wetness and humidity that occurred between 1.1 and 0.9Ma (Trauth *et al.* 2007).

Results indicate that the independent diversification of all endemic montane *Z. poliogaster* subspecies occurred within the last 0.88Ma. This corresponds to a period of reduced moisture availability associated with the cool and arid conditions that occurred after 0.9Ma (Trauth *et al.* 2007). Results further indicate that divergence of montane forms occurred in two phases. Mean age estimates place the independent divergences of *Z. p. silvanus* (node A4) and *Z. p. kulalensis* (node B2) at 0.88 and 0.78Ma respectively. These dates correspond to early stages of aridity that occurred shortly after a period of

extreme wetness that ended 0.9Ma. The second phase of divergence occurred in a more advance period of aridity between 0.63 and 0.29Ma (Trauth *et al.* 2007) and resulted in multiple divergence events between highland forest and lowland savannah specialists.

Divergence between the endemic montane *Z. p. kikuyuensis* (Mt Kenya and Aberdares Range) and a clade containing two *Z. senegalensis* subspecies (*Z. s. senegalensis* and *Z. s. jacksoni)* that occupy different elevational distributions and habitat types is estimated at 0.63Ma (node B3, CL 95% 0.42, 0.90). Divergence between highland *Z. s. jacksoni (Kenya*) and lowland *Z. s. senegalensis* (Ghana) is estimated at 0.39Ma (node B4, CL 95% 0.23, 0.58), coinciding with the time estimate for divergence between the endemic montane *Z. p. mbuluensis* (Chyulu Hills) and two lowland subspecies of *Z. abyssinicus* (*Z. a. jubaensis* and *Z. a. flavilateralis*) (node A2, 0.39Ma CL 95% 0.24, 0.59). Divergence between the endemic montane *Z. p. winifredae* (S. Pare Mountains) and a clade containing lowland *Z. pallidus* and highland *Z. s. stierlingi* is estimated at 0.37Ma (node A6, CL 95% 0.21, 0.56), with divergence between *Z. pallidus* (*Z. p. capensis* and *Z. p. virens*) and *Z. s. stierlingi* occurring soon afterwards (0.29Ma, CL 95% 0.17, 0.43).

2.5.5.Clock rate approach

For the conservative molecular clock rate (2.1.%) approach age estimates for the diversification of this group would coincide with Pliocene tropical forest retraction (Hamilton and Taylor 1991; Feakins *et al.* 2005; Sepulchre *et al.* 2006), while divergence of Montane forms would coincide with earlier cycles of precessional-forced climatic variability that occurred during the late Pliocene/early Pleistocene (Trauth *et al.* 2007). Divergence estimates place initial diversification of African *Zosterops* in the lower Pliocene (4.36Ma CI 95% 3.64, 5.10). Mean age estimates place initial divergence in clades A and B at 4.16 Ma (CI 95% 3.44, 4.94) and 3.19 Ma (CI 95% 2.54 3.87) respectively. These dates coincide with a period of pronounced aridity that is thought to have led to a substantial expansion of savannah with subsequent retraction of tropical forest (Hamilton and Taylor 1991; Feakins *et al.* 2005; Sepulchre *et al.* 2006). Mean divergence estimates place the independent diversification of all endemic montane *Z. poliogaster* subspecies within the last 2.34Ma corresponding with climatic fluctuations of the lower Pleistocene.

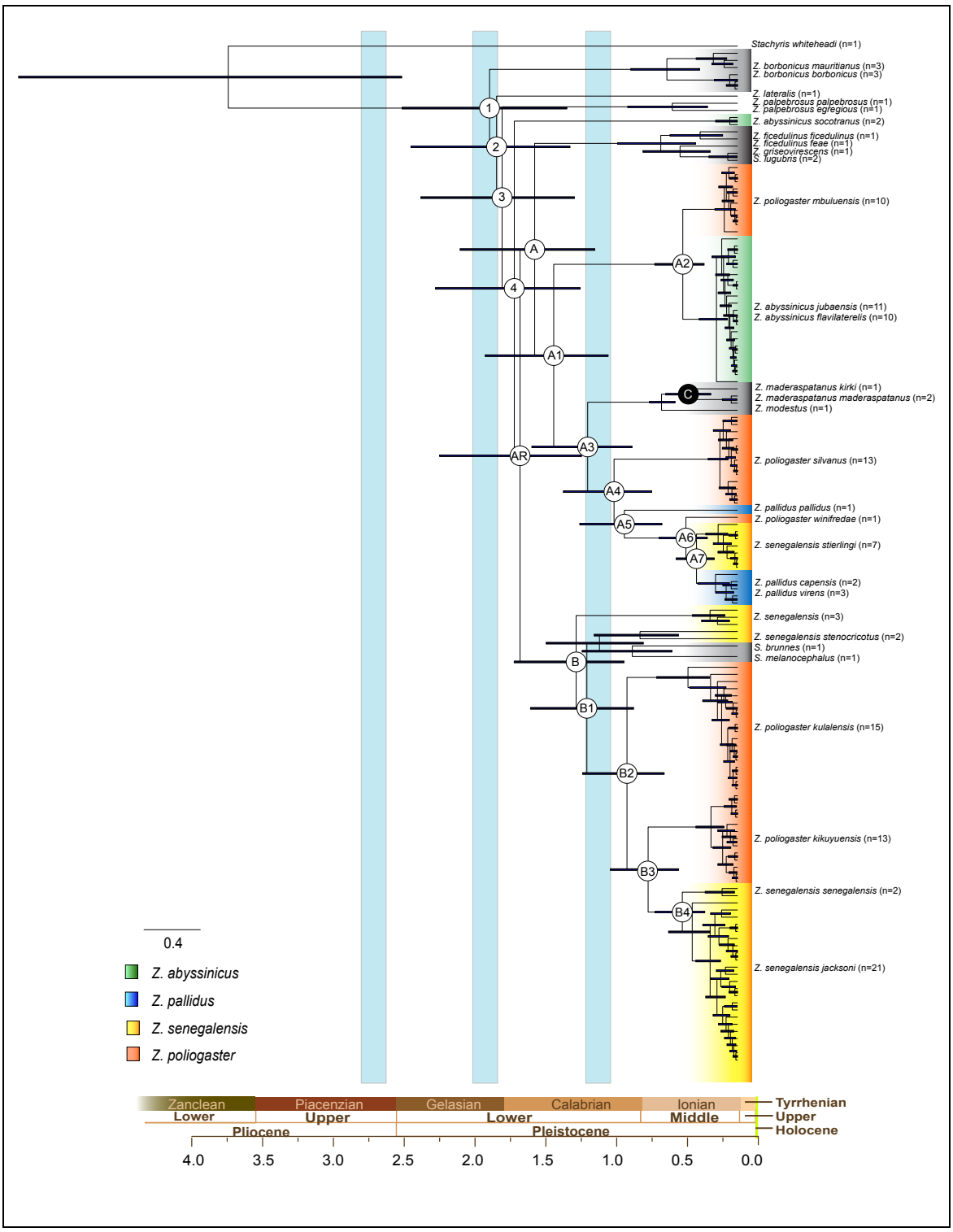


Figure 2.4: Temporal divergence of African Zosteropidae. Chronogram obtained under a Bayesian relaxed clock method (implemented in BEAST version 1.48) using the concatenated mtDNA dataset. The maximum estimate for the volcanic origin of Grande Comore at 0.5Ma (Nougier *et al.* 1986) was used to calibrate node C; the node separating the lowland Grande Comore white-eye (*Z. maderaspatanus kirki*) from other taxa in the *maderaspatanus* clade (Warren *et al.* 2006). Key nodes are labelled 1-4, AR, A-A7 and B-B4. Dark purple bars represent the 95% confidence intervals of age estimates. The pale blue background indicates three periods of extreme wetness and humidity that are estimated to have occurred between 2.7-2.5 Ma, 1.9-1.7 Ma and 1.1-0.9 Ma (Trauth *et al.* 2007).

Table 2.3. Divergence time estimates for African Zosteropidae

Node	Divergence time estimates (Ma)									
	Analysis A (Calibration of 0.5Ma on node C)		Analysis B (mutation rate of 2.1%)		Warren *et al.* (2006)		Melo *et al.* (2011)		Moyle *et al.* (2009)	
	Mean	95% CI	Mean	95% CL	Mean	95% CL	Mean	95% CL	Mean	95% CL
1	1.76	1.21, 2.38	4.90	4.07, 5.87	-	-	-	-	-	-
2	1.71	1.19, 2.32	-	-		-	-	-	-	-
3	1.67	1.16, 2.25	4.74	3.97, 5.67	1.73	-	-	-	-	-
4	1.58	1.12, 2.14	4.51	3.80, 5.33	1.63	-	-	-	-	-
AR	1.54	1.11, 2.11	4.36	3.64, 5.10	1.59	-	-	-	-	-
A	1.44	1.01, 1.97	4.16	3.44 4.94	-	-	1.60	1.34, 1.89	-	-
A1	1.30	0.92, 1.79	3.61	2.96, 4.38	-	-	-	-	-	-
A2	0.39	0.24, 0.59	1.12	0.69, 1.57	-	-	-	-	-	-
A3	1.06	0.75, 1.46	3.02	2.40, 3.70	1.20	-	1.20	0.98, 1.39	-	-
A4	0.88	0.61, 1.23	2.34	1.80, 2.95	-	-	-	-	-	-
A5	0.80	0.54, 1.12	2.05	1.50, 2.62	-	-	-	-	-	-
A6	0.37	0.21, 0.56	1.54	0.94, 2.23	-	-	-	-	-	-
A7	0.29	0.17, 0.43	1.09	0.68, 1.51	-	-	-	-	-	-
B	1.14	0.81, 1.59	3.19	2.54, 3.87	-	-	-	-	-	-
B1	1.06	0.74, 1.47	2.91	2.34, 3.58	-	-	1.25	1.02, 1.47	1.12	0.89,1.35
B2	0.78	0.52, 1.10	2.23	1.68, 2.79	-	-	-	-	-	-
B3	0.63	0.42, 0.90	1.68	1.25, 2.19	-	-	-	-	-	-
B4	0.39	0.23, 0.58	1.30	0.89, 1.81	-	-	-	-	-	-

CI: Lower and upper 95% confidence intervals. Node ID correspond to those presented in Fig 2.4

2.6. Discussion

2.6.1. Molecular phylogeny and taxonomic implications

This study represents the first densely sampled, strongly supported phylogenetic assessment of mainland African *Zosterops*. The results identify the presence of two major independent mainland African clades that both contain island radiations on nearby island archipelagos. Furthermore, results highlight significant non-monophyly of mainland Africa taxa, specifically *Z. poliogaster* and *Z. senegalensis*, with members of both species occurring in each major clade. These relationships support previous findings focusing on *Zosterops* island radiations (Warren *et al.* 2006; Melo *et al.* 2011) but which only included very limited mainland African sampling.

In the absence of molecular data, the various non-intergrading montane populations of *Z. poliogaster* are classified as subspecies of a wider species complex (Dickinson 2003). However, the extensive sampling in this study for five of the eight currently described *Z. poliogaster* subspecies strongly identifies this taxon as polyphyletic. Instead, strong support for the monophyly of individual subspecies indicates that the various non-intergrading montane populations should be considered as independent taxonomic units rather than intra-specific taxa. Further investigation using species delimitation methods (Wiens and Penkrot 2002; Pons *et al.* 2006; Knowles and Carstens 2007; Barraclough *et al.* 2009; Monaghan *et al.* 2009; Carstens and Dewey 2010; Leaché and Fujita 2010; Powell 2012) and involving subspecies absent from this study is required to accurately infer taxonomic boundaries.

The widespread taxon *Z. senegalensis* is also recovered as polyphyletic with the five subspecies sampled (14 currently recognised) falling into four distinct clades that are not related. The non-monophyly of *Z. senegalensis* suggests this group could be a cryptic species complex (Funk and Omland 2003). Denser sampling of subspecies within *Z. senegalensis* is needed to determine a more complete picture of intra-specific relationships.

This study also highlights discordance in the taxonomic treatment of *Z. abyssinicus,* revealing that two mainland African *Z. abyssinicus* subspecies (*Z. a. flavilateralis* and *Z. a. jubaensis*) are distinct from the insular member of *Z. a. socotranus*. Furthermore, results provide no support for division between *Z. a. flavilateralis* and *Z. a. jubaensis,* which brings into question their sub-specific

status. Further sampling of subspecies in Ethiopia and the Arabian Peninsula is required to fully resolve the systematic treatment of this group.

Finally, while two *Z. pallidus* (*Z. p. capensis* and *Z. p. virens*) subspecies form a strongly supported clade, the phylogenetic placement of the nominate subspecies (*Z. p. pallidus*) remains unclear. With no support for its placement as sister to the clade containing *Z. p. capensis* and *Z. p. virens*, these results support a previous taxonomic treatment (Moreau 1957), in which the western nominate group (*Z. p. pallidus*) and the eastern group (*Z. p. capensis*) were treated as separate species. Although the placement of *Z. p. pallidus* within clade A4 is unresolved, the non-monophyly of *Z. pallidus* demonstrated in this study is concordance with the recent molecular phylogeny of Oatley *et al.* (2012) that placed *Z. p. pallidus* as sister to *Z. senegalensis*. These findings add to existing questions (Melo *et al.* 2011) regarding the utility of traditional morphological characters used to delineate species within Zosteropidae and reinforce the need for complete systematic review of all African *Zosterops*.

2.6.2. Molecular phylogeny and models of speciation

The montane speciation model predicts that recently evolved montane populations should have similar ecological requirements and vagility, leading them to occupy congruent elevational distributions and habitat types that correspond with current and/or historical refugia (Roy 1997; Wiens and Donoghue 2004; Kozak and Wiens 2007; Fjeldså and Bowie 2008; Wiens *et al.* 2010). Phylogenetic results clearly indicate that this is not the case for *Z. poliogaster,* revealing that many endemic montane populations are more closely related to taxa with divergent habitat types, elevational distributions and dispersal abilities than they are to populations of restricted endemics that occur in neighbouring montane forest fragments (Fig. 2.3). This is in contrast to what has been reported for African bulbuls (Pycnonotidae: *Andropadus*), in which species and subspecies occupying montane forest fragments are recovered as a monophyletic group relative to species that occupy the dry and arid lowlands (Roy 1997). The phylogenetic placement of endemic montane forms, as sister to taxa with differing ecological requirements and vagility, provides clear evidence that *Z. poliogaster* subspecies are not relics of a previously widespread population as indicated in the current taxonomic arrangement (Fry *et al.* 2000; Dickinson 2003; van Balen 2008) and thus niche conservatism has

not played the primary role in divergence (Moritz *et al.* 2000; Kozak and Wiens 2007). Instead, results indicate that ancestral *Z. poliogaster* populations were adaptive, with niche divergence leading to aggregates of taxa with divergent habitat types, elevational distributions and dispersal abilities.

This is particularly evident for the endemic montane *Z. p. mbuluensis* (Chyulu Hills, Kenya) that is recovered as sister to two *Z. abyssinicus* subspecies (*Z. a. flavilateralis* and *Z. a. jubaensis*) that have a wide distribution throughout the dry and arid lowlands of Kenya and Ethiopia. Both the gradient speciation model and the vanishing refuge model have previously been used to explain the occurrence of sister taxa in adjacent but distinct habitats (Vanzolini and Williams 1981; Moritz *et al.* 2000; Ogden and Thorpe 2002; Hall 2005; Kozak and Wiens 2007). However in the absence of data on historical rate of gene flow, it is extremely difficult to distinguish between these two alternative hypotheses using phylogenetic inference.

Divergent selection is also demonstrated within two other mainland African sub-clades (Fig. 2.3, Nodes A4 and B2), yet the range of ecotypes within these clades makes historical relationships complex. Both clades (A4 and B2) contain two endemic montane *Z. poliogaster* subspecies that occur in neighbouring forest fragments. Despite the proximity of the forest fragments (<50km between Taita Hills/ S. Pare Mts and <100km between Mt Kulal/ N. Aberdares), *Z. poliogaster* populations in both clades are clearly divergent, conforming to the idea that lowland savannah habitat provides a barrier to gene flow causing divergence between isolated non-adaptive forms in neighbouring montane forest fragments (Fjeldså and Lovett 1997; Roy 1997; Fjeldså and Bowie 2008; Measey and Tolley 2011; Voelker *et al.* 2010). Despite this, in both sub-clades one of the *Z. poliogaster* subspecies is recovered as sister to a clade containing widely dispersed taxa (Fig. 2.3, Nodes A7 and B4).

Representatives of clades A7 and B4 differ from *Z. poliogaster* in both habitat type and elevational distribution. In clade A7 the two *Z. pallidus* subspecies (*Z. p. capensis* and *Z. p. virens)* have wide distributions, occurring in dry temperate grasslands of southern Africa, while in B4, *Z. s. senegalensis* has a wide distribution occurring in the moist and dry savannahs from Senegal to northwest Ethiopia. These clades also contain highland *Z. senegalensis* (*Z. s. jacksoni* and *Z. s. stierlingi*) that, like *Z. poliogaster,* occupy montane forest habitats throughout Kenya and Tanzania. However, the presence of highland *Z.*

senegalensis in multiple non-connected forest fragments indicates that, unlike *Z. poliogaster*, highland *Z. senegalensis* populations are less restricted by low dispersal abilities. Strong support for the basal placement of *Z. poliogaster* subspecies *Z. p. silvanus* and *Z. p. kulalensis* (restricted montane distributions) relative to clades A7 and B4 (divergent habitat types and/or wide distributions) challenges the montane speciation model, suggesting that ancestral lineages were in fact adaptive.

While the placements of clades A7 and B4 provide substantial support for mechanisms founded on niche divergence, they do not provide support for the gradient speciation model. Both *Z. pallidus* and *Z. s. senegalensis* have wide distributions occupying non-montane forest habitat. However, the restricted populations of *Z. p. winifredae* and *Z. p. kikuyuensis* are not contiguous with *Z. pallidus* and *Z. senegalensis* respectively and thus strong directional selection between habitat types along an altitudinal gradient is not reflected (Moritz *et al.* 2000; Kozak and Wiens 2007). Furthermore, the presence of highland *Z. senegalensis* forms (*Z. s. stierlingi and Z. s. jacksoni*) within clades A7 and B4 conflicts with the main predictions of the gradient speciation model that taxa should occur in distinct habitats that have elevationally non-overlapping geographical distributions (Moritz *et al*. 2000).

Instead, this study favours the vanishing refuge model to explain diversification within East African *Zosterops*. This model accounts for divergence between adjacent montane populations, which is interpreted as support for the theory that forest retraction served as a vicariant isolating mechanism for forest-adapted species that became geographically isolated in stable montane forest refugia (Fjeldså and Lovett 1997; Roy 1997; Fjeldså and Bowie 2008; Measey and Tolley 2011; Voelker *et al.* 2010). In addition, divergent selection toward tolerance of less favourable habitat as a result of habitat loss in less stable montane areas explains the sister relationship demonstrated between many endemic montane populations and clades containing taxa with non-montane habitat types. Under the assumption that a tolerance of arid conditions promotes habitat plasticity, directional selection of ancestral taxa would have led to wide ecological tolerances. This would have allowed for adaptation to arid environments during periods of extreme aridity with the subsequent re-colonisation of more favourable forested habitats when available.

2.6.3. Avian molecular clock

Broad application of 'traditional divergence rates', such as the avian molecular clock (2.1.%), has received widespread criticism (García-Moreno 2004; Lovette 2004; Ho *et al.* 2005; Perira and Baker 2006; Ho 2007). Previous studies indicate widespread variation in the rate of evolution among birds (Peterson 2006) and therefore the use of a standard clock rate across diverse avian taxa is questionable (Ho 2007). Mutation rate is also shown to vary depending on the taxonomic level with inter-specific comparisons giving lower rates relative to intra-specific comparisons (Lambert et al. 2002; Ho et al. 2005). Consequently, employment of the 2.1% interspecific rate in analyses that include intra-specific data is likely to produce overestimates of the ages of divergence events (Ho *et al.* 2005; Ho and Larson 2006).

The use of geological calibration, while often preferred, has marked weaknesses with divergence estimates being heavily dependent on the quality of calibration points available. Analyses often presume minimum error associated with geological ages of calibrations and do not take into account that lineage divergence may pre- or post-date the calibration set. Examining priors associated with calibrated nodes and using multiple calibration points is often used to try and increase the reliability of divergence estimates, however a lack of suitable calibration points means that this study was limited to the use of a single geological calibration. In spite of this, support for the use of the appearance of Grande Comore as a geological calibration to date divergence within African *Zosterops* comes from previous studies that demonstrate consistency in divergence estimate produced using independent calibrations (Moyle *et al.* 2009; Melo *et al.* 2011).

Using a different taxon set, additional genes, independent calibration points (New Georgia Group, Solomon Islands) and different analytical methods, Moyle *et al.* (2009) dated the divergence of the GGM clade between 0.89-1.35Ma. This is extremely similar to the estimate produce by the island-calibrated approach used in this study (0.7-1.47Ma). Even if the true rate of evolution in Zosteropidae was slower and similar to the more conservative 2.1% clock rate (Weir and Schluter 2008), then African *Zosterops* would still be estimated as a very recently diverged group (<5Ma). Based on the 2.1% clock rate approach, age estimates for the diversification of this group would coincide

with Pliocene tropical forest retraction (Hamilton and Taylor 1991; Feakins *et al.* 2005; Sepulchre *et al.* 2006), while divergence of Montane forms would coincide with earlier cycles of precessional-forced climatic variability that occurred during the late Pliocene/early Pleistocene (Trauth *et al.* 2007).

2.6.4. Evidence supporting the Pleistocene refuge model

The Pleistocene Refuge Hypothesis argues that dramatic changes in fauna composition during periods of reduced rainfall would have led to the temporary fractioning, and in some cases divergence, of previously widespread populations that became isolated in lowland forest refugia (Crowe and Crowe 1982; Mayr and O'Hara 1986; Diamond and Hamilton 2009). Previous avian studies addressing the montane speciation model have criticised the Pleistocene refuge hypothesis, arguing that divergence of many lowland species predates the Plio-Pleistocene (Amorim 1991; Hackett 1993; Fjeldså and Lovett 1997; Roy *et al.* 2001; Fjeldså and Bowie 2008).

However, results based on the calibrated approach indicate that this may not be the case with average estimates of divergence times for all African *Zosterops* (including lowland taxa) falling well within the Lower Pleistocene (Fig. 2.4., Table 3) Results based on the calibrated approach suggest that early divergence within this group took place during a phase of heightened aridity between 2.5 and 1.7Ma (deMenocal 1995; Trauth *et al.* 2007) (Fig. 2.4.). However, as previously discussed, discrepancies between the calibrated and clock rate approach mean that strict interpretation of results from either approach warrant caution. Nevertheless, even under the more conservative molecular clock approach several node estimates coincide with the Pleistocene and thus the role of Pleistocene refugia should not be altogether discounted.

2.6.5. Divergence of montane endemics

Results indicate that diversification of all endemic montane *Z. poliogaster* subspecies occurred within the last 0.88Ma (CL 95%, 0.61, 1.23). These age estimates coincide with a period of desiccation associated with the onset of arid conditions that occurred shortly after a period of extreme wetness (Trauth *et al.* 2007). Mean node ages suggest that divergence times of montane taxa were not contemporaneous: while early stages of dryness resulted in the independent diversifications of *Z. p. silvanus* and *Z. p. kulalensis* (0.88Ma); divergence of *Z.*

p. kikuyuensis, Z. p. mbuluensis and *Z. p. winifredae* occurred much later (0.63, 0.39 and 0.29Ma respectively). This relationship supports the vanishing refuge hypothesis, suggesting that later stages of aridity are associated with divergence between endemic montane forest specialists and clades containing widely dispersed lowland and/or highland taxa.

Under this model the ancestral population of clades containing widely dispersed lowland and/or highland taxa would have been a forest specialist. Following the assumption that increased aridity caused the contraction of forest habitat into montane forest refuges (Fjeldså 1994; Fjeldså Lovett 1997; Fjeldså and Bowie 2008; Roy *et al.* 2001; Voelker *et al.* 2010), this model predicts that the forest fragments occupied by these ancestral lineages would have been less stable than the forest fragments occupied by current *Z. poliogaster* subspecies. Periods of prolonged aridity would have resulted in refuges becoming too small to retain viable populations, leading to directional selection towards a tolerance of less favourable habitats (Vanzolini and Williams 1981; Moritz *et al.* 2000).

The lack of resolution in the spatio-temporal dynamics of key variables such as the structure and contiguity of rainforest habitat means that identifying the location of 'palaeoforest' fragments would be extremely difficult (Moritz *et al.* 2000). However, given the geographic proximity of *Z. poliogaster* subspecies in both Africa A and B, this study predicts that the ancestral populations of clades containing widely dispersed lowland and/or highland taxa would have occupied 'palaeoforest' fragments that were geographically close to forest fragments currently occupied by sister taxa.

2.7. Conclusions

In his review of African Zosteropidae, Moreau (1957) recognised the taxonomic complexities of this group. Moreau (1967) identified that the features used to determine relationships among African taxa may be problematic and advised others to 'deal with current taxonomy with great caution'. This study confirms that Moreau (1957) caution was not unjustified, with the phylogeny generated demonstrating significant non-monophyly of mainland African *Zosterops* species. Consequently, this study questions the utility of traditional characters, predominantly morphological, used to delineate species within Zosteropidae, with results indicating that the current taxonomic framework may have led to a severe underestimation of *Zosterops* diversity within mainland Africa. Denser sampling of *Zosterops* across continental Africa is necessary to determine a more comprehensive systematic framework, which would provide the basis for a complete systematic review of all mainland African taxa.

Phylogenetic analysis indicates that divergence leading to the current distribution of East African *Zosterops* is more complex that previously thought. The non-monophyly of *Z. poliogaster* indicates that the endemic montane populations of East Africa are not relics of a previously widespread population, as shown in African Bulbuls (Roy 1997), Akalats (Roy *et al* 2001) and Forest Robins (Voelker *et al.* 2010). This study therefore excludes the postulated montane speciation model in favour of the vanishing refuge model to explain lineage diversification in the focal group. Rather than being non-adaptive as predicted by the montane speciation model, this study indicates that ancestral *Zosterops* populations were in fact adaptive. Phylogenetic analysis identifies three key biotic diversification events within African *Zosterops,* where niche divergence has led to aggregates of taxa with divergent habitat types, elevational distributions and dispersal abilities. However, subsequent investigation into whether ancestral populations experienced severe bottlenecks with subsequent range expansion as a result of habitat loss is necessary.

Irrespective of method, divergence estimates recover African *Zosterops* as a very recently diverged group. Results indicate that the effect of climatic history on ancestral divergence within African *Zosterops* is not limited to divergence between montane endemics. Instead the unstable Plio-Pleistocene

African climate may have provided the primary driver for lineage diversification in all mainland African *Zosterops* lineages.

2.8. References

Amorim DS (1991) Refuge model simulations: testing the theory. *Revia Bras. Ent*, **35**, 803-815.

Barraclough TG, Hughes M, Ashford-Hodges N, Fujisawa T (2009) Inferring evolutionary significant units of bacterial diversity from broad environmental surveys of single-locus data. *Biology Letters*, **5**, 425-428.

Blackburn DC, Measley GJ (2009) Dispersal to or from an African biodiversity hotspot? *Molecular Ecology*, **18**, 1904-1915.

Carstens BC and Dewey TA (2010) Species delimitation using a combined coalescent and information-theoretic approach: an example from North American *Myotis* Bats. *Systematic Biology*, **59**, 400-414.

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

Crowe TM, Crowe AA (1982) Patterns of distribution, diversity and endemism in Afrotropical birds. *Journal of Zoology*, **198**, 417-442.

deMenocal PB (1995) Plio-Pleistocene African Climate. *Science*, **270**, 53-59.

Diamond AW, Hamilton AC (2009) The distribution of forest passerine birds and Quaternary climatic change in tropical Africa. *Journal of Zoology*, **191**, 379-402.

Dickinson EC (ed.) (2003) The Howard & Moore Complete Checklist of the Birds of the World, 3rd Edition. Christopher Helm, London.

Drummond AJ, Ho AYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, 0699-0710.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214,

Erixon P, Svennblad B, Britton T, Oxelman B (2003) Reliability of bayesian posterior probabilities and bootstraps frequencies in phylogenetics. *Systematic Biology*, **52**, 665-673.

Feakins SJ, deMenocal PB, Eglinton TI (2005) Biomarker records of late Neogene changes in northeast African vegetation. *Geology*, **33**, 977-980.

Fjeldså J (1994) Geographical patterns for relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodiversity and Conservation*, **3**, 207-226.

Fjeldså J, Lovett JC (1997) Geographical 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, Bowie RCK (2008) New perspectives on the origin and diversification of Africa's forest avifauna. *African Journal of Ecology*, **46**, 235-247.

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.

Fritz SA, JØnsson KA, Fjeldså J, Rahbek C (2011) Diversification and biogeographic patterns in four island radiations of passerine birds. *Evolution*, **66**, 179-190.

Fry CH, Keith S, Urban EK (Eds) (2000) The birds of Africa, Vol VI. London: Academic Press.

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

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

Hackett SJ (1993) Phylogenetic and biogeographic relationships in the neotropical genus *Gymnopithys* (Formicariidae). *The Wilson Bulletin*, **105**, 301-315.

Hall JPW (2005) Montane speciation patterns in *Ithomiola* butterflies (Lepidoptera: Riodinidae): are they consistently moving up in the world? *Proceedings of the Royal Society B*, **272**, 2457-2466.

Hamilton AC and Taylor D (1991) History of climate and forests in Tropical Africa during the last 8 million years. *Climatic Change*, **19**, 65-78.

Helm-Bychowski K, Cracraft J (1993) Recovering phylogenetic signal from DNA sequences: relationships within the Corvine assemblage (Class Aves) as inferred from complete sequences of the mitochondrial DNA cytochrome-*b* gene. *Molecular Biology and Evolution*, **10**, 1196-1214.

Ho SYW, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution*, **22**, 1561-1568.

Ho SYW, Larson G (2006). Molecular clocks: when times are a-changin'. *Trends in Genetics*, **22**, 79-83.

Ho SYW, Shapiro B, Phillips M, Cooper A, Drummond AJ (2007) Evidence for time dependency of molecular rate estimates. *Systematic Biology*, **56**, 515-522.

Huelsenbeck JP, Bull JJ, Cunningham CW (1996) Combining data in phylogenetic analysis. *Tree Reviews*, **11**, 152-157.

Huelsenbeck JP, Crandall KA (1997) Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review of Ecology and Systematics*, **28**, 437-466.

Huelsenbeck JP and Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics Applications Note*, **17**, 754-755.

Jetz W, Rahbek C, Colwell RK (2004) The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecology Letters*, **7**, 1180-1191.

Knowles LL, Carstens BC (2007) Delimiting species without monophyletic gene trees. *Systematic Biology*, **56**, 887-895.

Kozak KH, Wiens JJ (2007) Climatic zonation drives latitudinal variation in speciation mechanisms. *Proceedings of the Royal Society B*, **274**, 2995-3003.

Lambert DM, Ritchie PA, Millar CD, Holland B, Drummond AJ, Baroni C (2002) Rates of evolution in ancient DNA from Adélie penguins. *Science*, **295**, 2270-2273.

Lawon LP (2010) The discordance of diversification: evolution in the tropical-montane frogs of the Eastern Arc Mountains of Tanzania. *Molecular Ecology*, **19**, 4046-4060.

Leaché AD, Fujita MK (2010) Bayesian species delimitation in West Africa forest geckos (*Hemidactylus fasciatus*) *Proceedings of the Royal Society B,* **277**, 3071-3077.

Lerner HRL, Meyer M, James HF, Hofreiter M, Fleischer RC (2011) Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian Honeycreepers. *Current Biology*, **21**, 1838-1844

Linder HP, de Klerk HM, Born J, Burgess ND, Fjeldså J, Rahbek C (2012) The partitioning of Africa: statistically defined biogeographical regions in sub-Saharan Africa. *Journal of Biogeography*, **39**, 1189-1205.

Lovette IJ (2004) Mitochondrial dating and mixed support for the ‘‘2% rule’’ in birds. *The Auk,* **121**, 1-6.

Matthee CA, Tilbury CR, Townsend (2004) A phylogenetic review of the African leaf chameleons: genus *Rhampholeon* (Chamaeleonidae): the role of vicariance and climate change in speciation. *Proceedings of the Royal Society B*, **271**, 1967-1975.

Mayr E, O’Hara RJ (1986) The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. *Evolution*, **40**, 55-67.

Measey GJ, Tolley KA (2011) Sequential fragmentation of Pleistocene forests in an East Africa biodiversity hotspot: Chameleons as a model to track forest history. *PLoS ONE*, **6**, e26606.

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

Mittermeier RA, MyersN, Gil PR, Mittermeier CG (1999) Hotspots: Earth's biologically richest and most endangered terrestrial ecoregions (Cemex, Conservation International and Agrupacion Sierra Madre, Monterrey, Mexico).

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJG, Lees DC, Ranaivosolo R, Eggleton P, Barraclugh TG, Vogler AP (2009)

Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology*, **58**, 298-311.

Moreau RE (1957) Variation in the western Zosteropidae (Aves). *Bulletin of the British Museum,* **4**, 318–433.

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.

Moyle RG, Filardi CE, Smith CE, Diamond JC (2009) Explosive pleistocene diversification and hemispheric expansion of a ‘‘great speciator’’. *Proceedings of the National Academy of Science USA*, **106**, 1863–1868.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853-858

Nguembock B, Fjeldså J, Couloux A, Pasquet E (2009) Molecular phylogeny of Carduelinae (Aves, Passeriformes, Fringillidae) proves polyphyletic origin of the genera *Serinus* and *Cardelis* and suggests refined generic limits. *Molecular Phylogenetics and Evolution*, **51**, 169-181.

Nougier J, Cantagrel JM, Karche JP (1986) The Comores archipelago in the western Indian Ocean: volcanology, geochronology and geodynamic setting. *Journal of African Earth Sciences*, **5**, 135-145.

Oatley G, Voelker G, Crowe TM, Bowie RCK (2012) A multi-locus phylogeny reveals a complex pattern of diversification related to climate and habitat heterogeneity in southern African white-eyes. *Molecular Phylogenetics and Evolution*, **64**, 633-644.

Ogden R, Thorpe RS (2002) Molecular evidence for ecological speciation in tropical habitats. *Proceeding of the National Academy of Science*, **99**, 13612-13615.

Pereira SL, Baker AJ (2006). A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. *Molecular Biology and Evolution,* **23**, 1731-1740.

Peterson AT (2006) Application of molecular clocks in ornithology revisited. *Journal of Avian Biology*, **37**, 541-544.

Phillimore AB, Owens IPF, Black RA, Chittock J, Burke T, Clegg SM (2008) Complex patterns of genetic and phenotypic divergence in an island bird and the consequences for delimiting conservation units. *Molecular Ecology*, **17**, 2839-2853.

Plana V (2004) Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. *Philosophical Transactions of the Royal Society B*, **359**, 1585-1594.

Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, **55**, 595-609.

Powell JR (2012) Accounting for uncertainty in species delineation during the analysis of environmental DNA sequence data. *Methods in Ecology and Evolution*, **3**, 1-11.

Prager M, Johansson EIA, Andersson S (2008) Molecular phylogeny of the African windowbirds and bishops *Euplectes* spp. (Aves: Passeridae: *Ploceinae*). *Molecular Phylogenetics and Evolution*, **46**, 290-302.

Primmer CR, Borge T, Lindell J, Setre GP (2002) Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Molecular Ecology*, **11**, 603-612.

Rahbek C, Gotelli NJ, Colwell RK, Entsminger GL, Fernando T, Rangel TF, Graves GR (2007) Predicting continental-scale patterns of bird species richness with spatially explicit models. *Proceedings of the Royal Society B*, **274**, 165-174.

Roy MS (1997) Recent diversification in African greenbuls (Pycnonotidae: *Andropadus*) supports a montane speciation model. *Proceedings of the Royal Society B*, **264**, 1337-1344.

Roy MS, Sponer R, Fjeldså J (2001) Molecular systematics and evolutionary history of Akalats (Genus *Sheppardia*); A pre-Pleistocene radiation in a group of African forest birds. *Molecular Phylogenetics and Evolution*, **18**, 74-83.

Sepulchre P, Ramstein G, Fluteau F, Schuster M, Tiercelin JJ, Brunet M (2006) Tectonic uplift and eastern Africa aridification. *Science*, **313**, 1419-1423.

Shepard DB, Burbrink FT (2009) Phylogeographic and demographic effects of Pleistocene climatic fluctuations in a montane salamander, *Plethodon fourchensis*, *Molecular Ecology*, **18**, 2243-2262.

Swofford DL (2003) PAUP*: Phylogenetic analysis using parsimony (*and other methods) version 4 beta 10. Sinauer Associates, Sunderland, Massachusetts.

Thomas GH, Orme CDL, Davies RG, Olson VA, Bennett PM, Gaston KJ, Owens IPF, Blackburn TM (2008) Regional variation in the historical components of global avian species richness. *Global Ecology and Biogeography*, **17**, 340-351.

Trauth MH, Maslin MA, Deino AL, Strecker MR, Bergner AGN, Dühnforth M (2007) High- and Low- latitude forcing of Plio-Plistocene East African climate and human evolution. *Journal of Human Evolution*, **53**, 475-486.

van Balen S (2008) Family Zosteropidae (White-eyes). In Handbook of the birds of the world, Vol 13: 402–485. Del Hoyo J, Elliott A, Christie DA (Eds). Barcelona: Penduline Tits to Shrikes. Lynx Editions.

Vanzolini PE, Williams EE (1981) The vanishing refuge: a mechanism for ecogeographic speciation. *Papéis Avulsos de Zoologia*, **34**, 251-255

Voelker G, Outlaw RK, Bowie RCK (2010) Pliocene forest dynamics as a primary driver of African bird speciation. *Global Ecology and Biogeography*, **19**, 111-121.

Voje KL, Hemp C, Flagstad Ø, Setre G-P, Stenseth NC (2009) Climate change as an engine for speciation in flightless *Orthoptera* species inhabiting African mountains. *Molecular Ecology*, **18**, 93-108.

Warren B, Bermingham E, Prys-Jones RP, Thebauds C (2006) Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on the Indian Ocean islands. *Molecular Ecology,* **15**, 3769-3786.

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

Wiens JJ, Penkrot TA (2002) Delimiting species using DNA and morphological variation and discordant species limits in Spiny Lizards (*Sceloporus*). *Systematic Biology*, **51**, 69-91.

Wiens JJ, Donoghue MJ (2004) Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution*, **19**, 639-644.

Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Davies TJ, Grytnes J-A, Harrison SP, Hawkins BA, Holt RD, McCain CM, Stephens PR (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310-1324.

Xia X and Xie (2001) DAMBE: software package for data analysis in molecular biology and evolution. *The Journal of Heredity*, **92**, 371–373.

Xia X, Xie Z, Salemi M, Chen L, Wang Y (2003) An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution*, **26**, 1-7.

Yeung CKL, Lin R-C, Lei F, Robson C, Hung LM, Liang W, Zhou F, Han L, Li S-H, Yang X (2011) Beyond a morphological paradox: Complicated phylogenetic relationships of the Parrotbills (Paradoxornithidae, Aves). *Molecular Phylogenetics and Evolution*, **61**, 192-202.

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3		Cyt-b					
						1	2	1	2	3	4	5	6
Stachyris whiteheadi	Phillippines	AF094633	Sequence	NCBI		JN826960.1		NJ827210.1					
Z. palpebrosus													
palpebrosus	Nepal	RF2	Sequence	NCBI		DQ328448.1		DQ328348.1					
egregious	India	BMNH 1964.26.1	Sequence	NCBI		DQ328450.1		DQ328350.1					
Z. lateralis	Australia	SCB77	Sequence	NCBI		DQ328488.1		DQ328388.1					
Z. abyssinicus													
socotranus	Socota, Gulf of Aden	BW292	Blood	WARREN									
socotranus	Socota, Gulf of Aden	BW293	Blood	WARREN									
jubaensis	South Horr, Kenya	T60	Blood	COX									
jubaensis	South Horr, Kenya	T61	Blood	COX									
jubaensis	South Horr, Kenya	T65	Blood	COX									
jubaensis	South Horr, Kenya	T69	Blood	COX									
jubaensis	South Horr, Kenya	T70	Blood	COX									
jubaensis	South Horr, Kenya	T73	Blood	COX									
jubaensis	South Horr, Kenya	T76	Blood	COX									
jubaensis	South Horr, Kenya	T77	Blood	COX									
jubaensis	South Horr, Kenya	T84	Blood	COX									
jubaensis	South Horr, Kenya	T85	Blood	COX									
flavilateralis	Umani spring, Kenya	T4	Blood	COX									
flavilateralis	Umani spring, Kenya	T5	Blood	COX									
flavilateralis	Umani spring, Kenya	T11	Blood	COX									
flavilateralis	Umani spring, Kenya	T14	Blood	COX									
flavilateralis	Umani spring, Kenya	T15	Blood	COX									
flavilateralis	Umani spring, Kenya	T17	Blood	COX									

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3 1	ND3 2	Cyt-b 1	Cyt-b 2	Cyt-b 3	Cyt-b 4	Cyt-b 5	Cyt-b 6
flavilateralis	Umani spring, Kenya	T20	Blood	COX									
flavilateralis	Umani spring, Kenya	T21	Blood	COX									
flavilateralis	Umani spring, Kenya	T23	Blood	COX									
flavilateralis	Umani spring, Kenya	T30	Blood	COX									
flavilateralis	Nairobi, Kenya	RB3	Sequence	NCBI		DQ328439.1		DQ328339.1					
Z. pallidus													
pallidus	South Africa	AP50340	Blood	PFIAO									
capensis	Cape Town, South Africa	RB1	Sequence	NCBI									
capensis	Cape Town, South Africa	RB4	Sequence	NCBI									
virens	South Africa	AM36426	Blood	PFIAO									
virens	South Africa	AM36429	Blood	PFIAO									
virens	South Africa	AM36433	Blood	PFIAO									
Z. senegalensis													
senegalensis	Northern Region, Ghana	B39250	Sequence	NCBI		DQ328443.1		DQ328343.1					
senegalensis	Northern Region. Ghana	B39514	Sequence	NCBI		DQ328442.1		DQ328342.1					
stenocricotus	Mt Cameroon, Cameroon	STC01	Blood	CIBIO									
stenocricotus	Mt Cameroon, Cameroon	STC03	Blood	CIBIO									
jacksoni	Mathews Range, Kenya	BLS06	Blood	BORGHESIO									
jacksoni	Mt Nyiru, Kenya	BLS35	Blood	BORGHESIO									
jacksoni	Mt Nyiru, Kenya	BLS65	Blood	BORGHESIO									
jacksoni	Mathews Range, Kenya	BLS77	Blood	BORGHESIO									
jacksoni	Mathews Range, Kenya	BLS81	Blood	BORGHESIO									
jacksoni	Mt Nyiru, Kenya	ZMUC 131316	Blood	ZMUC									
jacksoni	Mt Nyiru, Kenya	ZMUC 131317	Blood	ZMUC									
jacksoni	Mt Nyiru, Kenya	ZMUC 131324	Blood	ZMUC									
jacksoni	Mt Nyiru, Kenya	ZMUC 131325	Blood	ZMUC									

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3 1	ND3 2	Cyt-b 1	Cyt-b 2	Cyt-b 3	Cyt-b 4	Cyt-b 5	Cyt-b 6
jacksoni	Mt Nyiru, Kenya	ZMUC 131331	Blood	ZMUC									
jacksoni	N. Aberdares, Kenya	ZMUC 146780	Blood	ZMUC									
jacksoni	N. Aberdares, Kenya	ZMUC 146784	Blood	ZMUC									
jacksoni	N. Aberdares, Kenya	ZMUC 146785	Blood	ZMUC									
jacksoni	N. Aberdares, Kenya	ZMUC 146786	Blood	ZMUC									
jacksoni	Kakamega, Kenya	T41	Blood	COX									
jacksoni	Kakamega, Kenya	T42	Blood	COX									
jacksoni	Kakamega, Kenya	T49	Blood	COX									
jacksoni	Kakamega, Kenya	T50	Blood	COX									
jacksoni	Kakamega, Kenya	T51	Blood	COX									
jacksoni	Kakamega, Kenya	T52	Blood	COX									
jacksoni	Kakamega, Kenya	T53	Blood	COX									
jacksoni	Kakamega, Kenya	T54	Blood	COX									
stierlingi	Tabora, Tanzania	ZMUC 145467	Blood	ZMUC									
stierlingi	Udzungwa Mts, Tanzania	ZMUC 140192	Blood	ZMUC									
stierlingi	Poroto Mts, Tanzania	ZMUC 142605	Blood	ZMUC									
stierlingi	Poroto Mts, Tanzania	ZMUC 142607	Blood	ZMUC									
stierlingi	W. Usambara Mts, Tanzania	ZMUC 129298	Blood	ZMUC									
stierlingi	W. Usambara Mts, Tanzania	ZMUC 129289	Blood	ZMUC									
stierlingi	Uhafiwa, Tanzania	O8255	Sequence	NCBI		DQ328438.1		DQ328338.1					
unknown	DRCongo	ZMUC 128660	Blood	ZMUC									
unknown	DRCongo	ZMUC 128632	Blood	ZMUC									
unknown	DRCongo	ZMUC 128658	Blood	ZMUC									
Z. poliogaster													
kulalensis	Mt Kulal, Kenya	K30	Blood	COX									
kulalensis	Mt Kulal, Kenya	K31	Blood	COX									

axon	Locality	Sample code	Sample type	Source	TGFß2	ND3 1	ND3 2	Cyt-b 1	Cyt-b 2	Cyt-b 3	Cyt-b 4	Cyt-b 5	Cyt-b 6
ılalensis	Mt Kulal, Kenya	K33	Blood	COX									
ılalensis	Mt Kulal, Kenya	K34	Blood	COX									
ılalensis	Mt Kulal, Kenya	K35	Blood	COX									
ılalensis	Mt Kulal, Kenya	K37	Blood	COX									
ılalensis	Mt Kulal, Kenya	K38	Blood	COX									
ılalensis	Mt Kulal, Kenya	K39	Blood	COX									
ılalensis	Mt Kulal, Kenya	K41	Blood	COX									
ılalensis	Mt Kulal, Kenya	K42	Blood	COX									
ılalensis	Mt Kulal, Kenya	2MK3	Blood	HABEL									
ılalensis	Mt Kulal, Kenya	2MK7	Blood	HABEL									
ılalensis	Mt Kulal, Kenya	2MK8	Blood	HABEL									
ılalensis	Mt Kulal, Kenya	2MK9	Blood	HABEL									
ılalensis	Mt Kulal, Kenya	2MK10	Blood	HABEL									
kuyuensis	Mt Kenya, Kenya	MK1	Blood	HABEL									
kuyuensis	Mt Kenya, Kenya	MK2	Blood	HABEL									
kuyuensis	Mt Kenya, Kenya	MK3	Blood	HABEL									
kuyuensis	Mt Kenya, Kenya	MK6	Blood	HABEL									
kuyuensis	Mt Kenya, Kenya	MK7	Blood	HABEL									
kuyuensis	Aberdares Range, Kenya	AB2	Blood	HABEL									
kuyuensis	Aberdares Range, Kenya	AB4	Blood	HABEL									
kuyuensis	Aberdares Range, Kenya	AB10	Blood	HABEL									
kuyuensis	Aberdares Range, Kenya	AB11	Blood	HABEL									
kuyuensis	Aberdares Range, Kenya	AB12	Blood	HABEL									
kuyuensis	Aberdares Range, Kenya	AB13	Blood	HABEL									
kuyuensis	Aberdares Range, Kenya	AB20	Blood	HABEL									
kuyuensis	Mt Kenya, Kenya	RB2	Sequence	NCBI		DQ328447.1			DQ328247.1				

axon	Locality	Sample code	Sample type	Source	TGFß2	ND3 1	ND3 2	Cyt-b 1	Cyt-b 2	Cyt-b 3	Cyt-b 4	Cyt-b 5	Cyt-b 6
vanus	Mbololo, Taita Hills, Kenya	O8580	Sequence	NCBI		DQ328437.1		DQ328337.1					
vanus	Taita Hills, Kenya	TH 1 2	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 1 3	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 1 10	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 1 21	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 2 3	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 2 10	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 2 12	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 2 15	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 3 12	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 3 13	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 3 20	Blood	HABEL									
vanus	Taita Hills, Kenya	TH 3 21	Blood	HABEL									
uluensis	Chyulu Hills, Kenya	CH1	Blood	HABEL									
uluensis	Chyulu Hills, Kenya	CH3	Blood	HABEL									
uluensis	Chyulu Hills, Kenya	CH6	Blood	HABEL									
uluensis	Chyulu Hills, Kenya	CH7	Blood	HABEL									
uluensis	Chyulu Hills, Kenya	CH8	Blood	HABEL									
uluensis	Chyulu Hills, Kenya	CH9	Blood	HABEL									
uluensis	Chyulu Hills, Kenya	CH11	Blood	HABEL									
uluensis	Chyulu Hills, Kenya	2CH10	Blood	HABEL									
uluensis	Chyulu Hills, Kenya	2CH12	Blood	HABEL									
uluensis	Chyulu Hills, Kenya	2CH13	Blood	HABEL									
ifredae	S. Pare Mts, Tanzania	O5899	Sequence	NCBI		DQ328436.1		DQ328336.1					
borbonicus													
rbonicus	La Réunion, Mascarenes	BWM46	Sequence	NCBI		DQ328497.1		DQ328397.1					

kon	Locality	Sample code	Sample type	Source	TGFß2	ND3 1	ND3 2	Cyt-b 1	Cyt-b 2	Cyt-b 3	Cyt-b 4	Cyt-b 5	Cyt-b 6
onicus	La Réunion, Mascarenes	BWM47	Sequence	NCBI		DQ328496.1		DQ328396.1					
onicus	La Réunion, Mascarenes	BWM54	Blood	WARREN									
ritianus	Mauritus, Mascarenes	BWM17	Blood	WARREN									
ritianus	Mauritus, Mascarenes	BWM25	Blood	WARREN									
ritianus	Mauritus, Mascarenes	BWM24	Blood	WARREN									
icedulinus													
dulinus	Annobón, Gulf of Guinea	FIP002	Sequence	NCBI		GU827278.1		GU827230.1					
	São Tomé, Gulf of Guinea	FIS003	Sequence	NCBI		GU827280.1		GU827232.1					
riseovirescens	Príncipe, Gulf of Guinea	GRI002	Sequence	NCBI		GU827282.1		GU827234.1					
naderaspatanus													
bransis	Aldabra, Seychelles	BW301	Sequence	NCBI		DQ328487.1		DQ328387.1					
leraspatanus	Mt Ankaratre, Madagascar	BW429	Blood	WARREN									
leraspatanus	Mt Ankaratre, Madagascar	BW445	Blood	WARREN									
i	Grande Comore, Comoros	BW146	Sequence	NCBI		DQ328476.1		DQ328376.1					
nodestus	Conception, Seychelles	BW344	Sequence	NCBI		DQ328463.1		DQ328363.1					
runneus	Bioko, Gulf of Guinea	BRU003	Sequence	NCBI		GU827268.1		GU827220.1					
nelanocephalus	Mt Cameroon, Cameroon	MEL002	Sequence	NCBI		GU827266.1		GU827218.1					
ugubris	São Tomé, Gulf of Guinea	LUG001	Blood	CIBIO									
ugubris	São Tomé, Gulf of Guinea	LUG042	Blood	CIBIO									

Chapter 3

The phylogenetic utility of Amplified Fragment Length Polymorphisms (AFLPs) in resolving relationships within East African *Zosterops*

Table of Contents

3.1 Abstract

Introduction: Estimating relationships from ncDNA sequence data can be particularly problematic for groups that have diversified relatively recently. The genus *Zosterops* (white-eyes) demonstrates the complexity involved in uncovering relationships within very recently diverged groups, with few appropriate ncDNA sequence markers available for phylogenetic construction. This study investigates the utility of AFLP characters in resolving relationships between East Africa *Zosterops* to address whether the non-monophyly of montane endemics, as demonstrated in the mtDNA phylogeny of Chapter 2, is congruent with the relationships inferred from AFLP data.

Methodology: In total 15 primer combinations were used to generate AFLP profiles for 92 *Zosterops* samples. MrBayes was used to construct a topology and to assess support for phylogenetic groupings. Bayesian hypothesis testing was used to investigate support for phylogenetic hypotheses surrounding the taxonomic validity of *Z. poliogaster*.

Results: Resolution of relationships across the AFLP phylogeny is generally poor. While endemic montane *Z. poliogaster* subspecies form independently well-supported clades there is no support for the broader clades recovered in the mtDNA phylogeny of Chapter 2. Bayesian hypothesis testing failed to provide support for the nine constrained topologies tested. Consequently this study was unable to confirm or reject the non-monophyly of East African montane *Zosterops*.

Discussion: Results are interpreted to suggest that the dramatically lower performance of AFLP analyses with respect to the mtDNA phylogeny generated in Chapter 2 is likely due to the low information content of the AFLP matrix generated. This study examines the various properties of AFLPs that may have limited the quantity and quality of data contained within the fragments generated and highlights: fragment length collision; fragment length co-dominance; co-dominant noise; and a predominance of private alleles as key factors.

3.2. Introduction

3.2.1. Molecular markers

In order to obtain an accurate assessment of genetic diversity to address questions regarding phylogenetic relationships, it is important to utilise the most appropriate marker for the research question. No single molecular technique is universally ideal, with each available technique exhibiting both strengths and weaknesses. Recent decades have seen extensive use of sequence-based data for phylogenetic reconstruction, with mitochondrial sequence data (mtDNA) being the most widely used genetic marker for phylogenetic inference (Moritz *et al.* 1987; Herbert *et al.* 2003; Ballard and Whitlock 2004; Brito and Edwards 2008). The widespread use of mtDNA can be attributed to numerous factors. When compared to nuclear sequence data (ncDNA), mtDNA genes have no recombination, higher mutation rates and a smaller effective population size (Ballard and Whitlock 2004; Brito and Edwards 2008) giving mtDNA a comparatively higher rate of evolution and relatively rapid coalescent times.

While the general utility of mtDNA markers for phylogenetic and phylogeographic studies is well established, the use of mtDNA genes has marked weaknesses (Knowles and Maddison 2002; Ballard and Whitlock 2004; Brito and Edwards 2008). The maternal inheritance of mtDNA means that phylogenetic reconstructions based solely on mtDNA genes only reflect the maternal lineage (Ballard and Whitlock, 2004). In groups that demonstrate sex-biased dispersal (i.e. female phylopatry), mtDNA phylogenies fail to take into account male gene-flow dynamics (Ruppell *et al.* 2003; Wilder *et al.* 2004). Furthermore, mtDNA is shown to be more readily affected by the interspecific hybridisation than ncDNA (e.g. Chan and Levin 2005; Linnen and Farrell 2007), which can result in inconsistencies between topologies produced by mtDNA and ncDNA datasets.

Conflicts between phylogenies produced from mtDNA and ncDNA sequence data can be seen throughout the literature (Sota *et al.* 2001; Spinks and Shaffer 2009; Jackson and Austin 2010; McKay and Zink 2010; Joyce *et al.* 2011; Yeung *et al.* 2011) and have been attributed to mechanisms including incomplete lineage sorting, introgression, hybridisation and a generally lower rate of mutation in the nuclear genes compared to the mitochondrial genome

(Takahashi *et al.* 2001; Sanderson and Shaffer 2002; Shaw 2002; Funk and Omland 2003; Ballard and Whitlock 2004; McKay and Zink, 2010).

The last decade has seen an increasing awareness that inference based on mtDNA alone will not always be sufficient to resolve the species tree (Doyle 1992; Funk and Omland, 2003; Chan and Levin 2005; Brito and Edwards 2008; Spinks and Shaffer, 2009). While recent years have seen an increasing tendency to include ncDNA when generating species-level phylogenies (García-Moreno *et al.* 2003; Beltrán *et al.* 2007; Alfaro *et al.* 2008; Hugall *et al.* 2008), the comparatively lower rate of evolution of ncDNA genes can often limit the usefulness of ncDNA sequence data. For groups that have diverged relatively recently (e.g. island radiations) the use of ncDNA sequence data can be particularly problematic with limited sequence variation leading to a lack of phylogenetic signal. Many studies of more recently diverged groups are therefore still heavily reliant on mtDNA sequence data (e.g., Warren *et al.* 2003; Arbogast *et al.* 2006; Barluenga *et al.* 2006; Dasmahapatra and Mallet 2006; Warren *et al.* 2006; Jackson and Austin 2010; Joyce *et al.* 2011; Melo *et al.* 2011).

Obtaining a nuclear assessment of genetic diversity that resolves relationships at more shallow taxonomic levels requires a large number of independent ncDNA loci (Shaffer and Thomson 2007; Brito and Edwards 2008). In situations where ncDNA sequence data is limited by a lack of available nuclear sequence markers, researchers have turned to other nuclear marker systems such as microsatellites or simple sequence repeats (SSR: Richard and Thorpe 2001), single nucleotide polymorphisms (SNP: Carstens and Knowles 2007), restriction fragment length polymorphisms (RFLP; Mori *et al.* 1997), inter simple sequence repeats (ISSR: Al-Daoude *et al.* 2012), random amplified polymorphic DNA (RAPD; Al-Daoude *et al.* 2012) and amplified fragment length polymorphisms (AFLP; Koopman 2005).

These nuclear marker systems differ in the amount and quality of information obtained, with some approaches allowing the examination of a few single-locus markers, while others allow for the simultaneous investigation of multiple loci (Gerber *et al.* 2000; Saliba-Colombani 2000; Sunnucks 2000; Campbell *et al.* 2003; Nybom 2004; Mendelson and Shaw 2005; Meudt and Clarke 2007). While these marker systems will all reflect differences or changes

in the nuclear genome, the appropriateness of each marker changes in relation to the research question and resources available (Sunnucks 2000).

Choosing the most appropriate marker for a given research question is often based on numerous factors, and can frequently involve a trade-off between precision and convenience (Mueller and Wolfenbarger 1999). One manifestation of this is the dichotomy between multi-locus approaches (i.e. AFLP, ISSR, RAPD and RFLP) and single-locus techniques (i.e. SNPs, RAPDs and ncDNA sequences). While single-locus techniques are often more informative than multi-locus approaches, providing information about both alleles present at a given loci (co-dominant marker systems), they can be technically demanding and often require prior knowledge of target regions (Sunnucks 2000; Mariette *et al.* 2002; Bensch and Åkesson 2005).

In contrast multi-locus approaches often require little or no prior sequence information, making them an attractive marker system in understudied groups where there is often limited prior knowledge of the nuclear genome (Bensch and Åkesson 2005; Mendelson and Shaw 2005). While multi-locus approaches can provide a wide genomic assessment of genetic variability, the dominant nature of such marker systems limits them to only reporting the presence or absence of a given allele (Sunnucks 2000).

3.2.2. Multi-locus marker systems

In instances where there is limited prior knowledge of the nuclear genome, multi-locus approaches, specifically AFLPs can be an attractive alternative to many co-dominant markers such as multi-gene DNA sequencing, microsatellites and SNPs (Jones *et al.* 1997; Mori *et al.* 1997; Albertson *et al.* 1999; Mueller and Wolfenbarger 1999; Gerber *et al.* 2000; Sunnucks 2000; Belaj *et al.* 2003; Campbell *et al.* 2003; Bensch and Åkesson 2005; Al-Daoude *et al.* 2012). In contrast to many co-dominant marker systems that require extensive screening of the genome for polymorphic regions, multi-locus dominant marker systems (ISSRs, RAPDs, RFLPs and AFLPs) require minimal prior genomic knowledge of the study group for primer design (Vos *et al.* 1995; Sunnucks 2000; Bensch and Åkesson 2005; Meudt and Clarke 2007).

Although dominant marker systems are all similar in that they are PCR-based techniques that use primers to amplify previously uncharacterised DNA fragments, they all vary in respect to their data quality, genetic variability and

discriminatory power (Meudt and Clarke 2007). While dominant marker systems have been shown to demonstrate similar patterns of genetic distance and informativeness (Belaj *et al.* 2003; Nybom 2004), several studies demonstrate that AFLP out-perform other dominant marker systems with respect to their higher specificity and reproducibility (Jones *et al.* 1997; Savelkoul *et al.* 1999; Belaj *et al.* 2003). As a result, researchers have increasingly turned to AFLPs in an attempt to obtain high-resolution investigation of relationships in recently evolved and/or non-model groups (Sullivan *et al.* 2004; Mendelson and Shaw 2005; Dasmahapatra *et al.* 2009; Quek *et al.* 2010; Smith *et al.* 2011).

Although the use of SNPs (co-dominant marker system) in non-model systems has been demonstrated (Emerson *et al.* 2010; Wagner *et al.* 2012), the absence of prior sequence information requires an additional data generation stage from which variable sites (SNPs) can be screened (Baird *et al.* 2008). This involves the use of high-throughput sequencing of restriction-site-associated DNA tags (RAD tags) that are subsequently screened to identify SNP sites. While the information extracted per locus is higher in SNPs compared to alternative dominant marker systems the generation of prior sequence information and the subsequent screening process can be extremely costly and labour intensive.

3.2.3. Amplified Fragment Length Polymorphism's (AFLP)

The AFLP method is a PCR based methodology that combines the strengths of RFLP and RAPD, and was first developed by Vos *et al.* (1995). It is a selective method that amplifies subsets of restriction fragments, resulting in a unique and reproducible fingerprint (or profile) for each individual (Mueller and Wolfenbarger 1999). Although there is a tendency for AFLP fragments to be concentrated around centromeric regions (Saliba-Colombani *et al.* 2000), their genome-wide distribution is thought to give a more complete picture of whole-genome diversity relative to other markers systems (sequence data or microsatellites) that concentrate on comparatively smaller regions of the genome (Meudt and Clarke 2007).

AFLPs have a broad range of applications and have been utilised in a wide range of disciplines, including linkage mapping (Alonso-Blanco *et al.* 1998; Saliba-Colombani *et al.* 2000), parentage analysis (Gerber *et al.* 2000), measuring genetic diversity (Nybom 2004), identifying hybrids (Goldman *et al.*

2004), population genetics (Parsons and Shaw 2001; Barluenga *et al.* 2006; Mila *et al.* 2010) population assignment (Campbell *et al.* 2003) and phylogenetic reconstruction (Albertson *et al.* 1999**;** Sullivan *et al.* 2004; Dasmahapatra *et al.* 2009; Genner and Tunner 2012).

AFLPs require a relatively short start-up time and have the ability to generate numerous (>1000) genome-wide di-allelic loci at moderate costs (compared to SNPs for example). While the dominant nature of AFLPs means that information extracted per locus is less informative when compared to co-dominant marker systems, it is argued that AFLPs derive their statistical power from the sheer number of loci that can be generated (Sunnucks 2000; Belaj *et al.* 2003; Bensch and Åkesson 2005; Meudt and Clarke 2007).

Despite their apparent usefulness however, there has been a non-random distribution in the utilisation of AFLPs relative to organism group**.** While AFLPs have been the choice method for many studies of plants, fungi, and bacteria, animal researchers have been relatively slow in embracing this method (Bensch and Åkesson 2005). In the past, micro-satellites have often prevailed as the molecular marker used by animal researchers (Brito and Edwards 2008), despite problems associated with isolation and transferability of markers between species.

The discovery that numerous AFLP data sets contain phylogenetic signal (Koopman 2005; Mendelson and Shaw 2005) has stimulated their use as a source of genetic information for phylogenetic reconstruction, particularly among closely related species or genera (Barluenga *et al.* 2006; Genner and Tunner 2010; Quek *et al.* 2010). As a result of their wide genomic distribution, AFLP markers are likely to uncover rare genetic variation in closely related groups and have been used to infer relationships in groups, which have previously been impossible to resolve with morphological or other molecular markers (Mendelson and Shaw 2005; Quek *et al.* 2010).

There is a general consensus that the usefulness of AFLP markers for phylogenetic and phylogeographic studies relates more to grouping of closely related lineages rather than distantly related taxa (Mueller and Wolfenbarger 1999; Bensch and Åkesson 2005; Meudt and Clarke 2007). While AFLPs have provided support for deep relationships in some groups (e.g., Dasmahapatra *et al.* 2009; Smith *et al.* 2011), there is significant evidence that demonstrates that phylogenetic inference becomes more problematic at higher taxonomic levels,

with homoplasy of AFLP profiles increasing significantly in comparisons of distantly related taxa (Koopman 2005; García-Pereira *et al.* 2009).

3.2.4. Study system - East African *Zosterops*

The genus *Zosterops* demonstrates the complexity involved in uncovering relationships within very recently diverged groups. With few ncDNA sequence markers available, attempts to uncover relationships within this highly speciose group have relied largely on mtDNA sequence data (Slikas *et al.* 2000; Warren *et al.* 2006; Phillimore *et al.* 2008; Melo *et al.* 2011) or microsatellites (Estoup and Clegg 2003; Clegg and Phillimore 2010) although Moyle *et al.* (2009) and Oatley *et al.* (2012) have used multi-locus approaches.

Chapter 2 attempted to generate a multi-locus species-level phylogeny for East African *Zosterops.* However, investigation into the potential use of the nuclear gene TGFß2 resulted in no informative sites, and the results therefore relied exclusively on mtDNA sequence data. While analyses resulted in a well-resolved strongly supported topology, phylogenetic relationships demonstrated significant discordance from the currently accepted taxonomy (predominantly based on phenotypic and ecological characters). Results showed that none of the forms previously identified as putative species corresponded to monophyletic groups. Instead results indicated that collection site rather than phenotype was often a better predictor of haplotype affinities.

The non-monophyly of taxa was interpreted to suggest that despite strong ecological and morphological similarities, in many cases montane *Z. poliogaster* subspecies are actually more closely related to lowland taxa than they are to neighbouring highland *Z. poliogaster* populations. Although the reciprocal monophyly of lowland and highland taxa provides evidence of lineage sorting (at least at the level of mtDNA loci), the polyphyletic placement of *Z. poliogaster* might also reflect incomplete lineage sorting or introgressive hybridisation as demonstrated in other radiations (Takahashi *et al.* 2001; Shaw 2002; Sullivan *et al.* 2004; Spinks and Shaffer 2009; Genner and Tunner 2012). Given that no independent estimate of phylogeny was available for this study, the possibility of local introgression of the mitochondrial genome across forms cannot be ruled out. Thus a nuclear phylogeny is still required to assess whether the non-monophyly of montane endemics is reflected across the genome.

3.3 Aims

Using amplified fragment length polymorphisms (AFLPs), the primary aim of this chapter is to obtain a nuclear assessment for genetic relationships within the East African *Zosterops*. In particular, this work aims to address whether the non-monophyly of montane endemics, as demonstrated in the mtDNA phylogeny of Chapter 2, is congruent to the relationships inferred from nuclear markers. If *Z. poliogaster* subspecies are more closely related to non-montane forms based on AFLP markers, this would support convergent evolution of the *Z. poliogaster* phenotype. Alternatively, if *Z. poliogaster* is recovered as monophyletic, indicating a single radiation of montane forms, results would instead provide support for either ancient mtDNA introgression or the retention of ancestral mtDNA polymorphism through incomplete lineage sorting.

3.4 Materials and methods

3.4.1. Taxonomic sampling

In total 116 DNA extracts were obtained for this study. The majority of DNA extracts (86) were obtained from blood samples that were taken during multiple collecting trips to Kenya between 2008 and 2010. Kenya is occupied by all three East African *Zosterops* species (*Z. poliogaster*, *Z. senegalensis* and *Z. abyssinicus*), including four of the eight currently recognised *Z. poliogaster* subspecies, one of the two East African *Z. senegalensis* subspecies (14 recognised across Africa) and two of the four mainland *Z. abyssinicus* subspecies (six subspecies in total). Specifically, this study obtained 49 DNA extracts for *Z. poliogaster* subspecies representing the four allopatric Kenyan subspecies: *Z. p. kulalensis* (Mt Kulal); *Z. p. kikuyuensis* (Aberdares range)*; Z. p. silvanus* (Taita Hills); and *Z. p. mbuluensis* (Chyulu Hills). Due to limited sampling outside Kenya, this study was unable to acquire samples for *Z. p. winifredae* and *Z. eurycricotus* from northern Tanzania or *Z. p. poliogaster* and *Z. p. kaffensis* from the Ethiopian highlands.

For the two highland *Z. senegalensis* subspecies that have fragmented distributions throughout Kenya and Tanzania, this study obtained 28 samples representing four populations of *Z. s. jacksoni* from across Kenya and three populations of *Z. s. stierlingi* from the highlands of Tanzania. Samples of *Z. s. jacksoni* from western Kenya were collected during fieldwork in 2011, while *Z. s. jacksoni* samples from the Kenyan highlands and Tanzanian *Z. s. stierlingi* samples were collected by various groups between 2000 and 2008. An additional five *Z. senegalensis* samples were also acquired from outside East Africa, representing three samples from DRC (not identified to sub-specific level) and two samples from Cameroon (*Z. s. stenocricotus*).

For the lowland species *Z. abyssinicus,* this study obtained DNA extracts (20 samples) for both Kenyan subspecies (*Z. a. jubaensis* and *Z. a. flavilateralis*) that have a parapatric distribution from southern Ethiopian through to northern Tanzania. It was not possible to obtain samples for *Z. a. abyssinicus* and *Z. a. omoensis* from the northern part of the range (Ethiopia, Sudan and Eritrea), although two samples were obtained for the insular subspecies *Z. a. socotranus.* An additional five samples were acquired for the southern Africa species *Z. pallidus*, in addition to seven samples representing *Zosterops*

species in the Gulf of Guinea and the Indian Ocean island systems. Details of voucher numbers, collection dates and sampling localities are listed in Appendix II.

3.4.2. DNA quality screening

AFLP-PCR is more sensitive to DNA quality and the presence of residual inhibitors in the DNA extract, than standard PCR-based applications (Vos *et al.* 1995: Bensch and Åkesson 2005). Poor quality DNA extracts can exhibit significant DNA degradation. The fragmented nature of degraded DNA (non-restriction fragments) can cause significant problems for the interpretation and analysis of AFLP datasets, with the presence of non-restriction fragments likely to increase levels of homoplasy (Bensch and Åkesson 2005). the presence of PCR-inhibitors in DNA extracts can indirectly affect AFLP profiles by reducing amplification efficiency and thus fragment generation between samples (Vos *et al.* 1995; Savelkoul *et al.* 1999; Meudt and Clarke 2007). A successful AFLP methodology therefore requires extracts that have high yields (~100 ng/µl) of good quality (non-degraded) DNA that is free of contaminants. All DNA extracts were screened using a Nanodrop spectrophotometer (Thermo Scientific NanoDrop 8000 V2.0) to quantify DNA concentration and to test for the presence of PCR inhibitors. In addition, DNA extracts were electrophoresed on a 1.5% agarose gel and assessed against Hyperladder I (Bioline) to obtain a measure of DNA quality. Only non-degraded, high molecular weight DNA extracts were used to generate AFLP profiles.

3.4.3. Generation of AFLP fragments

AFLP profiles were generated following the AFLP procedure of Vos *et al.* (1995), with certain modifications for fluorescent primers as detailed in Huang and Sun (1999) (Fig 3.1). Digestion of genomic DNA with restriction enzymes (Fig. 3.1, Step 1) was performed in 10µl volumes consisting of 5µl total genomic DNA (~10ng/ µl), 0.05µl of both the *MseI* (1 unit) and *EcoRI* (5 units) restriction enzymes, 1µl of both *MseI* and *EcoRI* restriction buffers (New England Biolabs), 1µl of BSA (10 mg/ml) and 2µl ddH_2O. Samples were incubated for 3 hours at 37˚C. Double stranded adapters were constructed from complementary single-stranded oligonucleotides (Table 3.1) heated to 95˚C for 10 min and then left to cool at room temperature. The ligation of the double stranded adapters (EcoRI

and MseI) to the end of the restriction fragments (Fig. 3.1, Step 2) was performed in 20µl volumes by adding 6.8µl ddH_2O, 1µl T4 DNA ligase buffer (New England BioLabs), 0.2µl of T4 DNA ligase (400 units), and 1µl of each adaptor to the restriction product (10µM) and incubating at 16˚C for 16 hours.

Following standard protocols (Vos *et al.* 1995), selective PCR was carried out in two stages. A sub-sample of all restriction fragments was obtained through a pre-selective amplification (Fig. 3.1, Step 3), followed by 15 selective amplifications (Fig. 3.1, Step 4) with each EcoRI primer labelled with either FAM ($+NNN^{*FAM}$) NED ($+NNN^{*NED}$) or HEX ($+NNN^{*HEX}$) fluorescent dyes. Pre-selective primers consisted of the adapter primer sequence with a single selective nucleotide (+N) at the 3' end. Selective amplifications were performed with primers containing the pre-amplification primer sequence with an additional 2 selective nucleotides at the 3' end, giving a 3 base pair extension (+NNN) from the original adapter sequence.

Pre-selective amplifications (Fig. 3.1, Step 3) were performed with 4µl of ligation product, 6.72µl ddH_2O, 2µl 10X PCR buffer (Bioline), 1µl $MgCl_2$ (25µM), 2µl dNTPs (2.5 µM of each), 1µl BSA (10 mg/ml), 0.6µl of each pre-selective primer (10µM), and 0.08µl of *Taq* DNA polymerase (0.4 units). PCR cycling parameters were a preliminary 72˚C extension for 2 minutes followed by 20 cycles of 20 seconds at 94˚C, 30 seconds at 56˚C and 2 minutes at 72˚C, with a final hold at 60˚C for 30 minutes. Following pre-amplification PCR products were diluted 1:5 with ddH_2O.

Selective amplification reactions (Fig. 3.1, Step 4) were performed in 10µl reaction volumes adding 5µl of the diluted (1:5) pre-selective PCR product to 1.04µl ddH_2O, 2.0µl dNTPs (2.5µM of each), 1µl 10 X PCR buffer (Bioline), 0.3µl MgCl2 (25µM), 0.06µl *Taq* DNA polymerase (0.3 units) and 0.3µl of both MseI and EcoRI selective amplification primers (10µM).

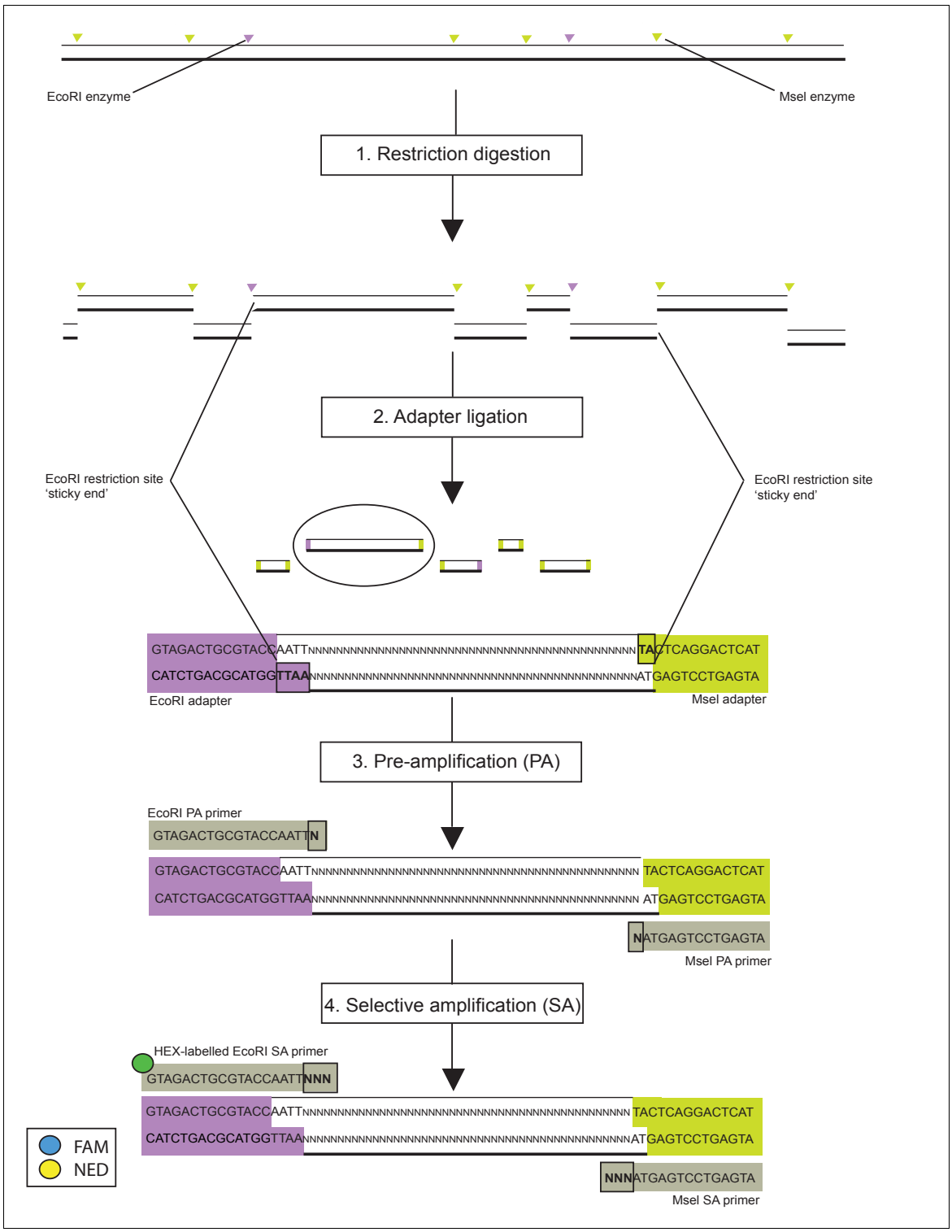


Figure 3.1. Overview of AFLP generation. Step 1: Digestion of genomic DNA with restriction endonucleases EcoRI (rare cutter) and MseI (common cutter). **Step 2:** Ligation of the double stranded adapters (EcoRI and MseI specific) to DNA fragment ends ('sticky ends') **Step 3:** Pre-amplification that amplifies a sub-set of EcoRI/MseI templates (~1/16). Primers match the adapter sequence with a single nucleotide extension (N) **Step 4:** Selective amplification further reduces the number of fragments using primers that have an additional two nucleotides. All EcoRI primers used for SA are labelled with a fluorescent dye (FAM, NED or HEX) thus ensuring all resulting fragments are dye-labelled.

Thermal cycling conditions for selective PCR consisted of 2 min at 94°C followed by 10 cycles with 20 sec at 94°C, 30 sec at annealing temperature, which decreased in each cycle by 1°C from 65°C to 56°C, and 2 min at 72°C. The PCR continued for 25 cycles with 20 sec at 94°C, 30 sec at 56°C and 2 min at 72°C, followed by a holding step at 60°C for 30 min. Five individuals were repeated from the restriction ligation stage onwards to obtain a relative assessment of the repeatability of AFLP profiles.

3.4.4. Primer testing

AFLP profiles vary widely in number of loci amplified between selective amplification products produced by different primer combinations (Meudt and Clarke 2007). Given that the EcoRI restriction site occurs less frequently in the genome than the MseI restriction site (frequent cutter), enzyme-specific primer pairings can have major consequences for the number of AFLP fragments generated (Meudt and Clarke 2007). Pairing two EcoRI-specific primers together (EcoRI-EcoRI) will result in fewer fragments than a pairing of EcoRI and MseI or MseI and MseI respectively.

The number of nucleotides used in primer selective base pair extensions will also dramatically affect the quality of AFLP profiles, with longer extensions reducing the number of AFLP fragments (Vos *et al.* 1995; Bonin *et al.* 2004). In good quality AFLP profiles, the number of AFLP fragments should be high enough to maximise resolution but low enough to minimise homoplasy. This study follows a previous AFLP study on the Mascarene grey white-eye (Mila *et al.* 2010), which used EcoRI-MseI primer combinations with three selective base pair extensions.

The combination of nucleotides in the selective base pair extension can also affect the quality of AFLP profiles (Bensch and Åkesson 2005). It is therefore necessary to screen different primer combinations to get a measure of amplification efficiency (Meudt and Clarke 2007). This study screened a total of 21 unique primer combinations that were generated from three selective amplification EcoRI+NNN primers (labelled with different fluorescent dyes) and seven MseI+NNN primers. A subset of eight DNA extracts was chosen to test all 21 AFLP primers combinations and resulting selective amplification products were electrophoresed on a 3.5% agarose gel against a Hyperladder V (Bioline)

size standard. The resulting electrophoresis runs were used to choose the most appropriate primer combinations. Good primer pairs were identified as those that produce numerous visual fragments of between 100-500 base pairs (bp) with little or minimal background smearing.

Table 3.1. AFLP adaptors and primers

Primer name	Primer sequences	Dye
Adaptors		
EcoRI A	5'-CTCGTAGACTGCGTACC-3'	
EcoRI B	5'-AATTGGTACGCAGTCTAC-3'	
MseI A	5'-GACGATGAGTCCTGAG-3'	
MseI B	5'-TACTCAGGACTCAT-3'	
Pre-amplification primers		
EcoRI	5'-GACTGCGTACCAATTCA-3'	
MseI	5'-GATGAGTCCTGAGTAAC-3'	
Selective-amplification primers		
SA_EcoRI_ACT	5'-GACTGCGTACCAATTCACT-3 '	HEX
SA_EcoRI_ACA	5'-GACTGCGTACCAATTCACA-3'	FAM
SA_EcoRI_AAC	5'-GACTGCGTACCAATTCAAC-3'	NED
SA_MseI_CGC	5'-GATGAGTCCTGAGTAACGC-3'	
SA_MseI_CTT	5'-GATGAGTCCTGAGTAACTT-3'	
SA_MseI_CAC	5'-GATGAGTCCTGAGTAACAC-3'	
SA_MseI_CTA	5'-GATGAGTCCTGAGTAACTA-3'	
SA_MseI_CAG	5'-GATGAGTCCTGAGTAACAG-3'	
SA_MseI_CAT	5'-GATGAGTCCTGAGTAACAT-3'	
SA_MseI_CTC	5'-GATGAGTCCTGAGTAACTC-3'	

Black = sticky end, purple= core adapter sequence, blue= enzyme-specific adapter sequence and red= selective base pair extension.

3.4.5. Pooling fluorescently labelled primer combinations

Fluorescent labelling of AFLP fragments has revolutionised the use of AFLP-PCR. This technique avoids the use of isotopes or silver staining (Huang and Sun 1999) and gives a much higher degree of resolution providing the AFLP banding patterns to the level of single nucleotide differences (Mueller and Wolfenbarger 1999). Using differently labelled primer combinations also allows selective amplification products to be pooled, allowing multiple AFLP primer combinations to be run in a single lane (Fig. 3.2). However, variation in amplitude of emission between fluorophores can often result in poor AFLP profiles and therefore investigating optimum pooling ratio is advocated (Meudt and Clarke 2007).

For this study, fragment analysis was conducted on a 3730 Applied Biosystems Sanger Sequencer using recommended fluorophores (FAM, NED HEX and LIZ). To account for differential amplitude of emissions between dyes (fluorophores), five samples were run using a series of dilution ratios (neat, 1:5 1:10) for each primer combination. From these dye ratio tests, an RFU emission standard could be identified for each dye, from which the optimum-pooling ratio could be calibrated.

3.4.6. Scoring AFLP profiles

The generation of an AFLP binary matrix can be a challenging process (Bonin *et al.* 2004; Pompanon *et al.* 2005; Meudt and Clarke 2007; Holland *et al.* 2008). Detecting homology of fragments across multiple taxa is compounded by the problem of homoplasy that makes identifying truly homologous characters (or alleles) difficult (Bensch and Åkesson 2005). Scoring profiles is also compounded by variation between samples, in peak amplitude and width, in addition to the presence of shoulder peaks or stuttering (Pompanon *et al.* 2005) (Fig. 3.3, A). The challenge in scoring AFLP profiles is to maximise the signal to noise ratio by optimising analysis parameters such as: peak amplitude threshold (the intensity above which a peak is scored); the bin width and position (size and position in which peaks are considered homologous); and the minimum fragment size recorded (Meudt and Clarke, 2007; Holland *et al.* 2008).

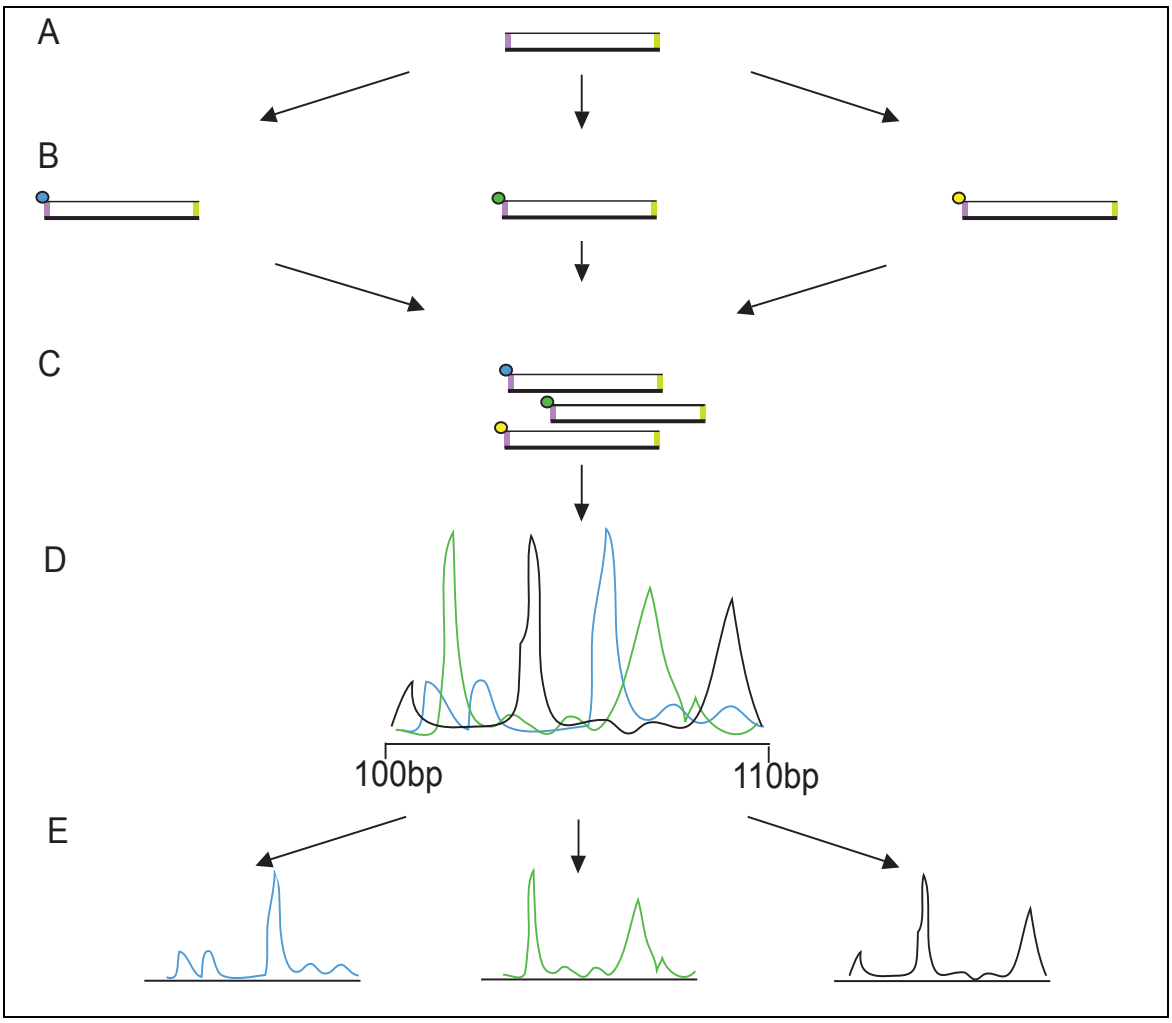


Figure 3.2. Overview of pool-plexing procedure. A: Pre-selective PCR amplification of a sample **B:** Selective PCR amplification using different primers, each labelled with different fluorescent dyes (FAM=blue, HEX=green and NED=yellow). **C:** Pooling of differently labelled PCR products **D:** capillary electrophoresis of pooled samples with resulting fragment read (FAM=blue, HEX=green and NED=Black). **E:** Extraction of data profile for each fluorescent dye/primer combination.

Peak amplitude thresholds, or more specifically variation in peak intensity between samples and fragments, provide a common source of error when scoring AFLP profiles (Bonin *et al.*, 2004; Meudt and Clarke 2007). Variation in fragment intensity between samples (Fig. 3.3,B) makes determining the peak amplitude threshold difficult. Peak intensity of AFLP data can vary widely and therefore special attention must be paid to the amplitude threshold, which should be low enough to detect the weakest peaks, but high enough above the background to eliminate noise (Holland *et al.* 2008). Fragment mobility is a second source of scoring error (Fig. 3.3,C) and is the result of variation in fragment length or poorly calibrated reads (Bonin *et al.* 2004; Meudt and Clarke 2007). The position and width of the bin dictates whether peaks from different samples are split into separate characters or grouped under a single character.

Consequently, bin width and bin position can have a significant effect of scoring efficiency (Holland *et al.* 2008).

Most AFLP automated scoring software allows for the control of scoring parameters such as amplitude threshold, bin width and minimum fragment size (Holland *et al.* 2008). In contrast to manual scoring, automated scoring is repeatable and far less time consuming however, while adjustable, threshold parameters are used for all fragment lengths across all samples. Scoring thresholds (amplitude and width) can vary widely both across fragment lengths and between samples and therefore using a single threshold can still result in significant scoring error. For example the AFLP-PCR procedure often results in numerous small fragments (e.g.150bp) and comparatively fewer larger fragments (e.g. 450bp). As a result AFLP reads often have more high intensity peaks with a lower signal to noise ratio at smaller fragment lengths when compared to larger fragments lengths. Optimising scoring thresholds for smaller fragments would result in a loss of larger fragments, while optimising based on larger fragments would result in a significant degree of noise. In such situations manual scoring allows for bin-specific thresholds, which would permit threshold variation between alleles whilst maintaining high specificity.

Peaks were visualised using GENEMAPPER version 3.7 and all primer combinations were analysed separately. An initial scoring panel was generated using the automatic panel generation feature of GENEMAPPER under default settings. This feature algorithmically generates panels and bins based on the collective peaks present from all samples. The resulting AFLP panels were then checked by eye and preliminary values were set for amplitude threshold and bin width that corresponded to average estimates of peak height and peak width for all fragments. Size standard concordance was checked by eye by overlaying all sample size standards (LIZ) to check for variation between samples. In addition Peak Quality Flags generated by GENEMAPPER were checked to assess the quality of sequence reads. All peak quality analyses were run using default setting.

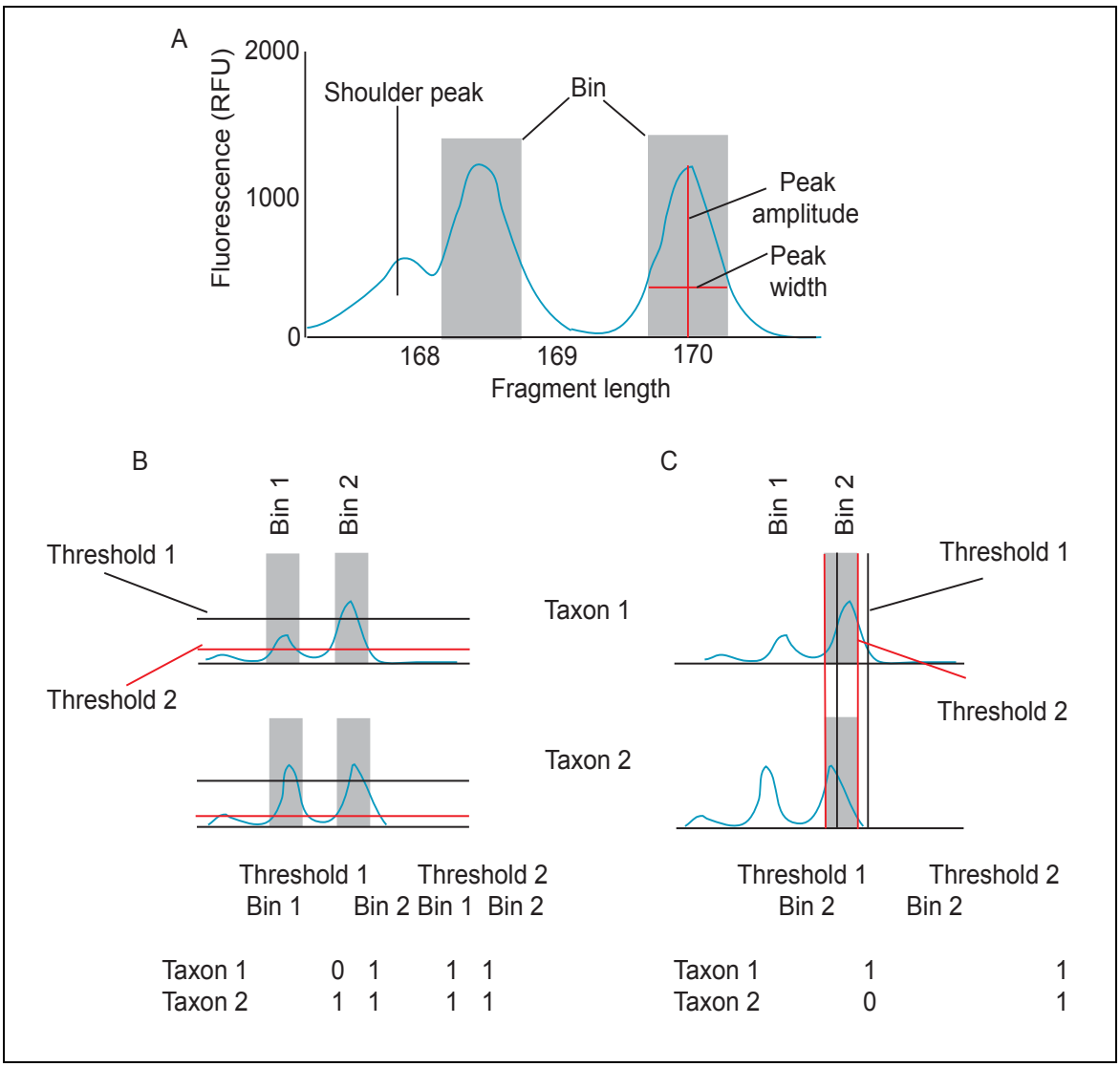


Figure 3.3. AFLP thresholds and common sources of scoring error. A: Example of AFLP profile with associated scoring parameters. **B:** Scoring errors associated with fragment intensity. If the upper threshold is used (black line) the peak in bin 1 for sample 1 is scored as absent (0) and for sample 2 as present (1). Although the threshold can be lowered (red line) to include the peak in bin 1 for sample 1 this can blur the boundary between noise and peaks and may cause the same problem to occur in other bins **C:** Scoring errors associated with peak mobility (fragments size differs between samples). Under threshold 1 the peaks are not considered to be the same character while under threshold 2 they are.

The AFLP samples were then re-analysed using the preliminary analysis parameters to remove much of the poor bin assignment before manual inspection. The resulting AFLP profiles were over-laid and the quality, position and width of each bin relative to the concatenated profile were manually assessed. For each fragment, bins were saved, modified or deleted depending on their fit to the data, with only unambiguously scorable loci (bins) retained for future analysis. In addition, bins corresponding to fragments less that 75bp were removed and peaks found in less than 2% of samples were discounted.

AFLP reads were subsequently re-analysed using the manually edited panel and an amplitude threshold that was approximately half that of the original estimated value. Given the variation in peak height both across samples and between bins, each bin within the panel was re-assessed and all fragments within each bin were manually scored using an independent bin-specific amplitude threshold. This threshold was identified by overlaying all peaks within each bin and assessing the signal to noise ratio.

3.4.7. Repeatability

Although AFLPs are considered highly reproducible **(**Mueller and Wolfenbarger 1999; Bonin *et al.* 2004), producing replicate or duplicate profiles is advised (Bonin *et al.* 2004; Meudt and Clarke 2007). Unlike DNA sequencing, where correct nucleotides can be determined with a high degree of confidence, the scoring of AFLP profiles is much more subjective. Factors such as plate position, reagents, PCR conditions, laboratory equipment, fluorescent dyes, size standard and capillary instrument can all affect reproducibility and comparability of AFLP profiles (Meudt and Clarke 2007). Therefore ensuring consistency throughout the study is critical. To remove any positional sources of error, plate position was randomised and samples were labelled anonymously. In addition, plate position, reagents, protocols and equipment were standardised throughout the duration of the study.

The technical aspect of generating AFLP profiles (PCR stutter, non-amplification), the subjectivity associated with scoring profiles, and differences in peak mobility and intensity of AFLP profiles all introduce further sources of error (Bonin *et al.* 2004; Pompanon *et al.* 2005). Although these factors may not directly bias the result of the analysis they cause a reduction in the signal to noise ratio (Bonin *et al.* 2004), which can result in a loss of resolution. Creating replicates and quantifying genotyping error rate is considered an essential component of an AFLP study, because replicates are the only objective measure of quality (Pompanon *et al.* 2005). In total five individuals were repeated from the restriction ligation stage onwards to obtain a relative assessment of the repeatability of AFLP profiles.

3.4.8. Phylogenetic analyses

MrBayes version 3.1.2 (Huelsenbeck and Ronquist 2001) was used for phylogenetic reconstruction of the 255-character AFLP dataset. Four independent Metropolis-coupled Markov Chain Monte Carlo (MCMC) chains (temp=0.2) were run for 10,000,000 generations, with a sampling frequency of 1000 and a relative burn-in of 25%. The binary matrix was coded as data-type=restriction and coding=no absence sites, with all other parameters set as default. The posterior probability branch support values (BPP) were estimated from a majority rule tree of the final 50% of trees generated.

3.4.9. Phylogenetic hypothesis testing

MrBayes was used to investigate support for phylogenetic hypotheses from AFLP data, using an approach that compares the harmonic mean log-likelihood of trees generated under constrained (hypothesised) and unconstrained (observed) topologies (Genner and Tunner 2012). This method differs from traditional approaches because it does not lead to the rejection of a null hypothesis in favour of an alternative hypothesis, but instead evaluates support for a given hypothesis based on available evidence (Genner and Tunner 2012).

Constrained topologies were constructed in MrBayes where Markov chain Monte Carlo analyses were run for 10,000,000 generations (sampling every 1000 generations) using coding data-type=restriction and coding=no absence sites. In total nine constrained topologies were generated (Table 3.2) to test hypotheses regarding the monophyly of species (constraints 1,2,3,4 and 8) as set out by the current taxonomy (monophyly of species) versus the grouping or relationships previously identified by mtDNA (constraints 5,6,7 and 9). For each constrained topology, Bayes factors were calculated as twice the harmonic mean of log-likelihoods between the null and the hypothesised topologies. In concordance with previous studies (Marek and Bond 2006; Genner and Tunner 2012), Bayes factors >10 were considered unsupportive of hypothesised or constrained topologies whereas those <10 were considered supportive.

3.5. Results

3.5.1. AFLP profiles

In total 116 samples were screened to get a measure of template DNA quality and quantity. Seven samples were removed from the study as the presence of smearing on the agarose gel indicated poor quality extracts. These samples represented three *Z. senegalensis* samples from the DRC (ZMUC128660, ZMUC128632 and ZMUC128658), three *Z. senegalensis jacksoni* samples from Kenya (T54, ZMUC131324 and ZMUC131324), and a single *Z. poliogaster kikuyuensis* sample from Mt Kenya (AB11). After manual inspection of AFLP profiles, a further 18 samples were removed from the dataset (*Z abyssinicus*: T14, T15, T76, T84, T85; *Z. poliogaster*: K41, 2MK3, 2MK7, 2MK8, 2MK9, 2MK10, MK1, MK2, MK7, AB13, TH1-2, CH1, CH3) as high levels of noise hindered efficient scoring.

The number of bins (alleles) for each of the 15 primer combinations identified by the initial scoring panel ranged from 211 to 563 across the 15 primer combinations used. In general, NED-labelled primer combinations gave the fewest number of fragments while FAM-labelled primer combinations gave the highest. Average peak amplitude was relatively uniform across primer combinations (~800 RFU), although the range of peak amplitude varied significantly between bins (100-5000). Shoulder stuttering was present in 11 of the 15 primer combinations used and were most frequently observed for FAM-labelled primers. The signal to noise ratio was lowest in FAM-labelled primers and was notably higher in HEX- and NED- labelled primers respectively. The degree or intensity of base line noise was highest at smaller fragment lengths and reduced significantly as fragment length increased. This was concordant with peak amplitude and the number of bins identified by the initial scoring panel, both of which decreased with increasing fragment length.

Manual examination of concatenated AFLP profiles identified a large variation in peak amplitude between samples, which subsequently led to a large proportion of the bin being discounted (~90%). Co-dominant alleles were evident across all primer combinations examined. However, peak amplitude variability between samples hindered assessments of frequency.

The final AFLP data set was constructed from 92 samples and contained 255 AFLP characters. Average estimates of genotyping error, measured as

recommended by Bonin *et al.* (2004), was 0.8%. The number of fragments scored per sample ranged from 45 to 84, with the mean number of fragment scored being 66.9. Of the 255 AFLP-loci examined, 31% (79 alleles) corresponded to private alleles for which scoring was limited to individuals from the same sampling locality.

3.5.2. Phylogenetic resolution and hypothesis testing

Bayesian Inference of the AFLP dataset recovers the two *Z. borbonicus* subspecies (*Z. b. borbonicus* and *Z. b. mauritianus*) (BPP=1.00) as phylogenetically distinct from all other Africa taxa (Fig. 3.4). This result is concordant with the mtDNA phylogeny of Chapter 2 that recovers the Ancient Indian Ocean (AIO) white-eye clade (including *Z. borbonicus*) at the base of all other African *Zosterops*. Mainland African taxa and representatives from the Gulf of Guinea and the *maderaspatanus* clade taxa are recovered as a single clade (BPP=1.00), but there is no support for the broader clades recovered in the mtDNA phylogeny of Chapter 2. Furthermore, Bayesian hypothesis testing failed to provide support for the nine constrained topologies tested (Table 3.2), and as a result this study is unable to support or reject the non-monophyly of *Z. poliogaster*.

While broader relationships across the AFLP-phylogeny are poorly resolved, there is good support for the monophyly of range-restricted taxa that have endemic distributions occupying oceanic islands (i.e. *Z. a. socotranus* and *Z. borbonicus*) and continental montane forest 'sky islands' (i.e. *Z. poliogaster* subspecies). Conversely, there is very little support for the monophyly of subspecies and or populations in more widely distributed taxa such as *Z. abyssinicus* and *Z. senegalensis* (Fig. 3.4). Extensive sampling of two subspecies of *Z. senegalensis* (*Z. s. jacksoni* and *Z. s. stierlingi*) and two subspecies of *Z. abyssinicus* (*Z. a. jubaensis* and *Z. a. flavilateralis*) across several sampling localities reveals limited phylogenetic clustering, and those clusters present, are not concordant with geographic or sub-specific divisions.

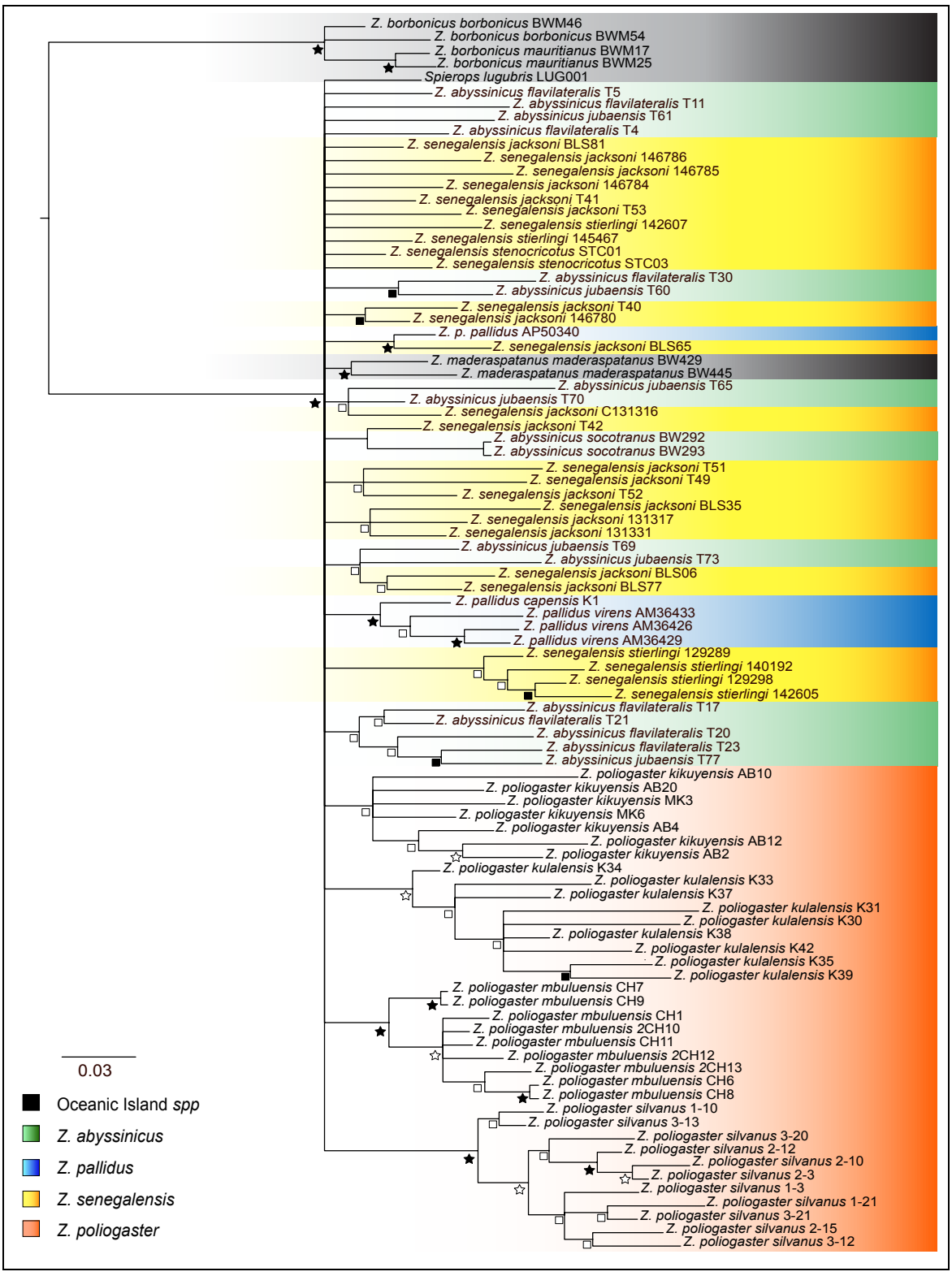


Figure 3.4: Phylogenetic reconstruction of East African *Zosterops* based on nuclear AFLP fragments. Node support in the form of Bayesian posterior probabilities (BPP) is displayed below corresponding nodes. ★ Indicates nodes with >95% BPP, ☆ indicates nodes with >90% BPP), ■ indicates nodes with >80% BPP and □ indicates nodes with > 50% BPP. Nodes with < 50% BPP are not shown. Taxa are labelled using full trinomial nomenclature, following the taxonomy of Dickinson (2003).

Table 3.2. Comparisons of mean likelihoods for unconstrained and constrained phylogenetic topologies generated from AFLP data.

Constraint(s) Employed	Mean LnL: Constrained	Bayes Factor	Phylogenetic Inference
1. *Z. poliogaster* monophyly	-4294.36	44.06	Not supported
2. *Z. senegalensis* monophyly	-4339.42	134.18	Not supported
3. *Z. abyssinicus* monophyly	-4309.45	74.24	Not supported
4. *Z. pallidus* monophyly	-4282.97	21.28	Not supported
5. *Z. p. mbuluensis*, *Z. a. jubaensis* and *Z. a. flavilateralis* combined monophyly	-4325.52	105.84	Not supported
6. *Z. p. silvanus*, *Z. pallidus* and *Z. s. stierlingi* combined monophyly	-4339.41	134.16	Not supported
7. *Z. p. kulalensis*, *Z. p. kikuyuensis* and *Z. s. jacksoni* combined monophyly	-4351.29	157.92	Not supported
8. (Constraint 1) (Constraint 2) (Constraint 3) (Constraint 4) = as taxonomy	-4359.86	175.06	Not supported
9. (Constraint 5)(Constraint 6)(Constraint 7) = as mtDNA phylogeny	-4422.04	299.42	Not supported

Note: Bayes factors are twice the difference in harmonic means of log-likelihoods between constrained and unconstrained trees. Harmonic mean of unconstrained LnL = -4272.33

The relationship between taxon distribution (or taxon vagility) and phylogenetic resolutions is also demonstrated when comparing support values for the monophyly of the four *Z. poliogaster* subspecies sampled in this study (eight subspecies currently recognised). There is a marked difference between support for the monophyly of the more widely distribution *Z. p. kikuyuensis* (BBP=0.63), that was sampled from both Mt Kenya and Aberdare range, relative to *Z. p. kulalensis* (BBP=0.94), *Z. p. silvanus* (BBP=0.99) and *Z. p. mbuluensis* (BBP=0.99), that have comparatively smaller distributions in Mt Kulal, Taita hills and the Chyulu hills respectively.

3.6. Discussion

The principle motivation for this study was to obtain a nuclear assessment of phylogenetic relationships for East Africa *Zosterops*. In comparison with the previous work presented in Chapter 2, this study aimed to identify concordance and discrepancies between phylogenies produced from ncDNA and mtDNA markers. Specifically, this study aimed to address whether a nuclear phylogeny of East Africa *Zosterops* supported the non-monophyly of *Z. poliogaster* as demonstrated in the mtDNA phylogeny of Chapter 2.

In this study AFLPs were chosen as the nuclear component for two reasons. Firstly, AFLPs have been shown to provide phylogenetic resolution in groups where nuclear sequence data resulted in poor phylogenetic resolution of taxa (Albertson *et al.* 1999; Sullivan *et al.* 2004; Mendelson and Shaw 2005; Quek *et al.* 2010; Genner and Tunner 2012). This is particularly evident among recent or rapidly evolving groups in which the comparatively lower rate of evolution of ncDNA genes limits the usefulness of ncDNA sequence data (Meudt and Clarke 2007). In such groups, the increased resolution of AFLPs has been associated with their genome-wide distribution, which is thought to overcome problems associated with locus-specific effects (Quek *et al.* 2010). Secondly, the use of AFLPs in this study was also attractive because in contrast to other marker systems, they can generate high numbers of loci per assay unit with high reproducibility in the absence of prior genomic information and at relatively low costs (Muller and Wolfenbarger 1999).

However, in contrast to other studies that describe good phylogenetic resolution at shallow taxonomic levels using AFLPs (Albertson *et al. 1999*; Sullivan *et al.* 2004; Mendelson and Shaw 2005; Dasmahapatra *et al.* 2009; Quek *et al.* 2010; Genner and Tunner 2012), the phylogenetic results from this study clearly demonstrate that this is not always the case. Although the AFLP methodology delivered in its generation of numerous loci per sample, identification of alleles (bins) in which peaks could be clearly scored (as present or absent) across samples proved extremely problematical.

As previously discussed, the process of scoring of AFLP profiles can be challenging (Bonin *et al.* 2004: Pompanon *et al.* 2005; Meudt and Clarke 2007; Holland *et al.* 2008). In an effort to maximise scoring efficiency, the peak topology (amplitude, width and shape) of bins identified by the initial scoring

panel across all samples were manually examined. This enabled the identification and removal of poor quality bins that were often characterised by wide variation in peak intensity (amplitude) between samples, inconsistencies in peak position or width across profiles and the presence of shoulder stuttering. In manually examining the concatenated reads it became evident that while some bins demonstrated fragment mobility and/or shoulder stuttering, in the majority of cases variation in fragment intensity between samples led to high levels of noise that limited identification of scoring thresholds. As a result, a large proportion of the bins (alleles) initially identified had to be discounted. In addition, while remaining bins contained clearly scorable peaks, many of these bins corresponded to alleles that had little or limited phylogenetic signal (private alleles).

In trying to assess why the AFLP methodology failed to provide adequate phylogenetic resolution, I will here examine the various properties of AFLPs that may have limited the quantity and quality of data contained within the fragments generated. These properties include: (a) non-homology of fragments; (b) fragment length co-dominance; (c) co-dominant noise; and (d) asymmetry in the probability of losing and gaining fragments. These features would increase the amount of stochastic noise in the data and limit the information content of alleles scored making them less likely to recover the correct phylogenetic relationship (Koopman 2005; Simmons *et al.* 2007; García-Pereira *et al.* 2009).

3.6.1 Non-homology of fragments

One assumption of AFLP datasets is that co-migrating bands (fragments of the same length) are homologous or that homoplasy is minimal. In the latter case, it is assumed that the collective signal of true bands is strong enough to overcome the noise generated by the few bands that demonstrate non-homology (Koopman 2005; García-Pereira 2009). For datasets that contain substantial non-homology of fragments, analyses can result in a considerable underestimation of genetic diversity, spurious phylogenetic relationships and/or a loss of phylogenetic resolution (Bonin *et al.* 2004; Bensch and Åkesson 2005). It may be unsurprising therefore that the most frequently discussed drawback of the AFLP technique is non-homology of profiles (Mueller and Wolfenbarger 1999; Bensch and Åkesson 2005; Meudt and Clarke 2007).

Homoplasy occurs when different accessions are incorrectly scored as having a shared character state as a result of either the co-migration of non-homologous fragments (fragment size homoplasy: Althoff *et al.* 2007; Gort and Eeuwijk 2011: and fragment size collision: Vekemans *et al.* 2002), or independent losses of a shared fragment (fragment length co-dominance: Simmons *et al.* 2007). Non-homology of shared fragments (fragment size homoplasy and fragment size collision) occurs when fragments of equal size do not originate from the same locus. Non-homology of AFLP loci can occur in comparisons among samples (Althoff *et al.* 2007; Gort and Eeuwijk 2011), when fragments of equal length actually represent independent DNA loci (Fig. 3.5,A), and may also occur within individuals (Vekemans *et al.* 2002), when co-migrating AFLP fragments of the same length do not represent the same genomic region (Fig. 3.5,B). This study follows Gort and Eeuwijk (2011) in using the terms 'homoplasy' and 'collision' to distinguish between fragment size non-homology among, and within, individuals respectively.

3.6.2 Fragment size homoplasy

Fragment size homoplasy is particularly concerning in studies of genetic diversity and phylogenetic construction, as high incidence of fragment size homoplasy can lead to poor groupings in which high similarity between individuals does not reflect shared ancestry (Bensch and Åkesson 2005). During the last decade various empirical and theoretical studies have tested the homology of co-migrating bands and have demonstrated that lack of homology and thus poor phylogenetic assignments increases dramatically among profiles from increasingly divergent taxa (Althoff *et al.* 1997; O'Hanlon and Peakall 2000). However, while there is a general consensus regarding the relationship between problems with homology assignments in relation to degree of taxonomic divergence, the taxonomic level at which AFLPs become unreliable is still a matter of debate (Dasmahapatra *et al.* 2009; García-Pereira *et al.* 2009; Smith *et al.* 2011).

Although resolution of relationships is generally poor across the AFLP phylogeny generated in this study, there is good support for the division of a clade containing two *Z. borbonicus* subspecies (*Z. b. borbonicus and Z. b. mauritianus)* at the base of the tree. This relationship is concordant with the mtDNA results of Chapter 2 and is interpreted to suggest that fragment size

homoplasy of profiles from distantly related taxa is unlikely to be the primary cause of the loss of resolution in this dataset.

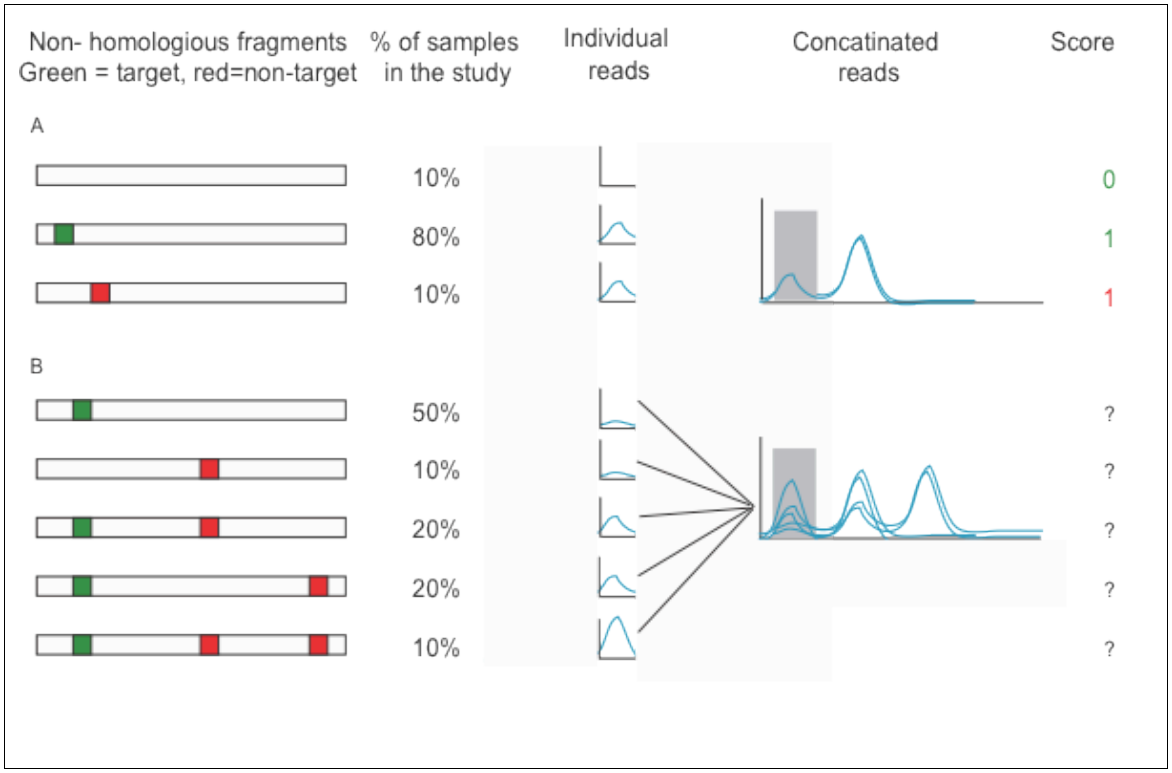


Figure 3.5 Scoring error associated with non-homology of AFLP fragments. A: scoring error associated with fragment length homoplasy. **B:** Loss of data associated with variation in peak amplitudes between samples resulting from fragment length homoplasy (between samples) and fragment length collision (within samples), which limit the identification of a scoring threshold.

3.6.3 Fragment size collision

Fragment size collision has been shown to occur regularly and to increases dramatically with density of amplified fragments (Vekemans *et al.* 2002; Althoff *et al.* 2007; Gort and Eeuwijk 2011). Profiles with more bands are thus more susceptible to collision. Collision can result in incorrect scoring assignments and can increase stochastic noise (Gort and Eeuwijk 2011) that makes identifying the peak height threshold of each bin extremely difficult (Fig. 3.5,B). The initial scoring panels produced in this study contained numerous fragments per primer combination (~300). However, variation in peak topology (peak height, peak width, peak shape) between samples meant that, in the majority of cases, the bins (alleles) identified by the initial scoring panel contained high levels of noise that limited identification of scoring thresholds. As a result, a large proportion of the bins (alleles) initially identified had to be discounted (~90%).

The number of fragments that a sample has for a given allele is directly related to the RFU emission, which in turn is related to peak intensity or peak amplitude. Fragment size collision can directly influence the number of fragments generated for a given fragment length. This can lead to variation in fragment intensity between samples that can limit identification of scoring thresholds. In following strict scoring parameters as suggested in Dasmahapatra *et al.* (2009), this study aimed to minimise the number of non-homologous AFLP bands by excluding any bins that were suspected of collision. While this approach may have significantly reduced the probability of non-homologous bands affecting the analysis, fragment size collision may have inadvertently led to a substantial reduction in the number of bands present in the final dataset. However, in the absence of sequence data for the AFLP fragments generated, it was not possible to quantify the effects of collision in this study.

3.6.4 Co-dominant fragment lengths

Scoring of AFLP data assumes that an absent allele really is absent from the data and does not take into account the different ways in which fragments can be lost (Mendelson and Shaw 2005; Simmons *et al.* 2007). For example, different types of mutations may result in AFLP fragments of different lengths (AFLP-length co-dominance). Under such a situation, two alleles at the same locus could mistakenly be scored as presence alleles at two different AFLP loci (Wong *et al.* 2001). Additionally, a substitution that creates a new cut site between primers may cause the absent allele for one locus to be scored as a presence allele at another AFLP locus (Bensch and Åkesson 2005).

Given that AFLP profiles are typically complex, containing numerous fragments, it is rarely possible to identify polymorphic loci (Bensch and Åkesson 2005). Mutations between primers may result in numerous non-independent loci, thus violating important assumptions regarding phylogenetic construction analyses (García-Pereira *et al.* 2009). The effect of AFLP-length co-dominance is assumed to be negligible as long as mutations between primers sites are rare (<10%) and a large number of informative bands (>100) have been studied (Parsons and Shaw 2001). However, without direct sequencing of AFLP fragments to identify co-dominant fragment lengths it is not possible to get an

accurate estimate of the effect of AFLP fragment length co-dominance on this dataset.

3.6.5 Co-dominant noise

As previously discussed, the number of fragments a sample has for a given allele is directly related to the RFU emission, which in turn is related to peak intensity or peak amplitude. In recognising this, we could predict that peak intensity differences should be positively correlated with allelic copy number (Piepho and Koch 2000). In other words, we would expect individuals that are homozygous (1/1) for a given allele to demonstrate a higher peak intensity value than heterozygous (1/0) individuals that only have one copy of the allele.

The presence of AFLP fragments that are the same size but that demonstrate distinctly different peak intensities (co-dominant allele) has stimulated investigation into the potential for co-dominant scoring of AFLP data (Bensch and Åkesson 2005; Meudt and Clarke 2007). While this dual threshold may be evident in most datasets, investigations into the potential for co-dominant scoring has generally been limited to model or crop organisms (Assuncão *et al.* 2006), for which pre-existing genetic information can be used to assess the accuracy of co-dominant scoring. These studies indicate that the proportion of co-dominant alleles in a given AFLP dataset is generally between 10% and 20% (Alonso-Blanco *et al.* 1998; Assuncão *et al.* 2006), but in some cases this value has been shown to be as high as 75% (Wong *et al.* 2001).

In theory, it should be possible to distinguish between a homozygous present (1/1), heterozygous present (1/0) and homozygous absent (0/0) based on relative peak intensities (Fig. 3.6,A). However, in practice peak amplitude often varies between samples and thus identification of the peak intensity threshold is often based on the range of peak amplitudes exhibited. For co-dominant alleles, the challenge is the fact that co-dominant variation adds an additional dimension of variability. In order to identify the peak threshold, the range of peak amplitudes between samples needs to be narrow. If the range of peak heights in a dataset gets too large, it makes distinguishing the boundary between present and absent all the more difficult and can result in incorrect assignment of samples and/ or loss of data.

Through manually examining concatenated AFLP profiles, the presence of co-dominant alleles is clearly evident. Given that variability in peak amplitude

was the primary factor leading to the rejection of bins identified by the initial scoring panel, co-dominant noise may have been a key factor responsible for the variability in peak amplitude that was widely demonstrated. However, the largely anonymous nature of AFLP fragments means that in the absence of corresponding genetic information it is extremely difficult to quantify the proportion of co-dominant alleles. As a result, the effect of co-dominant noise on peak amplitude variability cannot be quantified here.

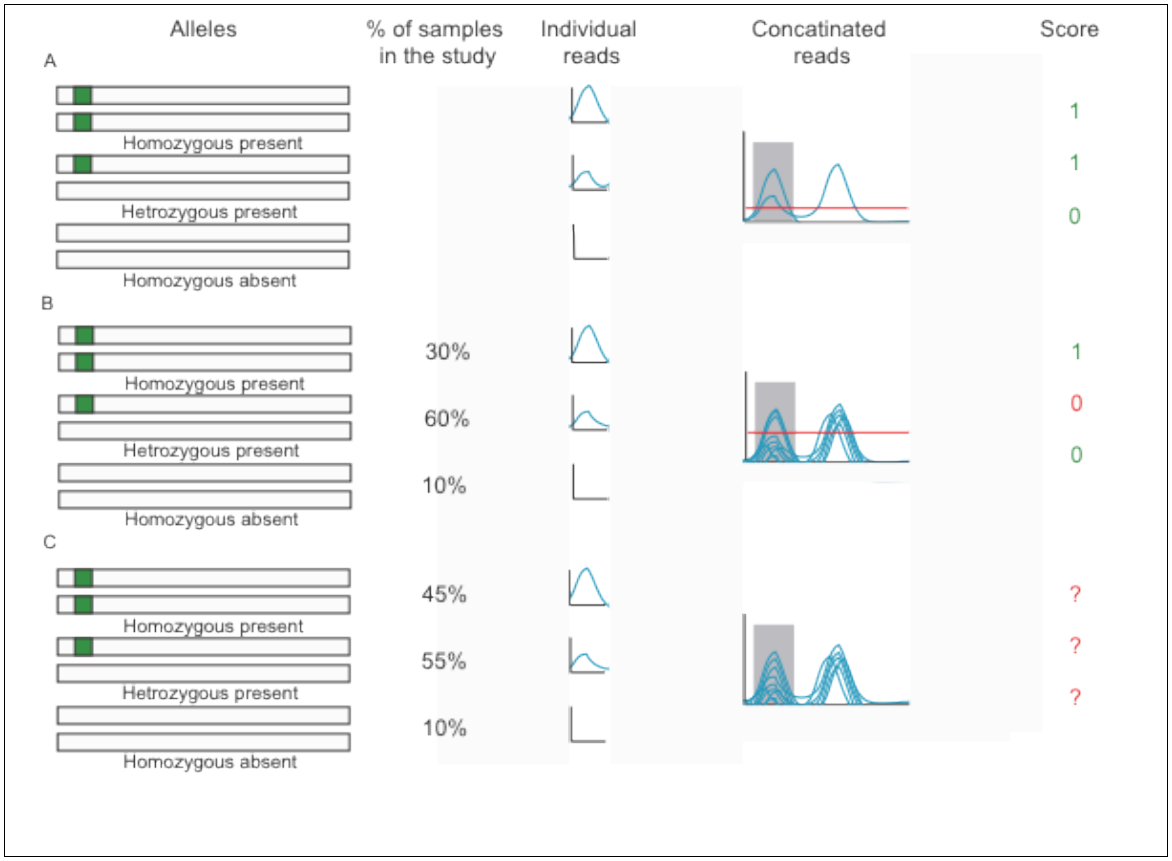


Figure 3.6 AFLP co-dominance and associated noise. A: Example of co-dominant AFLP loci with associated scores. **B:** Scoring errors associated with peak amplitude variability. Wide variation in fragment intensity between samples that are heterozygous for a given allele can result in the incorrect placement of peak amplitude scoring threshold, resulting in heterozygotes being scored as absent. **C:** Loss of data associated with peak amplitude variability. Variation in peak amplitude between samples for both heterozygotes and homozygotes can limit identification of peak amplitude scoring threshold making scoring extremely difficult.

3.6.6 Private alleles

This study aimed to maximise the AFLP signal by excluding any bins that were suspected of non-homology or co-dominance. In spite of this, results clearly demonstrate that alleles contained insufficient phylogenetic signal to adequately resolve phylogenetic relationships. While the scoring approach adopted by this study must have significantly reduced the number of poor

quality alleles, it inadvertently led to a substantial reduction in the number of alleles present in the final dataset.

Given that per-locus information content is relatively poor for AFLPs (Bensch and Åkesson 2005), an insufficient number of fragments may have led to poor resolution of relationships. However, in constructing an AFLP-phylogeny of seals, Dasmahapatra *et al.* (2009) used only 310 AFLP markers, which provided strong phylogenetic signal and resulted in a well-resolved AFLP-phylogeny that was concordant with trees constructed from mtDNA and ncDNA sequence data (Dasmahapatra *et al.* 2009).

Inspecting the decoded AFLP dataset (sample locality attached) revealed that a large proportion of characters contained within the AFLP matrix corresponded to alleles that were specific to a single population (private alleles). The predominance of private alleles may be attributed to asymmetry in the probability of losing or gaining fragments (Koopman 2005; Simmons *et al.* 2007; García-Pereira *et al.* 2009), with parallel losses of fragments occuring more frequently than parallel gains. The large number of private alleles in the dataset is thought to have resulted in strong support for the monophyly of independent populations, specifically range-restricted taxa, with little or limited phylogenetic resolution of relationships between populations. Results are therefore interpreted to suggest that the dramatically lower performance of AFLP analysis with respect to the mtDNA phylogeny generated in Chapter 2 and other AFLP studies with similar size datasets (Dasmahapatra *et al.* 2009) is likely due to the much lower information content of the AFLP dataset rather than insufficient data.

This result is concordant with the *in silico* AFLP simulations of García-Pereira *et al.* (2009), who identified lower information content of AFLP datasets as a factor that has considerable weight to phylogenetic accuracy. The results of García-Pereira *et al.* (2009) indicated that the poor performance of AFLP-based trees is not the result of sampling a much lower number of informative sites and that increasing the number of AFLP markers would still result in poor performance.

3.6.7. Future directions - RAD-sequencing

High throughput next generation sequencing (NGS) technologies offer an alternative method of sampling genome-wide diversity. Until recently the

restricted availability of whole genome sequences, which were previously required for assembling phylogenomic-scale data, has limited the potential use of NGS technologies for phylogenetic construction (Rokas *et al.* 2003; Prasad *et al.* 2008; Philippe *et al.* 2009). However, with the use of restriction-site-associated DNA (RAD) sequencing (Baird *et al.* 2008), it is now possible to assemble genome-wide sequence data from RAD-tags without the use of a reference genome (Sharma *et al.* 2012). Nevertheless, it should be noted that a lack of reference sequences is expected to increase errors in the assessment of gene orthology (Wagner *et al.* 2012).

Baird *et al.* (2008) developed a (RAD-tag) sequencing approach to simultaneously detect and genotype thousands of genome-wide SNPs. This approach focuses the sequencing effort on genomic regions flanking restriction sites, thereby reducing the representation of the genome to be sequenced. Use of the RAD-tag sequencing approach in the field of population genomics is rapidly expanding (Hohenlohe *et al.* 2012; Sharma *et al.* 2012) and studies are now possible even in non-model organisms. In contrast, the use of RAD-tags for phylogenetic inference is in its infancy. Nevertheless the few studies that have been conducted demonstrate the potential power that RAD-tag sequencing has to resolve even the most difficult of phylogenetic questions (Emerson *et al.* 2010; Rubin *et al.* 2012; Wagner *et al.* 2012).

As genomic approaches become cheaper and sequencing technologies allow for more effective surveys, this approach is likely to become an attractive alternative to other commonly used marker systems such as AFLPs or multi-locus gene sequencing. While the behaviour of such data sets in phylogenomic-scale analyses has not yet been systematically evaluated (Rubin *et al.* 2012; Wagner *et al.* 2012), current findings highlight the power that NGS-based data sets hold for resolving species boundaries and relationships, particularly in groups with challenging evolutionary histories.

3.7. Conclusion

In contrast to other studies that describe good phylogenetic resolution at shallow taxonomic levels using AFLPs, the phylogenetic results of this study clearly demonstrate that this is not always the case. Despite the generation of numerous AFLP loci per sample, this study was unable to resolve relationships between the three putative East African *Zosterops* species (*Z. poliogaster*, *Z. senegalensis* and *Z. abyssinicus*). While endemic montane *Z. poliogaster* subspecies form independently well-supported clades, resolution of relationships between taxa using AFLPs is generally poor. Consequently, this study was unable to confirm or reject the non-monophyly of montane endemics that was demonstrated in the mtDNA phylogeny of Chapter 2.

Efforts to maximise the signal to noise ratio resulted in the removal of poor quality bins which in turn led to a substantial reduction in the number of scorable alleles present in the final dataset. Fragment length collision, fragment length co-dominance and co-dominant noise have all been highlighted as factors that may have contributed to the high levels of noise demonstrated. However, in the absence of direct sequencing of AFLP fragments this study is unable to quantify the relative effects of these factors.

Of the characters present in the final dataset, a larger proportion corresponded to private alleles that were specific to a single population. This is thought to have led to strong support for the monophyly of independent populations of range-restricted taxa, with little or limited resolution of more broad-scale phylogenetic relationships. The high number of private alleles is interpreted to suggest that the low information content of this AFLP dataset was a factor with a much higher negative impact on phylogenetic accuracy than insufficient data.

3.8. References

Albertson RC, Markert JA, Danley PD, Kocher TD (1999) Phylogeny of a rapidly evolving clade: the cichlid fishes of Lake Malawi, East Africa. *Proceedings of the National Academy of Science USA*, **96**, 5107-5110.

Al-Daoude, Arabi MIE, Nabulsi I, MirAli N (2012) Molecular phylogeny of *Pyrenophora graminea* as determined by RAPD and ISSR fingerprints. *Journal of Plant Biology Research,* **1**, 25-35.

Alfaro ME, Karns DR, Voris HK, Brock CD, Stuart BL (2008). Phylogeny, evolutionary history, and biogeography of Oriental-Australian rear-fanged water snakes (Colubroidea: Homalopsidae) inferred from mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics and Evolution*, **46**, 576–593.

Alonso-Blanco C, Peeters AJ, Koorneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MT (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and constrution of a Ler/Cvi recombinant inbred line population. *Plant Journal*, **14**, 389-394.

Althoff DM, Gitzendanner MA, Segraves KA (2007) The utility of amplified fragment length polymorphisms in phylogenetics: a comparison of homology within and between genomes. *Systematic Biology*, **56**, 477-484.

Arbogast S, Drovetski SV, Curry RL, Boag PT, Seutin G, Grant PR, Grant BR, Anderson DJ (2006) The origin and diversification of Galapagos Mockingbirds. *Evolution*, **60**, 370-382.

Assuncão AG, Pieper B, Vromans J, Lindhout P, Aarts MG, Schat H (2006) Construction of a genetic linkage map of *Thlaspi caerulescens* and quantitative trait loci analysis of zinc accumulation. *New Phytologist*, **170**, 21-32.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson A (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, e3376.

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

Barluenga M, Stölting KN, Salzburger W, Muschick M, Meyer A (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, **439**, 719-723.

Belaj A, Satavic Z, Cipriani G, Baldoni L, Testolin R, Rallo L, Trujillo I (2003) Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and their effectiveness in establishing genetic relationships in olive. *Theoretical and Applied Genetics*, **107**, 736-744.

Beltrán M, Jiggins CD, Brower AVZ, Bermingham E, Mallet J (2007). Do pollen feeding and pupal-mating have a single origin in *Heliconius*? Inferences from multilocus sequence data. *Biological Journal of the Linnaean Society*, **92**, 221–239.

Bensch S, Åkesson M (2005) Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, **14**, 2899-2914.

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology Notes*, **13**, 3261-3273.

Brito PH, Edwards SV (2008) Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica*, **135**, 439-455.

Campbell D, Duchesne P, Bernatchez L (2003) AFLP utility for population assignment studies: analytical investigation and empirical comparison with microsatellites. *Molecular Ecology* **12**, 1979-1991.

Carstens BC, Knowles LL (2007) Shifting distributions and speciation: species divergence during rapid climate change. *Molecular Ecology*, **16**, 619- 627.

Chan KMA, Levin SA (2005). Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution,* **59**, 720–729.

Clegg SM, Phillimore AB (2010) The influence of gene-flow and drift on genetic and phenotypic divergence in two species of *Zosterops* in Vanuatu. *Philosophical Transactions of the Royal Society B*, **365**, 1077-1092.

Dasmahapatra KK, Mallet J (2006) DNA barcodes: recent successes and future prospects. *Heredity,* **97**, 254–255.

Dasmahapatra KK, Hoffman JI, Amos W (2009) Pinniped phylogenetic relationships inferred using AFLP markers. *Heredity*, **103**, 168–177.

Dickinson EC (ed.) (2003) The Howard & Moore Complete Checklist of the Birds of the World, 3rd Edition. Christopher Helm, London.

Doyle JJ (1992) Gene trees and species trees: molecular systematics as one-character taxonomy *Systematic Botany*, **17**, 144-163.

Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE, Holzapfel CM (2010) Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Science*, *USA*, **107**, 16196-16200.

Estoup A, Clegg S (2003) Bayesian inference on the recent island colonisation history by the bird *Zosterops lateralis lateralis. Molecular Ecology*, **12**, 657-674.

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

Garcia-Moreno J, Sorenson MD, Mindell DP (2003) Congruent avian phylogenies inferred from mitochondrial and nuclear DNA sequences. *Journal of molecular ecology*, **57**, 27-37.

García-Pereira MJ, Caballero A, Quesada H (2009) Evaluating the relationship between evolutionary divergence and phylogenetic accuracy in AFLP data sets *Molecular Biology and Evolution*, **27**, 988-1000.

Genner MJ, Turner GF (2012) Ancient hybridization and phenotypic novelty within Lake Malawi's cichlid fish radiation. *Molecular Biology and Evolution*, **29**, 195-206.

Gerber S, Mariette S, Streiff R, Bodénès C, Kremer A (2000) Comparison of microsatellites and amplified fragment length polymorphisms for parentage analysis. *Molecular Ecology* **9**, 1037-1048.

Goldman DH, Jansen RK, van den Berg C, Leitch IJ, Fay MF. Chase MW (2004) Molecular and cytological examination of *Calopogon* (Orchidaceae: Epidendroideae): circumscription, phylogeny, polyploidy, and possible hybrid speciation. *American Journal of Botany* **91**, 707-723.

Gort G, van Eeuwijk FA (2011) Review and simulation of homoplasy and collision in AFLP. *Euphytica,* **183**, 389-400.

Herbert P D N, Cywinska A, Ball S H, deWaard J R (2003) Biological identification through DNA barcodes. *Proceeding of the Royal Society of London B*, **270**, 313-321.

Hohenlohe PA, Bassham S, Currey M, Cresko WA (2012) Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **367**, 395–408.

Holland BR, Clarke AC and Meudt HM (2008) Optimizing automated AFLP scoring parameters to improve phylogenetic resolution. *Systematic Biology*, **57**, 347-366.

Huang J, Sun M (1999) A modified AFLP with fluorescence-labelled primers and automated DNA sequencer detection for efficient fingerprinting analysis in plants *Biotechnology Techniques*, **13**, 277-278

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics Applications Note*, **17**, 754-755.

Hugall AF, Foster R, Hutchinson M, Lee MSY (2008) Phylogeny of Australasian agamid lizards based on nuclear and mitochondrial genes: implications for morphological evolution and biogeography. *Biological Journal of Linnean Society,* **93**, 343–358.

Jackson ND, Austin CC (2010) The combined effect of rivers and refugia generate extreme cryptic fragmentation within the common ground skink (*Scincella lateralis*). *Evolution,* **64**, 409-428.

Jones CJ, Edwards KJ, Castaglione S, Winield MO, Sala F, Van DWC, Bredemeijer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevschi A, Marmiroli N, Aert R, Volckaert G, Rueda J, Linacero L, Vazquez A, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding,* **3**,381-390.

Joyce DA, Lunt DH, Genner MJ, Turner GF, Bills R, Seehausen O (2011) Repeated colonization and hybridization in Lake Malawi cichlids. *Current Biology,* **21,** 108-109.

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623-2635.

Koopman WJM (2005) Phylogenetic signal in AFLP data sets. *Systematic Biology,* **54**, 197-217.

Linnen CR, Farrell BD (2007). Mito-nuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (Hymenoptera: Diprionidae) sawflies. *Evolution,* **61,** 1417–1438.

Marek PE, Bond JE (2006) Phylogenetic systematics of the colourful, cyanide-producing millipedes of Appalachia (Polydesmida, Xystodesmidae, Apheloriini) using total evidence Bayesian approach. *Molecular Phylogenetics and Evolution*, **41**, 704-729.

Mariette S, Le Corre V, Austerlitz F, Kremer A (2002) Sampling with the genome for measuring within- population diversity: trade-offs between markers. *Molecular Ecology*, **11**, 1145-1156

McKay BD, Zink RM (2010) The causes of mitochondrial DNA gene tree paraphyly in birds. *Molecular Phylogenetics and Evolution,* **54,** 647-650.

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

Mendelson TC, Shaw KL (2005) Use of AFLP markers in surveys of arthropod diversity. *Methods in Enzymology* **395**, 161-177.

Meudt HM, Clarke AC (2007) Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science*, **12**, 106-117.

Mila B, Warren BH, Heeb P, Thébaud C (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: *Zosterops borbonicus*). *BMC Evolutionary Biology*, **10**, 158-171.

Mori N, Moriguchi T, Nakamura C (1997) RFLP analysis of nuclear DNA for study of phylogeny and domestication of tetraploid wheat. *Genes and Genetic Systems,* **72**, 153-161.

Moritz C, Dowling T E, Brown W M (1987) Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. *Annual Review of Ecology and Systematics,* **18** 269-292.

Moyle RB, Filardi CE, Smith CE, Diamond J (2009) Explosive Pleistocene diversification and hemispheric expansion of a 'great speciator'. *Proceedings of the National Academy of Science*, 106, 1863-1868.

Mueller UG and Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. *Tree Reviews*, **14**, 389-394.

Nybom H (2004) Comparisons of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**, 1143-1155.

Oatley G, Voelker G, Crowe TM, Bowie RCK (2012) A multi-locus phylogeny reveals a complex pattern of diversification related to climate and habitat heterogeneity in southern African white-eyes. *Molecular Phylogenetics and Evolution,* **64**, 633-644.

O' Hanlon PC Peakall R (2000) A simple method for the detection of size homoplasy among amplified fragment length polymorphism fragments. *Molecular Ecology*, **9**, 815-816.

Parsons YM, Shaw KL (2001) Species boundaries and genetic diversity among Hawaiian crickets of the genus *Laupala* identified using amplified fragment length polymorphisms. *Molecular Ecology*, **10**, 1765-1772

Philippe H, Derelle R, Lopez P, Pick K, Borchiellini C, Boury-Esnault N, Vacelet J, Renard E, Houliston E, Quéinnec E, Da Silva C, Wincker P, Le Guyader H, Leys S, Jackson DJ, Schreiber F, Erpenbeck D, Morgenstern B,

Wörheide G, Manuel M (2009) Phylogenomics revives traditional views on deep animal relationships. *Current Biology*, **19**, 706–712.

Phillimore AB, Owens IPF, Black RA, Chittock J, Burke T, Clegg SM (2008) Complex patterns of genetic and phenotypic divergence in an island bird and the consequences for delimiting conservation units. *Molecular Ecology*, **17**, 2839-2853.

Piepho HP, Koch G (2000) Codominant analysis of banding data from a dominant marker system by normal mixtures. *Genetics*, **155**, 1459-1468.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, **6**, 847-859.

Prasad AB, Allard MW, Green ED, Program NCS (2008) Confirming the phylogeny of mammals by use of large comparative sequence data sets. *Molecular Biology and Evolution*, **25**, 1795–1808.

Quek SP, Counterman BA, Albuquerque de Moura P, Cardoso MZ, Marshall CR, Mc Millian WO, Kronforst MR (2010) Dissecting comimetic radiations in *Heliconius* reveals divergent histories of convergent butterflies. *Proceedings of the National Academy of Science USA*, **107**, 7365-7370.

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**, 351-360.

Rokas A, Williams BL, King N, Carroll SB (2003) Genome-scale approaches to resolving incongruence in molecular phylogenies *Nature*, **425**, 798–804.

Rubin BER, Ree RH, Moreau CS (2012) Inferring phylogenies from RAD sequence data. *PLoS ONE*, **7**, e33394.

Ruppell O, Stratz M, Baier B, Heinze J (2003) Mitochondrial markers in the ant *Leptothorax rugatulus* reveal the population genetic consequences of female philopatry at different hierarchical levels. *Molecular Ecology*, **12**, 795–801.

Saliba-Colombani V, Causse M, Gervais L, Philouze J (2000) Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. *Genome* **43**, 29-40.

Sanderson MJ, Shaffer HB (2002) Troubleshooting molecular phylogenetic analyses. *Annual Review of Ecology and Systematics,* **33,** 49-72.

Savelkoul PHM, Aarts HJM, de Haas J, Dijkshoorn L, Duim B, Otsen M, Rademaker JLW, Schouls L, Lenstra JA (1999) Amplified fragment length polymorphism analysis: the state of an art. *Journal of Clinical Microbiology*, **37**, 3083-3091.

Shaffer HB, Thomson RC (2007) Delimiting species in recent radiations. *Systematic Biology* **56**, 896-906.

Sharma R, Goossens B, Kun-Rodrigues C, Teixeira T, Othman N, Boone JQ, Jue NK, Obergfell C, O'Neilli RJ, Chikhi L (2012) Two different high throughput sequencing approaches identify thousands of *de novo* genomic markers for the genetically depleted Bornean Elephant. *PLoS ONE*, **7**, e49533.

Shaw K (2002) Conflict between nuclear and mitochondrial DNA phylogenies of recent species radiations: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proceedings of the National Academy of Science USA*, **99**, 16122-16127.

Simmons MP, Zhang LB, Webb CT, Müller K (2007) A penalty of using anonymous dominant markers (AFLPs, ISSRs, and RAPDs) for phylogenetic inference. *Molecular Phylogenetics and Evolution*, **42**, 528-542.

Slikas B, Jones IB, Derrickson SR, Fleischer RC (2000) Phylogenetic relationships of Micronesian white-eyes based on mitochondrial sequence data. *The Auk*, **117**, 335-365.

Smith TA, Mendelson TC, Page LM (2011) AFLPs support deep relationships among darters (Percidae: Etheostomatinae) consistent with morphological hypothesis. *Heredity*, **107**, 579-588.

Sota TR, Ishikawa M, Ujiie F, Kusumoto F, Vogler AP (2001) Extensive trans-species mitochondrial polymorphisms in carabid beetles *Carabus* subgenus *Ohomopterus* caused by repeated introgressive hybridization. *Molecular Ecology*, **10**, 2833-2847.

Spinks PQ, Shaffer HB (2009) Conflicting mitochondrial and nuclear phylogenies for the widely disjunct *Emys* (Testudines: Emydidae) Species Complex, and what they tell us about biogeography and hybridization. *Systematic Biology,* **58,** 1-20.

Sullivan JP; Lavoué S, Arnegard ME, Hopkins CD (2004) AFLPs resolve phylogeny and reveal mitochondrial introgression within a species flock of African electric fish (Mormyroidea: Teleostei). *Evolution*, **58**, 825-841.

Sunnucks P (2000) Efficient genetic markers for population biology. *Tree Reviews,* **15**, 199-203.

Takahashi K, Terai Y, Nishida M, Okoada N (2001) Phylogenetic relationships and ancient incomplete lineage sorting among Cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. *Molecular Biology and Evolution,* **18,** 2057-2066.

Vekemans X, Beauwens T, Lemaire M, Roldán-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, **11**, 139-151.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research,* **21**, 4407-4414.

Wagner CE, Keller I, Wittwer S, Selz OM, Mwiko S, Greuter L, Sivasaunder A, Seehausen O (2012) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, doi: 10.1111/mec.12023

Warren BH, Bermingham E, Bowie RCK, Prys-Jones RP, Thebaud 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 B, Bermingham E, Prys-Jones RP, Thebaud C (2006) Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on the Indian Ocean islands. *Molecular Ecology,* **15**, 3769-3786.

Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, Hammer MF (2004) Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of female's versus males. *Nature Genetics*, **36**, 1122-1125.

Wong A, Forbes MR, Smith ML (2001) Characterisation of AFLP markers in damselflies: prevalence of co-dominant markers and implication for population genetic applications. *Genome*, **44**, 677-684.

Yeung CKL, Lin RC, Lei F, Robson C, Hung LM, Liang W, Zhou F, Han L, Li SH, Yang X (2011) Beyond a morphological paradox: Complicated phylogenetic relationships of the parrotbills (Paradoxornithidae, Aves). *Molecular Phylogenetics and Evolution*, **61**, 192-202.

3.9 Appendix II

Taxon	Locality	Sample code	Sample type	Source
Z. abyssinicus				
socotranus	Socota, Gulf of Aden	BW292	Blood	WARREN
socotranus	Socota, Gulf of Aden	BW293	Blood	WARREN
jubaensis	South Horr, Kenya	T60	Blood	COX
jubaensis	South Horr, Kenya	T61	Blood	COX
jubaensis	South Horr, Kenya	T65	Blood	COX
jubaensis	South Horr, Kenya	T69	Blood	COX
jubaensis	South Horr, Kenya	T70	Blood	COX
jubaensis	South Horr, Kenya	T73	Blood	COX
jubaensis	South Horr, Kenya	T76	Blood	COX
jubaensis	South Horr, Kenya	T77	Blood	COX
jubaensis	South Horr, Kenya	T84	Blood	COX
jubaensis	South Horr, Kenya	T85	Blood	COX
flavilateralis	Umani spring, Kenya	T4	Blood	COX
flavilateralis	Umani spring, Kenya	T5	Blood	COX
flavilateralis	Umani spring, Kenya	T11	Blood	COX
flavilateralis	Umani spring, Kenya	T14	Blood	COX
flavilateralis	Umani spring, Kenya	T15	Blood	COX
flavilateralis	Umani spring, Kenya	T17	Blood	COX
flavilateralis	Umani spring, Kenya	T20	Blood	COX
flavilateralis	Umani spring, Kenya	T21	Blood	COX
flavilateralis	Umani spring, Kenya	T23	Blood	COX
flavilateralis	Umani spring, Kenya	T30	Blood	COX
Z. pallidus				
pallidus	South Africa	AP50340	Blood	PFIAO
virens	South Africa	K1	Blood	CIBIO
virens	South Africa	AM36426	Blood	PFIAO
virens	South Africa	AM36429	Blood	PFIAO
virens	South Africa	AM36433	Blood	PFIAO
Z. senegalensis				
stenocricotus	Mt Cameroon, Cameroon	STC01	Blood	CIBIO
stenocricotus	Mt Cameroon, Cameroon	STC03	Blood	CIBIO
jacksoni	Mathews Range, Kenya	BLS06	Blood	BORGHESIO
jacksoni	Mt Nyiru, Kenya	BLS35	Blood	BORGHESIO
jacksoni	Mt Nyiru, Kenya	BLS65	Blood	BORGHESIO
jacksoni	Mathews Range, Kenya	BLS77	Blood	BORGHESIO
jacksoni	Mathews Range, Kenya	BLS81	Blood	BORGHESIO
jacksoni	Mt Nyiru, Kenya	ZMUC 131316	Blood	ZMUC
jacksoni	Mt Nyiru, Kenya	ZMUC 131317	Blood	ZMUC
jacksoni	Mt Nyiru, Kenya	ZMUC 131324	Blood	ZMUC

Taxon	Locality	Sample code	Sample type	Source
jacksoni	Mt Nyiru, Kenya	ZMUC 131325	Blood	ZMUC
jacksoni	Mt Nyiru, Kenya	ZMUC 131331	Blood	ZMUC
jacksoni	N. Aberdares, Kenya	ZMUC 146780	Blood	ZMUC
jacksoni	N. Aberdares, Kenya	ZMUC 146784	Blood	ZMUC
jacksoni	N. Aberdares, Kenya	ZMUC 146785	Blood	ZMUC
jacksoni	N. Aberdares, Kenya	ZMUC 146786	Blood	ZMUC
jacksoni	Kakamega, Kenya	T41	Blood	COX
jacksoni	Kakamega, Kenya	T42	Blood	COX
jacksoni	Kakamega, Kenya	T49	Blood	COX
jacksoni	Kakamega, Kenya	T50	Blood	COX
jacksoni	Kakamega, Kenya	T51	Blood	COX
jacksoni	Kakamega, Kenya	T52	Blood	COX
jacksoni	Kakamega, Kenya	T53	Blood	COX
jacksoni	Kakamega, Kenya	T54	Blood	COX
stierlingi	Tabora, Tanzania	ZMUC 145467	Blood	ZMUC
stierlingi	Udzungwa Mts, Tanzania	ZMUC 140192	Blood	ZMUC
stierlingi	Poroto Mts, Tanzania	ZMUC 142605	Blood	ZMUC
stierlingi	Poroto Mts, Tanzania	ZMUC 142607	Blood	ZMUC
stierlingi	W. Usambara Mts, Tanzania	ZMUC 129298	Blood	ZMUC
stierlingi	W. Usambara Mts, Tanzania	ZMUC 129289	Blood	ZMUC
unknown	DRCongo	ZMUC 128660	Blood	ZMUC
unknown	DRCongo	ZMUC 128632	Blood	ZMUC
unknown	DRCongo	ZMUC 128658	Blood	ZMUC
Z. poliogaster				
kulalensis	Mt Kulal, Kenya	K30	Blood	COX
kulalensis	Mt Kulal, Kenya	K31	Blood	COX
kulalensis	Mt Kulal, Kenya	K33	Blood	COX
kulalensis	Mt Kulal, Kenya	K34	Blood	COX
kulalensis	Mt Kulal, Kenya	K35	Blood	COX
kulalensis	Mt Kulal, Kenya	K37	Blood	COX
kulalensis	Mt Kulal, Kenya	K38	Blood	COX
kulalensis	Mt Kulal, Kenya	K39	Blood	COX
kulalensis	Mt Kulal, Kenya	K41	Blood	COX
kulalensis	Mt Kulal, Kenya	K42	Blood	COX
kulalensis	Mt Kulal, Kenya	2MK3	Blood	HABEL
kulalensis	Mt Kulal, Kenya	2MK7	Blood	HABEL
kulalensis	Mt Kulal, Kenya	2MK8	Blood	HABEL
kulalensis	Mt Kulal, Kenya	2MK9	Blood	HABEL
kulalensis	Mt Kulal, Kenya	2MK10	Blood	HABEL
kikuyuensis	Mt Kenya, Kenya	MK1	Blood	HABEL
kikuyuensis	Mt Kenya, Kenya	MK2	Blood	HABEL
kikuyuensis	Mt Kenya, Kenya	MK3	Blood	HABEL
kikuyuensis	Mt Kenya, Kenya	MK6	Blood	HABEL

Taxon	Locality	Sample code	Sample type	Source
kikuyuensis	Mt Kenya, Kenya	MK7	Blood	HABEL
kikuyuensis	Aberdares Range, Kenya	AB2	Blood	HABEL
kikuyuensis	Aberdares Range, Kenya	AB4	Blood	HABEL
kikuyuensis	Aberdares Range, Kenya	AB10	Blood	HABEL
kikuyuensis	Aberdares Range, Kenya	AB11	Blood	HABEL
kikuyuensis	Aberdares Range, Kenya	AB12	Blood	HABEL
kikuyuensis	Aberdares Range, Kenya	AB13	Blood	HABEL
kikuyuensis	Aberdares Range, Kenya	AB20	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 1 2	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 1 3	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 1 10	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 1 21	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 2 3	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 2 10	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 2 12	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 2 15	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 3 12	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 3 13	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 3 20	Blood	HABEL
silvanus	Taita Hills, Kenya	TH 3 21	Blood	HABEL
mbuluensis	Chyulu Hills, Kenya	CH1	Blood	HABEL
mbuluensis	Chyulu Hills, Kenya	CH3	Blood	HABEL
mbuluensis	Chyulu Hills, Kenya	CH6	Blood	HABEL
mbuluensis	Chyulu Hills, Kenya	CH7	Blood	HABEL
mbuluensis	Chyulu Hills, Kenya	CH8	Blood	HABEL
mbuluensis	Chyulu Hills, Kenya	CH9	Blood	HABEL
mbuluensis	Chyulu Hills, Kenya	CH11	Blood	HABEL
mbuluensis	Chyulu Hills, Kenya	2CH10	Blood	HABEL
mbuluensis	Chyulu Hills, Kenya	2CH12	Blood	HABEL
mbuluensis	Chyulu Hills, Kenya	2CH13	Blood	HABEL
Z. borbonicus				
borbonicus	La Réunion, Mascarenes	BWM46	Blood	WARREN
mauritianus	La Réunion, Mascarenes	BWM54	Blood	WARREN
mauritianus	Mauritus, Mascarenes	BWM17	Blood	WARREN
mauritianus	Mauritus, Mascarenes	BWM25	Blood	WARREN
maderaspatanus	Mt Ankaratre, Madagascar	BW429	Blood	WARREN
maderaspatanus	Mt Ankaratre, Madagascar	BW445	Blood	WARREN
S. lugubris	São Tomé, Gulf of Guinea	LUG001	Blood	CIBIO

Chapter 4

Molecular phylogeny and species status of continental African *Zosteropidae* using museum skins

Table of Contents

4.1 Abstract

Background: The current taxonomy of the African white-eyes (Zosteropidae) is contentious; involving the recognition of many putative species based largely on poor morphological characters. The pace at which morphological characters, particularly plumage, appear to change within Zosteropidae has led to considerable taxonomic instability. From the few phylogenetic studies that have included sampling of African Zosteropidae, it is becoming increasing apparent that traditional taxonomic approaches have led to some poor taxonomic groupings. However, previous efforts to resolve the systematics of this group have been hindered by its broad distribution, which makes extensive sampling for fresh material extremely difficult.

Methods: Using DNA extracted from museum material, this study greatly expands on the sampling of Chapter 2 providing unprecedented sampling of the African Zosteropidae system (inc. Arabian Peninsula, Gulf of Aden, Gulf of Guinea and Indian Ocean). Eight novel primer sets were designed to break down mitochondrial genes cytochrome b (Cyt b) and NADH dehydrogenase subunit III (ND3) into a series of smaller overlapping fragments that could be amplified using DNA obtained from museum material. This dataset was supplemented with sequence data for Cyt b and ND3 previously generated in Chapter 2. A comparatively smaller subset of sequences for the nuclear transforming growth factor-beta 2 (TGFß2) was also included. Phylogenetic analysis was performed using Bayesian Inference and a GMYC approach was used to examine the transition between Yule and coalescent processes across the tree.

Results: The Bayesian topology generated from the concatenated Cyt b, ND3 and TGFß2 dataset is generally well supported, and is largely congruent with the mtDNA phylogeny produced in Chapter 2. Extensive sampling across Sub-Saharan Africa, the Indian Ocean region, the Arabian Peninsular and the Gulf of Guinea region has revealed that the Africa Zosteropidae complex is comprised of six major clades. Relationships within these six clades are well supported, although resolution of more broad-scale relationships between these clades is less clear. Within the African Zosteropidae system GMYC analysis provided strong support for 14 distinct evolutionary lineages. While genetic cluster probabilities (neutral coalescence and Yule diversification) were strongly

supported at the base of the tree and within some of the terminal taxa (P>0.80 tips and P<0.15 base), branching relationships in much of the tree yielded intermediate probabilities between 0.5 and 0.3.

Discussion: The use both archive and fresh material has enabled the largest genetic assessment of western Zosteropidae to date. This work provides the first comprehensive molecular framework of mainland African taxa and has enabled the assessment of current and previous taxonomic arrangement. The widespread non-monophyly of mainland African species renders the current taxonomic groupings invalid. GMYC analysis recovers 14 distinct evolutionary lineages within this group. However, further analysis using model-based species delimitation approaches is required to see whether these distinct lineages represent species or taxa at a different hierarchical level.

4.2. Introduction

4.2.1. Species delimitation

Defining species and estimating their phylogenetic relationships is a major aim of systematics and plays an important role in every field of biology (Agapow 2004; de Queiroz 2007; Wiens 2007). Species are the fundamental unit of evolutionary biology, and their delimitation can have broad implications ranging from biological conservation (Agapow 2004; Balakrishnan 2005) to comparative evolutionary analyses (Leaché and Fujita 2010; Hamilton *et al.* 2011). While species discovery is fundamental to the basic understanding of biodiversity, progress in species delimitation methods have previously been hindered by issues regarding the concept of a species itself (de Queiroz 2007).

In recent years discussion of species concepts has shifted from philosophical and conceptual questions to a more pragmatic approach (Wiens 2007). This shift has come from recognising the distinction between what species are and the evidence used to recognise them (de Queiroz 2007). Speciation is a continuous process and the characteristics affected by this process are highly diverse. Disagreement between rival species concepts comes from the use of these characteristics as defining criteria (Agapow 2004). The problem lies in the fact that changes in these characteristics occur at different stages of speciation and thus these criteria often come into conflict (Wiens 2007).

The unified species concept identifies that all modern species definitions are variations on the same general lineage concept of a species, because the various alternative definitions equate species, either explicitly or implicitly, with separately evolving metapopulation lineages (de Queiroz 2007). By combining properties that previously created incompatibilities among alternative species concepts, the unified species concept provides various 'operational criteria' or lines of evidence that can be used in assessing lineage divergence (de Queiroz 2007). While the 'unified species concept' will not resolve species delimitation in practice, it provides a unified context for understanding the relevance of integrating the various methods used in solving the problem of species delimitation.

In practice, morphological and molecular approaches can be mutually informative and are often the most feasible (Wiens and Penkrot 2002; Wiens

2007). Genetic data is among the most common source of evidence used in delimiting species. However, it is only recently that general and objective methods for delimiting species using comparative phylogenetic data have been proposed (Wiens and Penkrot 2002; Pons *et al.* 2006; Knowles and Carstens 2007; Barraclough *et al.* 2009; Monaghan *et al.* 2009; Carstens and Dewey 2010; Leaché and Fujita 2010; Powell 2012).

Model-based species delimitation approaches (Rach *et al.* 2008; Carstens and Dewey 2010; Leaché and Fujita 2010; Yang and Rannala 2010) can be particularly challenging in situations where hypothesised species boundaries are problematic. This is particularly evident for groups, such as African Zosteropidae, which lack a practical species-level taxonomic framework. In contrast, methods based on the general mixed Yule-coalescent (GMYC) model (Pons *et al.* 2006; Monaghan *et al*. 2009 Powell *et al.* 2012) are particular suitable where a *priori* specification of hypothesised species boundaries is inconvenient as they require no prior assumptions regarding the probable placement of species boundaries (Powell 2012).

Developing from early comparisons of branching patterns contained within phylogenetic trees, GMYC models are a model-based likelihood approach that combines phylogenetics and coalescence theory. These approaches examine transitions in tree-branching patterns between long inter-specific branches and short intra-specific branches (Pons *et al.* 2006), and have been used extensively to estimate species boundaries from DNA sequence data (Pons *et al.* 2006; Monaghan *et al.* 2009; Hamilton *et al.* 2011; Ceccarelli *et al.* 2012; Harrington and Near 2012; Powell 2012). In groups where an ambiguous taxonomic framework exists, GMYC approaches can be used to generate probabilistic taxonomic hypotheses, which can then be utilised to form a series of a *priori* hypotheses which can be tested using model-based species delimitation approaches (Powell 2011).

4.2.2. Zosteropidae (white-eyes)

The hyper-diverse avian family Zosteropidae (Aves: Passeriformes) is made up of small, gregarious, arboreal birds that exhibit remarkable uniformity in their morphological structure, plumage and behaviour (van Balen 2008). The genus *Zosterops* dominates within the family, including over 70% of all recognised species, and is notorious for the phenotypic uniformity that has

traditionally made designation of taxon boundaries within the family difficult. While members of this family are generally sedentary in nature, this group is best known for its exceptional colonisation abilities (Slikas *et al.* 2000; Warren *et al.* 2006; Phillimore *et al.* 2008; Moyle *et al.* 2009; Melo *et al.* 2011), which gives the family a wide distribution occupying the Afrotropics, southern and eastern Asia (from Indian subcontinent, through to Japan), Australasia and the tropical islands of the Indian Ocean, the western Pacific Ocean, and the Gulf of Guinea region.

Resolving the systematics of this family is notoriously difficult, particularly at the species-level where broad geographic ranges and homogenous appearance complicate efforts to identify natural groupings (Moreau 1957; Mayr 1965). Morphology and distribution have guided much of the current taxonomy, supported by facts regarding general behaviour, nesting, and vocalisations. However, despite extensive analyses by Moreau (1957) for western Zosteropidae, and Mayr (1965) and Mees (1957; 1961; 1969) for eastern Zosteropidae, the affinities of numerous taxa remain unresolved (van Balen 2008).

4.2.3. Taxonomic complexities

Species delimitation in this group has relied heavily on fine morphological differences, particularly within the genus *Zosterops* (van Balen 2008). Structural variation across the range of the family is slight and characteristics such as body size and wing length can often be linked to abiotic variables such as attitude and temperature (Moreau 1957). Nevertheless, divergent phenotypes have been demonstrated in numerous insular taxa, which have been attributed to rapid morphological change associated with exploitation of novel habitats (Clegg *et al.* 2002; 2008; Phillimore *et al.* 2008; Milá *et al.* 2010).

The use of plumage colouration as a taxonomic character has received considerable attention and has been used widely as a tool for facilitating taxonomic arrangements (Mees 1957; Moreau 1957; Mees 1961; 1969). While Zosteropidae possess relatively simple plumage patterns, the distribution and gradation of colouration between forms appears to change readily (van Balen, 2008). The white eye-ring that is typical of Zosteropidae is highly exaggerated in some taxa (e.g. *Z. poliogaster*) and reduced or absent in others (e.g. *Chlorocharis* or *Tephrozosterops*). Within the genus *Zosterops,* several forms

have lost, or are in the process of losing, the typical yellow-green pigmentation and variation in belly colour is frequently observed (Mees 1957; Moreau 1957; Mees 1961; 1969).

Despite the widespread use of plumage colouration as a taxonomic tool within Zosteropidae, it has long been recognised that variation in plumage colouration, both within and between populations, is complex (Mees 1957; Moreau 1957; Mees 1961; Mayr 1965; Mees 1969). Geographically disparate forms can often look very similar, while neighbouring taxa (which might be expected to be closely related) often show very different plumage patterns. This is aptly summed up by Mees (1961) who stated "I know of no other group of birds in which close relatives, may differ more from each other than do distantly related species". The seemingly infinite number of groupings makes it practically impossible to confidently postulate relationships within this family based on plumage differences alone. Nevertheless, in many cases plumage variation has provided the primary characters for species delimitation (van Balen 2008).

Although Zosteropidae members occupy a diverse range of habitats (e.g. remote island archipelagos, continental montane forests, arid lowland savannahs and semi-deserts), ecology has not proven to be a reliable guide to taxonomic relationships. For example, while several subspecies within *Z. chloris* (Lemon-bellied white-eye) are restricted to small coral islands in the Indo-pacific, this species also occurs throughout the mainland of larger islands such as Sulawesi (van Balen 2008). The lack of concordance between ecology and systematics has also been noted in continental forms. The mtDNA phylogeny of chapter 2 demonstrates that the endemic populations of *Z. poliogaster*, which occupy isolated montane forest fragments in East Africa, are in many cases more closely related to taxa occupying arid lowland savannahs than they are to montane endemics that occupy neighbouring forest fragments.

Investigation into the use of vocalisations for species delimitation has seen varying success. While calls appear to have little taxonomic significance (van Balen 2008), song is shown to be a more reliable character of relationships (Mayr 1965; Pratt *et al.* 1987; Fry *et al.* 2001; van Balen 2008) and has been used to split several species, including the Micronesian species *Z. semperi* and *Z. hypolais* (Pratt *et al.* 1987). In a more total evidence approach, Mayr (1965) suggested that song should be used in conjunction with other traits in an

attempt to provide a more robust taxonomic framework. However, while the potential use of song as a taxonomic character within Zosteropidae is well recognised (Moreau 1957; Mayr 1965; Pratt *et al.* 1987; Fry *et al.* 2001; van Balen 2008), the relative utility of song as a taxonomic character in Zosteropidae still remains poorly understood.

4.2.4. African Zosteropidae

Following the taxonomy of Dickinson (2003), Africa and its associated island systems (Gulf of Guinea and Indian Ocean) encompass 14 *Zosterops* species and four *Speirops* species. More than half the African *Zosterops* species are offshore island endemics with only four species restricted to mainland Africa (Dickinson 2003). Recent molecular work (Warren *et al.* 2006; Melo *et al.* 2011; Oatley *et al.* 2012) has highlighted significant problems in the current taxonomy of African Zosteropidae, indicated by the non-monophyly of many species. The systematics of Zosteropidae occurring in the Indian Ocean (Warren *et al.* 2006) and Gulf of Guinea (Melo *et al.* 2011) regions have benefited greatly from recent molecular insights, yet limited sampling across mainland Africa has hindered more broad scale assessments of species validity in mainland Africa.

In his review of African Zosteropidae, Moreau (1957) recognised the taxonomic complexities of this group, which resulted in much uncertainty regarding taxonomic arrangements. Moreau (1957) identified that the features often used to determine relationships among African taxa may be problematic and advised others to 'deal with current taxonomy with great caution'. In the past decade various forms have been switched across species complexes, however a definitive arrangement still appears to be elusive (van Balen 2008).

4.2.5. Mainland African Zosteropidae

Much of Sub-Saharan Africa is occupied by *Z. senegalensis* (Yellow white-eye); a yellow-bellied bird with a green back (Table 4.1), which has a broad range occupying a diverse range of habitats from acacia woodland to evergreen forest (Moreau 1957). There are currently fourteen recognised subspecies of *Z. senegalensis* (Dickinson 2003) that inter-grade widely across the species range (Moreau 1957; Clancey 1967). The nominate subspecies, *Z. s senegalensis* (Senegal to Northwest Ethiopia) intergrades with subspecies *Z.*

s. demeryi (Sierra Leone, Liberia, Ivory Coast), *Z. s. stenocricotus* (Bioko, southeast Nigeria to Gabon) and *Z. s. stuhlmanni* (East Zaire, South Uganda, Northwest Tanzania), while *Z. s. quanzae* (central Angola) intergrades with *Z. s. kasaicus* (Southwest Zaire, Northeast Angola), *Z. s. heinrichi* (Northwest Angola) and *Z. s. anderssoni* (South Angola, Southeast Zaire and West Mozambique to North Namibia, Zimbabwe and East Transvaal).

In the past *Z. s. gerhardi* (South Sudan and Northeast Uganda) and *Z. s. jacksoni* (western Kenya, northern Tanzania) have been included in the *Z. poliogaster* species complex (van den Elzen and König 1983), but as a result of their resemblance to *demeryi*, *stuhlmanni* and *stierlingi* were later placed in *Z. senegalensis* (Fry *et al* 2001). Alternatively, Z. s. *stenocricotus* has been put forward as a candidate for full species status, in recognition of differences in song, with respect to other *Z. senegalensis* subspecies (Fry *et al* 2001),

In the molecular phylogeny of chapter 2, the five *Z. senegalensis* subspecies sampled are recovered into independent lineages rendering *Z. senegalensis* an invalid species. Nevertheless, a more broad scale phylogenetic assessment including denser sampling across the range of *Z. senegalensis* is needed to adequately resolve the taxonomic status of this group.

In the lowlands of the dry Northeast, *Z. senegalensis* is replaced by *Z. abyssinicus* (White-breasted white-eye) that has a duller green back and locally can have either a yellowish or whitish belly (Table 4.1). This species is confined to lowland (<1000m) scrubland or semi-desert habitat of Northeast Africa but is also found outside mainland Africa on the island of Socotra in the Gulf of Aden and also in the southern tip of the Arabian peninsula (Moreau 1957). There are currently six recognised subspecies of *Z. abyssinicus* (Dickinson 2003), although some authors argue that this group represents two separate species based on belly colour (one white or pale-bellied - *Z. abyssinicus*; and the other yellow-bellied - *Z. smithi*; Sclater 1930).

Phenotypic variation in the white or pale-bellied forms has led to the recognition of four subspecies: *Z. a. abyssinicus* (Northeast Sudan, Eritrea and North and central Ethiopia), *Z. a. omoensis* (Ethiopia, Lake Tana to the Omo Valley), *Z. a. arabs* (Arabia), and *Z. a. socotranus* (Socotra and northern Somalia), while yellow-bellied *Z. abyssinicus* forms have been further subdivided into two subspecies: *Z. a. flavilateralis* (southern Kenya to northern

Tanzania) and *Z. a. jubaensis* (southern Ethiopia and southern Somalia to northern Kenya).

The molecular phylogeny of chapter 2 identified *Z. abyssinicus* as a non-monophyletic taxon, recovering yellow-bellied *Z. a. flavilateralis* and *Z. a. jubaensis* as distinct from insular members of *Z. a. socotranus* (white-bellied). At presence, there is no molecular assessment of *Z. abyssinicus* subspecies from Ethiopia and the Arabian Peninsula and therefore very little is known regarding the phylogenetic placement of other *Z. abyssinicus* forms (*Z. a. omoensis, Z. a. arabs* and mainland *Z. a. socotranus*) relative to the two lineages identified in chapter 2.

Much of southern Africa (southern Namibia, southern Botswana, South Africa and southern Mozambique) is occupied by *Z. pallidus* (Cape white-eye). This species is generally much duller in colouration, relative to other Africa *Zosterops*, but exhibits significant variation in belly colouration (Table 4.1). Members of *Z. pallidus* fall into three well-marked phenotypic groupings (Oatley *et al.* 2011) that have historically been considered as separate species (Gill 1936). The nominate subspecies, *Z. p. pallidus*, is white on the belly grading to buff laterally, *Z. p. capensis* is light grey on the belly darkening laterally while *Z. p. virens* is yellow on the belly grading to olive laterally (Table 4.1). Marked variation exists within these broad groupings, which has led to the recognition of additional sub-specific forms *Z. p. atmorii*, *Z. p. sundevalli*, and *Z. p. caniviridis* (Dickinson 2003).

In his taxonomic revision of African *Zosterops*, Moreau (1957) arrived at the tentative conclusion that both *pallidus* and *sundevalli* formed a monotypic species, differing both vocally and in plumage from other *Zosterops* (Fry *et al.* 2001). In contrast, evidence of interbreeding between *capensis* and *atmorii* in upland West Natal and South East Orange-Free State led Moreau (1957) to challenge the species rank of Z. *capensis* and Z. *virens.* Consequently, the white or buff bellied races; *pallidus* and *sundevalli* from the western part of the range (Namibia and central South Africa) were combined under the name *Z. pallidus*, while the grey and green bellied races; *capensis, virens* and *atmorii,* occupying the central and eastern part of the range (southwest South Africa to southwest Mozambique) became subspecies of *Z. virens*. Shortly after, Clancey (1967) reviewed all southern African *Zosterops* and identified several areas where *Z. pallidus* intergraded freely with *Z. virens*. As a result the two species

recognised by Moreau (1957) are now treated as a single species under the earliest name *Z. pallidus* (Dickinson 2003).

The arrangement of this group remains heavily contested (Hockey *et al*. 2005; van Balen 2008; Oatley, 2011) and has driven recent molecular investigation into the relationships of southern African *Zosterops* (Oatley *et al.* 2012). The molecular phylogeny of Oatley *et al.* (2012) included three of the six *Z. pallidus* subspecies (*Z. pallidus*, *Z. capensis* and *Z. virens*), and recovers the nominate subspecies *Z. p. pallidus* as distinct from *Z. p. virens* and *Z. p. capensis*. Genetic analyses confirmed hybridisation in areas of sympatry, which is concordant with the presence of intermediate phenotypes. However, in contrast to previous studies that lumped taxa based on the presence of hybrids, Oatley *et al.* (2012) interprets hybridisation as evidence of habitat type (and associated climatic conditions) driving diversification in southern African *Zosterops*.

The most complex geographic situation exists in northeast Africa where, from the mountains of Ethiopia through the Kenyan Highlands (east of the Rift Valley) and down to several isolated mountains in southern Kenya and northern Tanzania are occupied by members of the *Z. poliogaster* species complex. These comparatively large birds with rich green backs, yellow or grey bellies and some with very broad white-eye rings and bright golden feathers (Table 4.1) are endemic to montane forest habitat and are ecologically segregated from neighbouring *Z. senegalensis* or *Z. abyssinicus* (Hall and Moreau 1970).

The current taxonomic treatment regards these non-intergrading montane populations as eight subspecies of a wider species complex, under the oldest name *poliogaster* (Dickinson 2003). However, given the presence of highland *Z. senegalensis* (*Z. s. jacksoni*) in northern Kenya, Moreau (1957) previously chose to treat the montane populations of northeast Africa as conspecific with *Z. senegalensis*. More recently, some authors have argued to split this group into several species (Collar *et al* 1994; BirdLife International. (2000); Borghesio and Laiolo 2004) on the bases of vocal differences and ecology. However plumage variation within this group was not considered sufficient to warrant species status (David and Gosselin 2002).

Chapter 2 included extensive sampling for five of the eight currently described *Z. poliogaster* subspecies. Results confirmed the non-monophyly of *Z. poliogaster,* but provided strong support for the monophyly of individual

subspecies. Results were interpreted to suggest that the various non-intergrading montane populations should be considered as independent taxonomic units rather than intra-specific taxa. However, it was acknowledged that further investigation using species delimitation methods would be required to adequately resolve taxonomic boundaries.

From the few phylogenetic studies that have included sampling of African *Zosterops* (Warren *et al.* 2006; Moyle *et al.* 2009; Melo *et al.* 2011; Oatley *et al.* 2012; Chapter 2) it is clear that traditional taxonomic approaches have led to some poor taxonomic grouping. Denser sampling of *Zosterops* across continental Africa will be necessary to determine a more comprehensive systematic framework, which should then provide the basis for a complete systematic review of all mainland African taxa.

Table 4.1 Catalogue of mainland African Zosterops subspecies

Zosterops abyssinicus

Z. a. abyssinicus (BMNH 1915 12 24 1198)

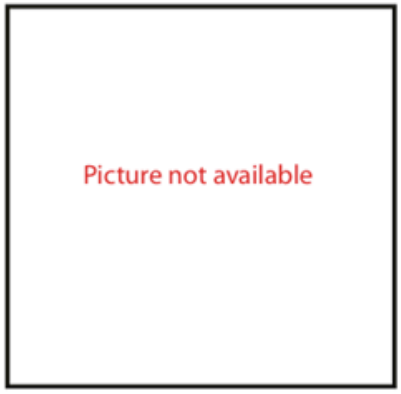
Z. p. arabs (BMNH 1913 7 18 44)

Z. a. socotranus - mainland (BMNH 1982 3 44)

Z. a. socotranus - insular (BMNH 1899 8 11 23)

Z. a. omoensis (BMNH 1927 11 5 580)

Z. a. jubaensis (BMNH 1946 5 2722)



Z. a. flavilateralis

Zosterops pallidus

Z. p. sundervalli (BMNH 1904 11 19 56)

Z. p. virens (BMNH 1905 12 29 1725)

Z. p. atmorii (BMNH 1950 50 666)

Zosterops senegalensis

Z. s. senegalensis (BMNH 1938 8 3 10)

Z. s. demeryi (BMNH 1977 20 2492)

Z. s. stenocricotus (BMNH 1966 16 3408)

Z. s. gerhardi (BMNH 1939 10 1 284)

Z. s. toroensis (BMNH 1936 2 21 237)

Z. s. reichenowi (BMNH 1906 12 23 718)

Z. s. kasaicus

Z. s. heinrichi (BMNH 1957 35 527)

Z. s. quanzae (BMNH 1957 35 531)

Z. s. anderssoni (BMNH 1932 5 5 128)

Z. s. tongensis (BMNH 1905 12 29 1713)

Zosterops poliogaster

Z. p. poliogaster (BMNH 1954 20 26)

Z. p. kaffensis (BMNH 1913 7 18 44)

Z. p. kulalensis (examplar)

Z. p. silvanus (examplar)

Z. p. winifredae (BMNH 1935 12 23 10)

4.3. Aims

The African Zosteropidae system illustrates the trends and complexities of both continental and island species within the family. While recent molecular studies are starting to uncover genetic relationships, to date a lack of sampling has hindered resolution of broad scale relationships within mainland Africa. Furthermore, while recent molecular insights have highlighted taxonomic conflicts, there have been no attempts to test the validity of old or newly formed taxonomic arrangements using quantitative species delimitation methods. Using a combination of fresh and archive samples that give an exceptional coverage of Zosteropidae diversity across mainland Africa and its associated island systems, this chapter aims to generate a robust and comprehensive molecular phylogeny for African Zosteropidae. Using a GMYC approach with multi-model inference and model averaging, this study aims to use the topology generated to make predictions regarding the probable placement of species boundaries in this poorly understood group.

4.4. Materials and methods

4.4.1. Museum collections

Documenting avian diversity and distribution has been fundamental in driving the expanse and maintenance of ornithological collections. For decades these collections have served as an invaluable multi-dimensional resource for research and education providing ecological, morphological and biogeographic data (Winker 2005; Wandeler *et al.* 2007). Since the advent of PCR (Polymerase chain reactions) based methods, developments in molecular techniques have led to an increased interest in museum collections as a source of genetic material (Graves and Braun 1992; Mundy *et al.* 1997; Moum 2002; de Moraes-Barros and Morgante 2007; Lee and Prys-Jones 2008; Töpfer *et al.* 2011). The use of biological collections over fresh samples not only avoids costly fieldwork but also enables researchers to include taxa that are rare and/or extinct (Moum 2002) or difficult to obtain due to either the inaccessibility of habitat or political instability within countries of interest.

Advancements in PCR based techniques have allowed for the extraction and amplification of DNA from poor quality sources including; dried skins (de Moraes-Barros and Morgante 2007; Töpfer *et al.* 2011), toe pads (Mundy *et al.* 1997), eggs (Lee and Prys-Jones 2008) and feathers (Sefc *et al.* 2003). It should be noted however, that although DNA is a chemically stable molecule, DNA from museum specimens ('archive DNA') is subject to DNA degradation (Mandrioli *et al.* 2006; Rohland and Hofreiter 2007). Consequently, DNA isolated from museum material is typically present in low amounts, heavily fragmented, chemically modified and contaminated with environmental DNA (Rohland and Hofreiter 2007; Zimmermann *et al.* 2008; Töpfer *et al.* 2011). In spite of the apparent difficulties of working with archive DNA, museum material has been used extensively to address a wide range of biological questions in Aves, from species verification (Norman *et al.* 1998; Irwin *et al.* 2001; Hennache *et al.* 2003), to broad systematic relationships (Irestedt *et al.* 2006; Slikas 2002).

4.4.2. Taxon sampling

A total of 74 museum specimens (Natural History Museum, Tring), 115 blood sample sequences (Chapter 2) and 79 NCBI sequences (obtained from National Centre for Biotechnology Information) were obtained for this study,

representing 25 species and 53 subspecies (Appendix III). In an attempt to sample the breadth of African *Zosterops* distribution and phylogenetic diversity, attempts were made to obtain three individuals per subspecies for all mainland African forms. Where possible, attempts were made to obtain samples that cover the geographical and phenotypic range of each subspecies (rather than the 'type' locality). Samples were preferentially taken based on their localities (Fig. 4.1), therefore specimens that had poor or ambiguous locality data were not used. In cases were the museum collection contained a few specimens of a particular subspecies (e.g. *Z. senegalensis tongensis*) only one specimen was used in order to minimise damage to the collection. This sample was selected based on the proximity to the 'type' locality.

It was not always possible to obtain multiple samples for all subspecies and as a result this study only obtained limited sampling (≤ 2) for *Z. abyssinicus arabs* (n=2), *Z pallidus caniviridis* (n=1), *Z. senegalensis reichenowi* (n=2), *Z. s. quanzae* (n=1) and *Z. senegalensis tongensis* (n=2). Museum specimen collection dates ranged from 1899 to 1982. The age and locality of specimens can be found in Appendix III.

Blood samples were obtained from numerous research groups that have collected in Africa and its associated island systems over the last decade. Blood samples were taken from mist-netted specimens and were typically stored in ETOH (99%) or Queen's lysis buffer. Sampling localities are listed in Appendix III. A further 79 sequences were obtained from the NCBI database giving coverage of principle lineages in the Indian Ocean and Gulf of Guinea island system, in addition to out-group sampling of Asian and Indo-Pacific member of the genus *Zosterops*.

Given inconsistencies between the primers and target genes used in previous molecular studies of Zosteropidae, taxa whose sequence data was obtained from NCBI often had significant sampling gaps when compared to sequence data that was generated using the primers designed in this study. Where possible, for each putative species, sequences for multiple individuals from various past studies were obtained in an attempt to cover the breath of sequences data targeted in this study.

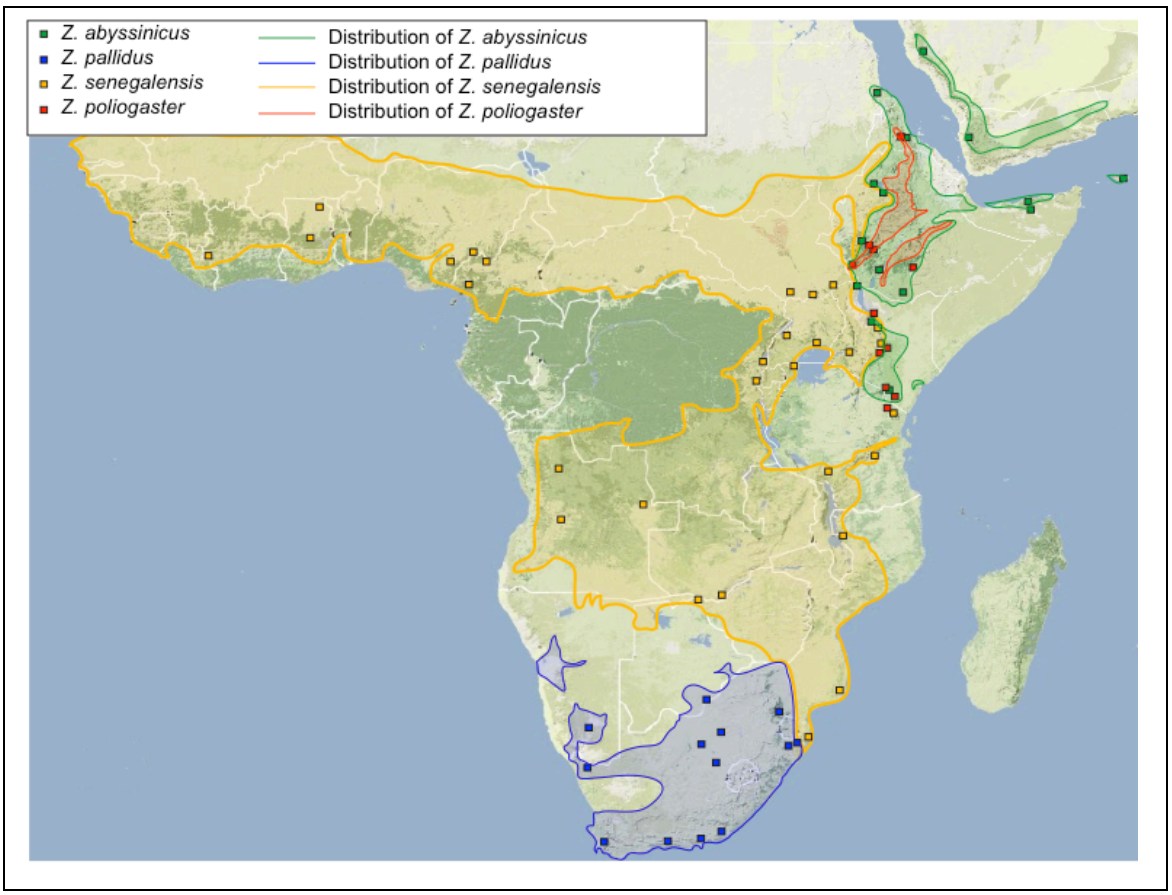


Figure 4.1: Distribution of mainland African Zosteropidae with sampling localities. Image modified from www.googlemaps.com.

4.4.3. Molecular markers

Mitochondrial genes have long served as the preferred marker for phylogeographic and species-level phylogenetic analyses of young systems (Moritz *et al.* 1987; Herbert *et al.* 2003; Ballard and Whitlock 2004; Brito and Edwards 2008). Their use has been recommended in taxonomic studies, with the proposal that all described species are given a mitochondrial DNA (mtDNA) sequence tag or bar-code (Hebert *et al.* 2003). MtDNA sequences have a comparatively higher evolutionary rate and smaller effective population size when compared to the nuclear DNA sequences (ncDNA) (Ballard and Whitlock 2004). Consequently, utilising mtDNA genes increases the chance of recovering relationships and patterns of divergence without an extensive sequencing effort.

In contrast, the use of ncDNA genes often requires researchers to develop primers for multiple genes and sequences across the taxonomic range of the focus group to identify markers with an appropriate evolutionary rate. MtDNA can also be easily amplified across a variety of taxa often making it an attractive marker in more broad-scale taxonomic studies (Brito and Edwards

2008). Conversely, efforts to develop and amplify ncDNA markers often involve extensive refinement of primers between taxa that can substantially increase project start-up times and costs.

While mtDNA is a very useful marker, its use is not without complication. It should be recognised that should male and female history differ in a species, then the use of mtDNA genes as markers fails to reflect the history of the species as a whole (Ballard and Whitlock 2004). There have also been technical issues arising from the presence of nuclear integrations of mtDNA (numts) (Bensasson *et al.* 2001). In a more total evidence approach, recent years have seen an increasing tendency to include ncDNA when generating species-level phylogenies (García-Moreno *et al.* 2003; Beltrán *et al.* 2007; Alfaro *et al.* 2008; Hugall *et al.* 2008). However as seen in Chapter 2, the comparatively lower rate of evolution of ncDNA genes can often limit the usefulness of ncDNA sequence data, with minimal sequence variability between taxa limiting phylogenetic signal.

Available nuclear markers that allow divergence, coalescence, or gene-tree issues to be addressed within Zosteropidae are limited and the degraded nature of 'archive DNA' means that extensive screening for alternative ncDNA sequence markers is beyond the scope of this project. A previous investigation into the potential use of nuclear transforming growth factor-beta 2 (TGFß2) gene revealed it to be of little use at lower taxonomic levels, with results indicating minimal sequence variability between taxa (Chapter 2). Nevertheless, this marker has been useful in assessing more broad-scale relationships with Zosteropidae (Moyle *et al.* 2009) and therefore may provide resolution of relationships between more divergent taxa. Sequences data for the mtDNA genes cytochrome b (Cyt b) and NADH dehydrogenase subunit III (ND3) has been generated for all samples used in this study. This mtDNA dataset is supplemented with sequence data for ncDNA gene TGFß2, which has been compiled for a subset of samples from sequenced data generated in chapter 2 in addition to sequence data available on the NCBI database.

4.4.4. Primer design

'Archive DNA' imposes great difficulties for the retrieval of large amounts of sequence data. The degraded nature of museum extracts means that PCR-based amplification is confined to comparatively short fragments of DNA, rarely

exceeding a couple hundred base pairs (Mundy and Woodruff 1997; Sefc *et al.* 2003; Irestedt *et al.* 2006; Rohland and Hofreiter 2007; Bantock *et al.* 2008; Lee and Prys-Jones 2008). Amplifying and sequencing DNA from museum samples required a series of new primers to be designed. In an attempt to obtain greater amounts of sequence data, eight primer sets were designed to break down the Cyt b and ND3 genes into a series of smaller overlapping fragments (150 – 250 bp) (Fig. 4.2).

The effectiveness and sensitivity of PCR largely depends on the efficiency of the primers (Dieffenbach *et al.* 1993). To generate primers with high specificity and a sufficiently high melting temperature (T_m, 57-62°C), optimal length for primer design was set between 17-28 bp. Where possible attempts were made to design primers with a 40-60% GC content with several G or C bases at the 3' end (GC clamp). The stronger hydrogen bonding of G and C with respect to A and T bases helps promote correct binding at the 3' end and results in more efficient priming (Dieffenbach *et al.* 1993).

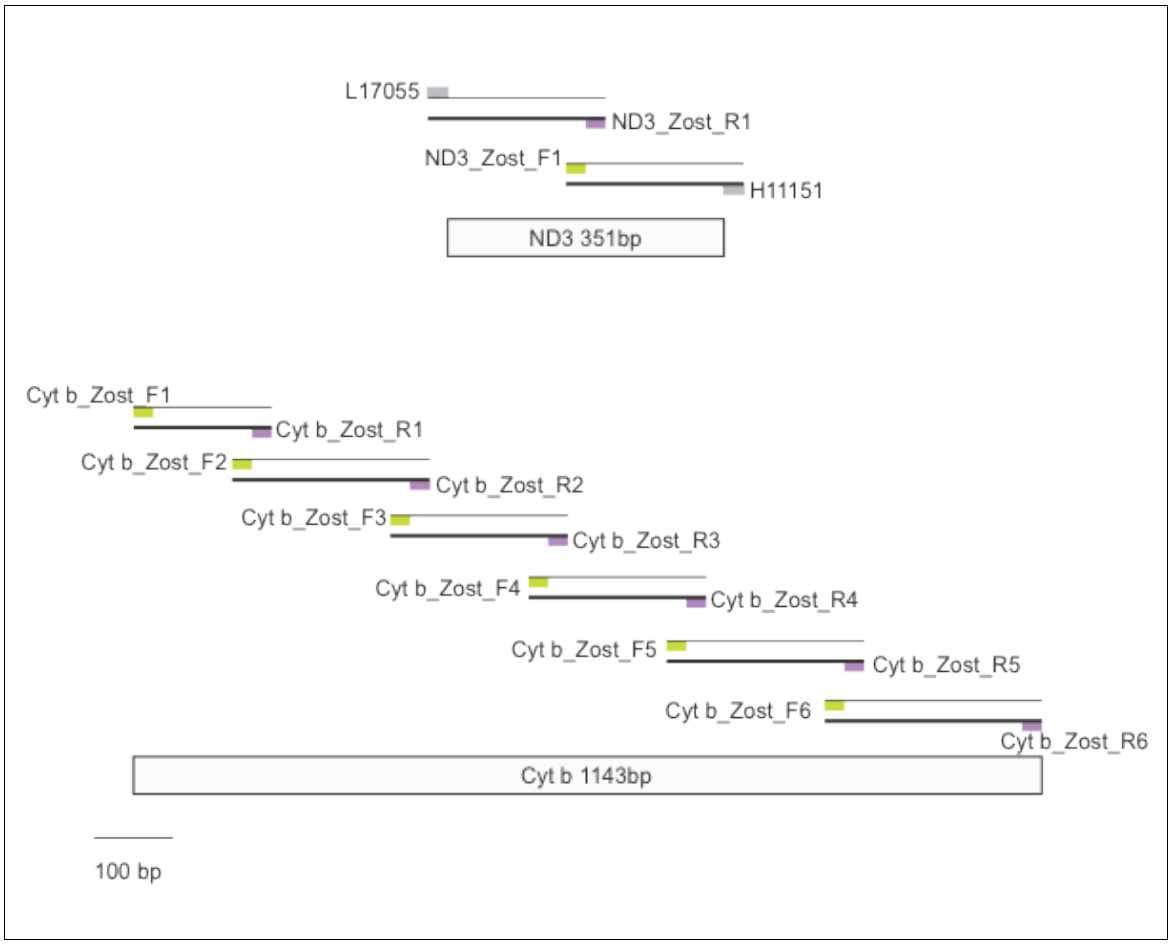


Figure 4.2 Series of overlapping fragments and associated primers that allow to for the amplification of the two mitochondrial genes Cyt b and ND3 genes from museum material.

Table 4.2 Primers used in the amplification of DNA from museum material

Gene	Primer Name	Primer sequence	Reference	PCR			
				Length	Tm	G-C content	GC clamp
ND3	L10755	5'-GACTTCCAATCTTTAAAATCTG-3'	Chesser 1999	22 bp	52.64°C	31.82%	
	H11151	5'-GATTTGTTGAGCCGAAATCAAC-3'	Chesser 1999	22 bp	60.84°C	40.91%	
	ND3_Zost_F1	5'-CCATTCTCAATCCGTTTCTTCCT-3'	Cox unpublished	23 bp	62.77°C	43.48%	
	ND3_Zost_R1	5'-GTCGAATAGTAGGAATAGGATTGC-3'	Cox unpublished	24 bp	57.52°C	41.67%	Y
Cyt b	Cyt b_Zost_F1	5'-GCTCTCAATCTTCGTAAAAACCAC-3'	Cox unpublished	24 bp	60.51°C	41.67%	
	Cyt b_Zost_F2	5'-ACACACTACACAGCAGACACCA-3'	Cox unpublished	22 bp	59.31°C	50.00%	
	Cyt b_Zost_F3	5'-TAGGAGTTGTCCTTCTACTAGC-3'	Cox unpublished	22 bp	52.44°C	45.45%	Y
	Cyt b_Zost_F4	5'-CCGATTCTTTGCCCTACACTTC-3'	Cox unpublished	22 bp	62.16°C	50.00%	
	Cyt b_Zost_F5	5'-CGTAGCTCTATTCGCTCCAAAC-3'	Cox unpublished	22 bp	60.40°C	50.00%	
	Cyt b_Zost_F6	5'-GTTACACACGTCCAAGCTACGT-3'	Cox unpublished	22 bp	59.24°C	50.00%	
	Cyt b_Zost_R1	5'-GTGAGCTACTGAAGCAAAGGCTA-3'	Cox unpublished	23 bp	60.22°C	47.83%	
	Cyt b_Zost_R2	5'-CTACGAATGCGGTAGCCATTAG-3'	Cox unpublished	22 bp	60.97°C	50.00%	
	Cyt b_Zost_R3	5'-CTAATGTGAGCCCTGCGATTAC-3'	Cox unpublished	22 bp	60.99°C	50.00%	
	Cyt b_Zost_R4	5'-AATTTTCTGGGTCCCCTAGTAG-3'	Cox unpublished	22 bp	57.75°C	45.45%	
	Cyt b_Zost_R5	5'-AGAGAGGTCGGAAGGTCATTGA-3'	Cox unpublished	22 bp	62.41°C	50.00%	
	Cyt b_Zost_R6	5'-TTATAGTTTGAGTAGTTTATTTTCTAGGA-3'	Cox unpublished	28 bp	54.59°C	24.14%	Y

The primer-designing program Primer 3 version 0.4 was used to check primer melting temperature (T_m), GC content and the presence of palindromes (sequence regions that may be read the same way in either direction), and hairpin loop structures (structures formed by the complementary binding of regions along a primer sequence). Palindromic and self-complementary regions in primer sequences can cause "primer dimers" to form, where the primer anneals with itself (homodimers) or other primers (heterodrimers) (Dieffenbach *et al.* 1993; Singh *et al* 2000). The generation of spurious products uses up reagents within the reaction mix which leads to a decrease in amplification efficiency and thus amplification yield. To guard against mispriming, the program Primer3-BLAST was used to test for significant homology between candidate primers and non-target regions. Primer sequences used in this study are listed in Table 4.2 and their locations in the gene regions are shown in Fig. 4.2

4.4.5. Tissue sampling

For the purpose of DNA analysis, the majority of preserved avian specimens are sampled by partially cutting off toe-pads (Mundy *et al.* 1997). While toe-pad sampling causes a degree of structural loss, especially in small birds (Payne and Sorenson 2002; Mann 2007), PCR success has proven to be significantly better with DNA from toe-pads compared to body skin and feathers (Mundy *et al.* 1997; Töpfer *et al.* 2011). Specimen preservation has included the use of a variety of chemical reagents (e.g. arsenic), which are thought to increase DNA degradation (Töpfer *et al.* 2011) and act as inhibitors in PCR reactions (Mundy *et al.* 1997). During preparation of museum skins the feet of birds are not always chemically treated, therefore tissue samples from the feet, often provide higher yields of DNA and exhibit lower levels of degradation when compared to other sampling areas (body skin and feathers) (Mundy *et al.* 1997; Töpfer *et al.* 2011).

In order to minimise damage to the collection, this study used a small tissue sample from the proximal phalanx (large pad) on the hind digit as the source for DNA extraction (Fig. 4.3), which is shown to contain minimal taxonomic characters (Mundy *et al.* 1997; Mann 2007). Tissue samples were taken from each specimen using sterile scalpels and forceps. Specimens were

sampled individually on a clean sheet of paper. The paper and equipment were exchanged between specimens.

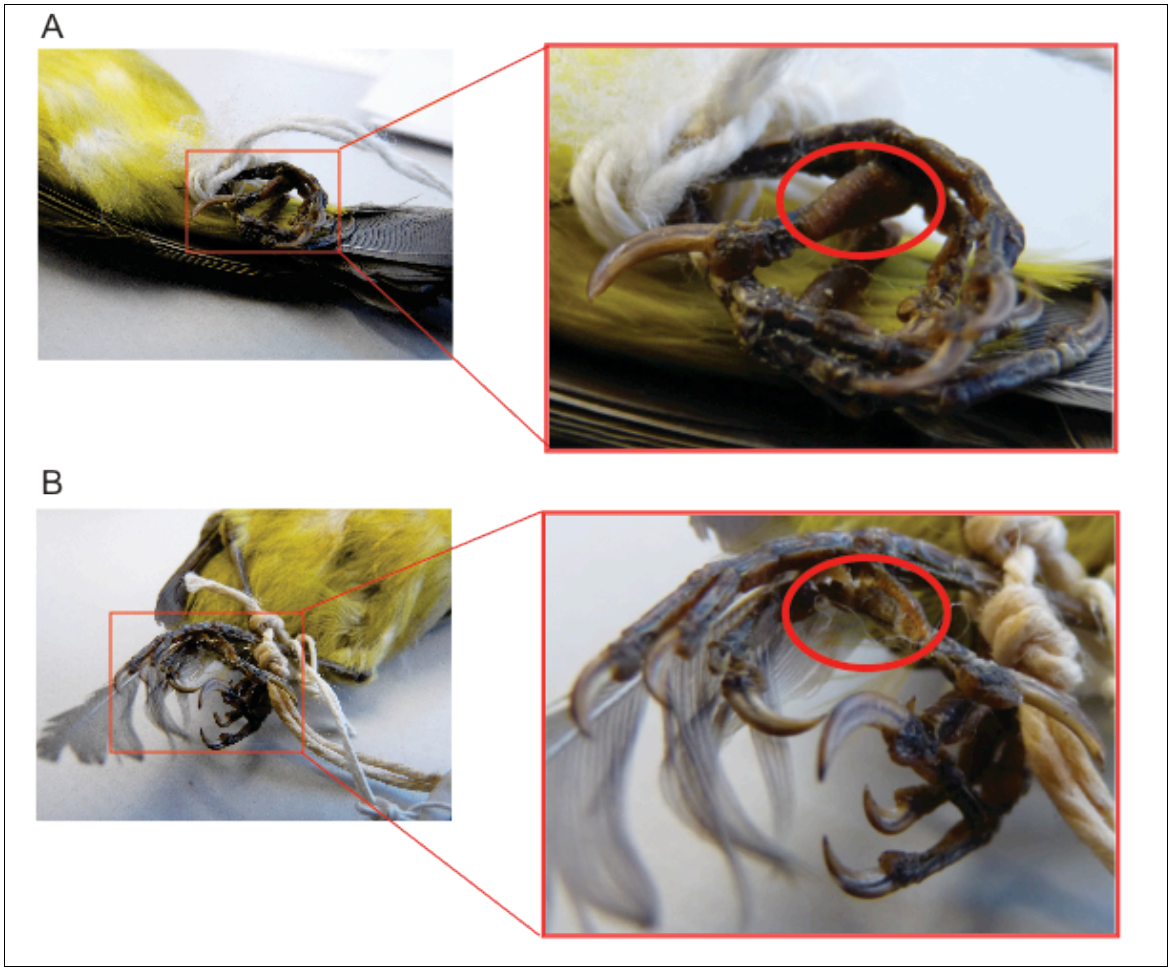


Figure 4.3. Sampling of museum specimens. A: proximal phalanx of specimen prior to sampling **B:** Damage to specimen after removal of the proximal phalanx.

4.4.6. DNA extraction

Prior to DNA extraction, dried toe-pad samples were soaked in ddH_2O for 30 minutes to re-hydrate. Whole genomic DNA was extracted from tissue samples using the QIAamp DNA Micro kit (Qiagen) that has been optimised for purifying DNA from very small and/or degraded samples. Adjustments to standard protocol included the addition of 20µl DTT (DL-Dithiothreitol) during tissue lysis. DTT is used to stabilise enzymes and aids the breakdown of disulfide bonds. This makes the digestion of proteins more efficient and speeds up digestion time (Nagai *et al.* 1998). Two additional incubation steps were also included. Following the addition of buffer AL, lysates (digested products) were incubated at 72°C for 10 minutes to ensure optimal binding of DNA to the spin column membrane. Prior to elution of DNA, buffer AE was added to spin columns and incubated at room temperature for 15 minutes. This ensured that

the elution buffer was fully absorbed into the silica-gel membrane, ensuring the highest possible yields of DNA. The extractions and amplification reactions were performed in rooms dedicated to working with old material (sterile environment free of DNA), with appropriate facilities such as a UV-bench used for sterilising equipment.

4.4.7. Generation of mitochondrial sequence data

PCR amplifications were performed using puReTaq Ready-To-Go PCR Beads (GE Healthcare). Beads contain buffers, dNTPs, enzyme, stabilisers, and BSA, all of which had been pre-treated to minimise contamination. Single beads were combined with 21µl ddH_2O, 0.5µl of each primer (10mM) and 3µl of template DNA giving a final reaction volume of 25µl. A hot-start touchdown PCR approach was used, where annealing temperatures for the first cycles were generally 1–2 °C below the T_m of primers. Given that primers were designed to have similar T_m's, all reactions were run under the same thermal cycling conditions. This program started with an initial denaturation at 95°C for 4 minutes, followed by six cycles of 95°C for 30 seconds, 58-53°C for 30 seconds and 72°C for 30 seconds, where the annealing temperature was lowered to 53°C in two-cycle increments. A further 34 cycles consisting of 95°C for 30 seconds, 51°C for 30 seconds and 72°C for 30 seconds were performed, followed by a final extension at 72°C for 9 minutes. PCR products were electrophoresed on a 1.5% agarose gel and visualised under ultra-violet light.

Purification of amplified PCR products was performed using an ExoSAP-IT PCR Clean-up (GE Healthcare). ExoSAP-IT (5µl) was added directly to PCR products and incubated at 37°C for 30 minutes. Inactivation of ExoSAP-IT enzymes was performed by heating to 80°C for 15 minutes. Cycle sequencing reactions were performed in 20µl volumes using 1µl BigDye™ Terminator (PE Applied Biosystems), 4µl ABI sequencing buffer, 1µl primer (1.6µM), 2µl of the purified PCR product and 12µl of ddH_2O. Cycle sequencing reactions consisted of an initial denaturation at 96°C for 1 minute, followed by 25 cycles of 95°C for 30 seconds, 50°C for 15 seconds and 60°C for 4 minutes. Cycle sequencing products were purified using a DyeEX 96 kit cleanup (Qiagen) following standard protocols. To ensure the accuracy of amplification of the ND3 and Cyt b genes, both the heavy and light strands were sequenced using an ABI 3730 DNA analyser (Applied Biosystems).

4.4.8. Sequences and alignment

For each individual, multiple sequences were obtained as a result of sequencing with several different primers. Valid sequences were considered to be clear DNA sequence reads, with no specimen ambiguities that could be aligned with control DNA sequences. The program Sequencher version 4.8 was used to check chromatograms for each primer pair before producing contigs of complementary fragments. To ensure that amplified fragments represented target regions and not nuclear copies of mitochondrial genes (numts), contig's were checked to ensure that they contained no gaps, insertions or deletions. In addition, sequences were translated into amino acids using the vertebrate mitochondrial translation table in MacClade version 4.08a to check they contained no stop codons. These sequences were then assembled to produce two consensus mtDNA gene sequences (ND3 and Cyt b) for each individual. Consensus sequences were aligned in Clustal W version 1.83 using default parameters with the resulting alignment checked by eye in the program SE-AL version 2.0.

4.4.9. Phylogenetic analysis

Variation in base composition for both genes was assessed using the X^2 test of homogeneity, implemented in PAUP (Swofford 2003). To account for differences in evolutionary processes experienced by the different sites in the alignment, PartitionFinder version 1.0 (Lanfear *et al*. 2012) was used to select the best-fit partitioning scheme and models of molecular evolution across all possible partitioning strategies. Branch lengths were estimated independently for each partition (unlinked). Model selection was limited to those available in MrBayes and the best scheme was calculated according to the Greedy algorithm using the Alkaike Information Criterion (AIC). Splitting the third codon for Cyt b and ND3 from all other sites was recovered as the best partitioning strategy and the GTR+I+G model was selected as the best model of sequence evolution across all partitions.

Bayesian analysis was conducted using MrBayes 3.1 (Huelsenbeck and Ronquist 2003) as implemented in the CIPRES portal using the models of evolution and partitioning strategies recommended by PartionFinder version

1.0. Base frequencies were estimated and evolutionary rates were allowed to vary across partitions under a dirichlet prior. Eight simultaneous Metropolis-coupled Markov Chain Monte Carlo (MCMCMC) chains were run for five million generations, starting from random trees, sampling every 100 generations with a heating parameter of 0.4. Stationarity of the Markov process was evaluated using average split frequencies (<0.05) and convergence of MCMC chains was assessed graphically in the program TRACER version 1.4.1 (Drummond and Rambaut 2007). A burn-in of 25% was applied, with the final tree constructed from 37500 post burn-in trees. Support is assessed using Bayesian posterior probabilities (BPP).

4.4.10. Generating an ultra-metric tree

Coalescent analysis requires an ultrametric tree, which has equal root-to-tip path lengths for all lineages. In this study an ultrametric tree was generated from the concatenated dataset using Bayesian methods implemented in the program BEAST version 1.48 (Drummond and Rambaut 2007). A log-likelihood ratio test implemented in PAUP* version 4.0b10 failed to reject the null hypothesis of rate constancy, resulting in enforcement of the molecular clock.

Starting from the BI tree, two independent Markov Chain Monte Carlo (MCMC) analyses were run for 2,000 000 generations, using a constant rate Yule speciation prior (assumes a constant speciation rate per lineage). Analyses were run using the GTR+I+G model of molecular evolution where the number of gamma categories was set to 6. Trees and corresponding parameters were sampled every 1,000 generations with a burn-in of 10%. Convergence of the two independent MCMC runs was assessed graphically in TRACER version 1.4.1 (Drummond and Rambaut 2007). A consensus tree was obtained from the post burn-in tree sample (rejecting the first 10%) using TREE ANNOTATOR version 1.4.8 (Drummond and Rambaut 2007). The posterior probability threshold was set to 0.5, the target tree type was set to maximum clade credibility, and the heights of the nodes were retained.

4.4.11. General mixed Yule-coalescent (GMYC): Multi-model inference and model averaging approach

GMYC analyses attempt to measure the degree of genetic clustering with the goal of delimiting independent evolutionary clusters, which can be used to

infer species boundaries (Powell *et al.* 2011). By estimating lineage birth rates associated with speciation events (Yule diversification) and coalescent processes within the tree, the GMYC method calculates the likelihood of speciation-coalescent transitions at each node within the phylogeny (Pons *et al.* 2006). The recently modified method of Powell *et al.* (2011) accounts for model uncertainty by using multi-model inference and model averaging based on information-theoretic approaches. This approach assigns weights and ranks to models based on their contribution to the estimation of parameters and their ability to account for variation in the data. Unlike the single and multiple threshold approaches, this method uses model averaging to estimate transition boundaries for each node and provides confidence estimates associated with these boundaries. GYMC analysis was run on the ultrametric tree that was generated in BEAST, using the SPLITS package (available from http://r-forge.r-project.org/projects/splits/) in the program R 2.10.1 (R Core Development Team, 2009). The multi-model inference and model averaging approach utilised the modified GMYC source code of Powell *et al.* (2011).

4.5. Results

4.5.1. Sequence data

A total of 74 specimens were extracted, resulting in 592 amplification reactions being sequenced. No amplification of non-target DNA (i.e. bacteria) was detected, although several sequence reads were very short or messy. Several of these sequences were consistent with certain toe-pad samples and were interpreted as sequence reads from severely degraded DNA and rejected from further analysis. The remaining 420 sequenced fragments, representing 64 individuals (86% of specimens sampled), were concatenated to produce full sequence reads for each target gene. Primer fidelity across taxa was not always consistent. Consequently, the concatenated gene sequences for some specimens have gaps or slightly truncated sequence lengths (Appendix III). The museum specimen dataset was supplemented with sequence data for 115 individuals whose DNA was previously obtained from blood samples extracted in Chapter 2, in addition to 39 individuals whose sequence data was obtained from the NCBI database. In total, sequence data was obtained for 218 individuals providing a 2056 base pairs (bp) dataset (TGFß2- 25 individuals (582bp); ND3- 214 individuals (347bp); Cyt b- 207 individuals (1115bp)).

4.5.2. Phylogenetic analysis

The Bayesian topology generated from the concatenated Cyt b, ND3 and TGFß2 dataset is generally well supported, and is largely congruent with the mtDNA phylogeny produced in Chapter 2. Extensive sampling across Sub-Saharan Africa, Indian Ocean region, southern Arabian Peninsular and the Gulf of Guinea region has revealed that the Africa Zosteropidae complex is comprised of six major clades (Fig. 4.4; clades A – F). However, while there is good support for relationships within these six clades, resolution of more broad scale relationships between these clades, specifically between clades C, D, E and F, is poorly.

Members of *Z. abyssinicus* form three independently well-supported clades that are polyphyletic with respect to each other. Four of the six currently recognised subspecies form clade D (BPP=1.00). This clade contains the nominate subspecies *Z. a. abyssinicus* that is distributed throughout north-eastern Sudan, Eritrea and northern and central Ethiopia, insular *Z. a.*

socotranus from the island of Socotra in the Gulf of Aden, *Z. a. omoensis* occurring in western Ethiopia and *Z. a. arabs* that is found throughout the southern tip of the Arabian peninsula (Saudi Arabia, Yemen and southern Oman).

Within Clade D the nominate subspecies *Z. a. abyssinicus* is non-monophyletic. Sample BMNH 1952 32 3 from Eritrea is recovered at the base of clade D (BPP=1.00), while samples BMNH 1927 11 5 577 and BMNH 1915 12 24 1198 from Ethiopia and Sudan respectively form a clade with *Z. a. arabs* (BPP=0.98). The monophyly of subspecies *Z. a. omoensis* and *Z. a. arabs* is well supported (BPP=1.00 and 0.99, respectively), with results placing *Z. a. omoensis* as sister to the clade containing *Z. a. arabs* and *Z. a. abyssinicus* (BPP=1.00).

There is strong support for the division of insular members of *Z. a. socotranus* from the clade containing *Z. a. omoensis*, *Z. a. arabs* and *Z. a. abyssinicus* (BPP=0.99), however mainland members of *Z. a. socotranus* (northern Somalia) fall outside clade D, rendering this subspecies non-monophyletic. Two samples of mainland *Z. a. socotranus* from northern Somalia forms clade B (BPP=1.00), which nests between the AIO *Zosterops* (clade A) and all other major African lineages (BPP=0.60). The two remaining *Z. abyssinicus* subspecies, *Z. a. jubaensis* and *Z. a. flavilateralis*, that have a parapatric distribution throughout East Africa (Ethiopia, Kenya and Tanzania) are recovered as a monophyletic group (BPP=0.99) within clade E, with no support for any division between these two 'subspecies'.

All *Z. pallidus* samples are recovered into two independent clades within the major clade E. Results demonstrate maximum support (BPP=1.00) for a clade containing samples of *Z. p. pallidus* and *Z. p. sundevalli* that are found throughout Namibia, western and central South Africa. Resolution of branching patterns within this clade is generally poor and there is no support for division of the two subspecies. The remaining *Z. pallidus* subspecies (*Z. p. capensis*, *Z. p. atmorii*, *Z. p. virens* and *Z. p. caniviridis*) that have a broad distribution throughout southeast Africa, (South Africa, Botswana and Mozambique) form a single clade that includes three samples of *Z. senegalensis* sampled from different southern African localities (BPP=0.78). Resolution of relationships within this clade is poor, and there is no support for the monophyly of sub-species.

Widespread sampling for members of *Z. senegalensis* confirms the polyphyletic nature of this species, with members found in two of the six major clades (E and F). Within clade E, resolution of relationships between members of *Z. senegalensis,* those distributions is throughout southern Africa (*Z. s. quanzae*, *Z. p. kasaicus*, *Z. p. heinrichi* and *Z. p. anderssoni*), is generally poor. Analysis supports a close relationship between southern African *Z. senegalensis* and members of *Z. pallidus* (Oatley *et al.* 2012), with two *Z. senegalensis* clades recovered as sister to a clade containing *Z. pallidus* and *Z. senegalensis*. However, analysis fails to resolve branching patterns between the three clades and thus relationships between these southern African forms remains ambiguous.

Within major clade F, members of *Z. senegalensis* form three independently well-supported clades. Samples of *Z. s. stenocricotus* from Mt Cameroon in West Africa receive maximum clade support (BPP=1.00) and are recovered as sister to *S. brunneus* from the island of Bioko in the Gulf of Guinea (BPP=0.98). The second clade contains an assemblage of central African *Z. senegalensis* (DRC, Uganda and Tanzania). This clade includes: *Z. s. reichenowi* from eastern DRC; *Z. s. toroensis* from northeast DRC and western Uganda; *Z. s. stuhlmanni*, that has a distribution from southern and central Uganda towards Tanzania; and *Z. senegalensis* samples from the DRC (not identified to the sub-specific level).

This central African *Z. senegalensis* clade is strongly supported (BPP=1.00) and contains a significant degree of genetic structure. While the independent monophyly of subspecies *Z. s. toroensis* and *Z. s. reichenowi* is strongly supported (BBP=1.00, BBP=1.00), the subspecies *Z. s. stuhlmanni* is recovered as non-monophyletic. Sample BMNH 1934 1 17 27 from central Uganda recovered as sister to two *Z. s. reichenowi* samples (BPP = 0.99), while samples BMNH 1913 7 16 140 and NRM 552125 from western Uganda and northern Tanzania respectively, form a clade with the three *Z. senegalensis* samples from the DRC (BPP=0.99). Analysis recovers *Z. s. toroensis* as sister to a clade containing *Z. s. stuhlmanni* and *Z. senegalensis*, although support for this relationship is poor (BPP = 0.50).

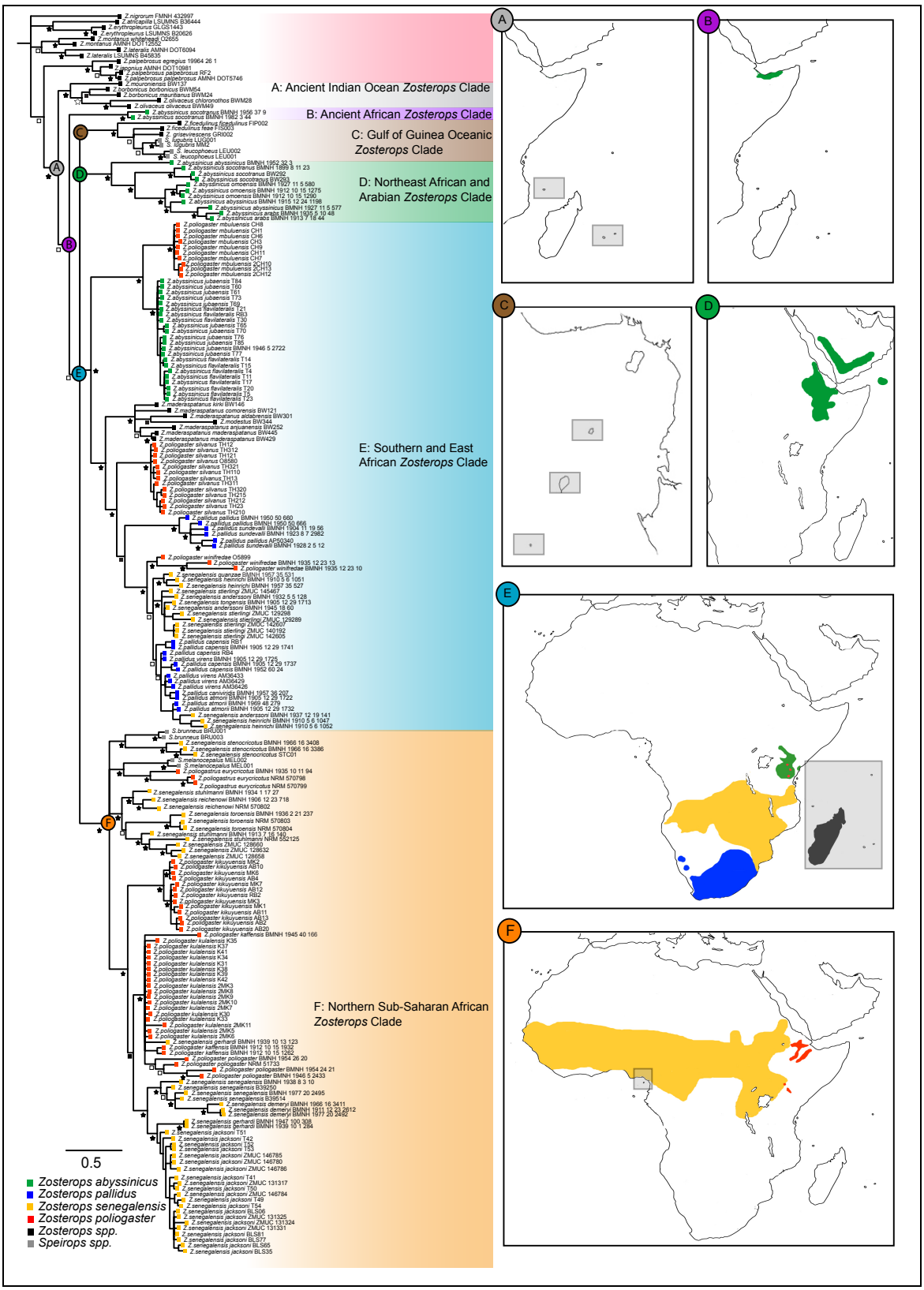


Figure 4.4. Bayesian inference (BI) tree of African Zosteropidae Branch lengths are proportional to the degree of sequence divergence. Bayesian posterior probabilities (BPP) are displayed below branches. ★ Indicates nodes with >95% BPP, ☆ indicates nodes with >90% BPP, ■ indicates nodes with >80% BPP and □ indicates nodes with > 50% BPP. Nodes with < 50% BPP are not shown. Labelling of distribution maps corresponds to key clades, which are labelled A-F. Taxa are labelled using full trinomial nomenclature, following the taxonomy of Dickinson (2003).

The most northerly members of *Z. senegalensis* form the third assemblage within clade F. This group contains the nominate subspecies *Z. s. senegalensis,* that has a wide distribution across much of northern Sub-Saharan Africa, *Z. s. demeryi* that is found in western Africa (Sierra Leone, Liberia, and Ivory Coast), *Z. s. gerhardi* from southern Sudan and northeast Uganda and *Z. s. jacksoni* from Kenya and northern Tanzania.

Samples of *Z. s. senegalensis* and *Z. s. demer*yi form a strongly supported monophyletic group (BPP=1.00) that is recovered as sister to a clade containing *Z. s. jacksoni* and *Z. s. gerhardi* (BPP=1.00). There is good support for the monophyly of subspecies *Z. s. demeryi* (BPP=0.99), *Z. s. gerhardi* (BBP=1.00) and *Z. s. jacksoni* (BBP=0.84), however the nominate subspecies *Z. s. senegalensis* is recovered as paraphyletic. Considerable structure is noted within *Z. s. senegalensis* and *Z. s. jacksoni,* although analysis fails to provide resolution of relationships.

For *Z. poliogaster*, results are largely concordant with the molecular phylogeny of Chapter 2, supporting the non-monophyly of *Z. poliogaster*. Clade E contains three endemic montane *Z. poliogaster* subspecies: *Z. p. mbuluensis* from the Chyulu Hills (southern Kenya); *Z. p. silvanus* from the Taita Hills (southern Kenya); and *Z. p. winifredae* from the South Pare Mountains (northern Tanzania). The independent monophyly of all three subspecies received maximum support (BPP=1.00), and their phylogenetic placement is congruent with the topology generated in Chapter 2. Samples of *Z. p. mbuluensis* are recovered as sister to a clade containing *Z. a. jubaensis* and *Z. a. flavilateralis* (BPP=1.00), while *Z. p. silvanus* nests between the Indian Ocean *maderaspatanus* white-eyes and a clade containing *Z. p. pallidus* and *Z. p. sundevalli* (BPP=0.99). Increased sampling has improved support for the placement of *Z. p. winifredae*, which is recovered at the base of a clade containing *Z. pallidus* and southern African *Z. senegalensis* samples (BPP=1.00).

Extensive sampling for *Z. poliogaster,* including subspecies that were absent from previous molecular analysis (*Z. p. poliogaster, Z. p. kaffensis*, *Z. p. eurycricotus*) reveals that all other *Z. poliogaster* subspecies fall into three independent clades within the major clade F. Samples of *Z. p. eurycricotus* from Mount Meru in northern Tanzania received maximum support (BPP=1.00) and

are recovered as sister to *S. melanocephalus* from Mt Cameroon in West Africa (BPP=0.95).

Samples of *Z. p. kikuyuensis* (BPP=1.00) from Mt Kenya and the Aberdare Range (central Kenya) form the second strongly supported monophyletic group and are recovered as sister to a large clade containing *Z. poliogaster* from the most northerly part of the range (*Z. p. poliogaster, Z. p. kaffensis and Z. p. kulalensis*) and northern Sub-Saharan members of *Z. senegalensis* (*Z. s. senegalensis, Z. s. demeryi, Z. s. gerhardi, Z. s. jacksoni*). The placement of *Z. p. kikuyuensis* in relation to *Z. p. kulalensis* conflicts with the in molecular phylogeny generated in Chapter 2, which alternatively places *Z. p. kulalensis* as basal to *Z. p. kikuyuensis*.

The three most northerly members of *Z. poliogaster* species complex (*Z. p. poliogaster, Z. p. kaffensis* and *Z. p. kulalensis*) form a well-supported clade with a single sample of *Z. s. gerhardi* that was collected from Lomoling in the Imatong Mountains in southern Sudan (BBP=0.98). This clade contains the nominate subspecies *Z. p. poliogaster*, that occurs throughout Eritrea, the Ethiopian highlands and in isolated mountains within Southeast Sudan, *Z. p. kaffensis* found in Southwest Ethiopia and *Z. p. kulalensis* that is endemic to Mt Kulal in northern Kenya. In contrast to other *Z. poliogaster* subspecies there is no support for the independent monophyly of *Z. p. kaffensis* or *Z. p. kulalensis* and only weak support for the monophyly of *Z. p. poliogaster* (BPP=0.55).

In agreement with Indian Ocean-centred phylogeny of Warren *et al.* (2006), all Indian Ocean *Zosterops* are recovered into two independent assemblages commonly referred to as the Ancient Indian Ocean (AIO) white-eyes and the Indian Ocean *Maderaspatanus* (IOM) clade. Results provide strong support from the monophyly of both clades (AIO, BPP=0.96; IOM, BPP=1.00) and the placement of taxa within these clades is largely congruent with previous the molecular studies (Warren *et al.* 2006; Melo *et al.* 2011). In agreement with the topology produced in the phylogenetic analyses of Chapter 2, the IOM white-eyes that are distributed throughout Madagascar, the Comoros and the Seychelles nests within the major clade E (BPP=1.00). Analysis also provides improved support for the placement of the AIO white-eyes (Grand Comoro and the Mascarenes) that are recovered at the base of all other African Zosteropidae (BPP=0.97)

Within the Gulf of Guinea Zosteropidae system, results provide additional support for the non-monophyly *of Speirops* (Melo *et al.* 2011), with members of this genus occurring in two of the six major clades (C and F). However, in contrast to the molecular phylogeny of Melo *et al.* (2011) that places all Gulf of Guinea Zosteropidae in two independent assemblages (GGM, Gulf of Guinea mainland; GGO, Gulf of Guinea Oceanic), results spilt Gulf of Guinea white-eyes into three independent clades. In contrast to previous topologies (Melo *et al.* 2011; Chapter two) which recovered *S. melanocephalus*, *S. brunneus* and *Z. s. stenocricotus* as a monophyletic group (GGM), results split *S. melanocephalus* from *S. brunneus* and *Z. s. stenocricotus* and place it as sister to *Z. p. eurycritus* from East Africa (BPP=0.93). The two newly formed clades (*S. brunneus* and *Z. s. stenocricotus*; *S. melanocephalus* and *Z. p. eurycritus*) are both recovered in the major clade F, however lack of resolution at the base of this clade means that it is not possible infer relationships between these two groups.

Relationships within the GGO clade are in agreement with the topology produced by Melo *et al.* (2011); *Z. f. ficedulinus* is recovered at the base of the GGO clade (BBP=1.00) and is clearly divergent from *Z. f. feae* rendering Z. *ficedulinus* non-monophyletic. The relationship between *Z. f. feae* and *Z. griseovirescens* is unclear although both are placed between *Z. f. ficedulinus* and a clade containing *S. lugubris* and *S. leucophaeus.* The sister relationship between the two *Speirops* species is only weakly supported (BPP=0.54), although there is strong support for the monophyly of both species (BPP= 0.95 and BBP=1.00, respectively).

4.5.3. GMYC analyses

Support for genetic clusters predicted by this method varied across the tree. While genetic cluster probabilities (neutral coalescence and Yule diversification) were strongly supported at the base of the tree and within some of the terminal taxa (P>0.80 tips and P<0.15 base), branching relationships in much of the tree yielded intermediate probabilities between 0.5 and 0.3. Where there is no support for neutral coalescence of samples within lineages (P<0.85), GMYC coalescence estimates are interpreted as evidence of variation (or genetic variability) that makes daughter lineages distinct from one another. However, in the absence of strong support for Yule diversification (P>0.15) this

study considers daughter lineage to not contain sufficient genetic variation to warrant them 'distinct evolutionary units'. Within the African Zosteropidae system GMYC analysis provided strong support (<0.15) for Yule diversification of daughter lineages at 10 nodes (Fig. 4.5; Node 1-10), leading to the resolution of 14 distinct evolutionary lineages (Fig. 4.5; L1-L14). Although several coalescence clusters were detected within these lineages, the majority of samples were recovered as poorly supported singletons (<0.80 P >0.15) suggesting a certain degree of genetic variation within these 13 distinct assemblages.

The AIO white-eyes (Clade A) are recovered into two distinct lineages (L1-L2). The node connecting *Z. mouroniensis* from (Grande Comore) to a clade containing *Z. borbonicus* and *Z. olivaceus* (Mascarenes), provides support for Yule divergence of daughter lineages (node 1, P= 0.145), indicating that *Z. mouroniensis* is a distinct evolutionary unit relative to other members the AIO white-eyes clade. Mainland *Z. a. socotranus* (Clade B) is also recovered as an independent lineage (L3), with strong support for Yule diversification of daughter lineages at node 2 (P= 0.011). GMYC estimates place all GGO white-eyes (Clade C) in lineage L4. All GGO *Zosterops* sampled are recovered as poorly supported singleton, with nodes corresponding to intermediate coalescence estimates (0.301≥ P ≤0.334). In contrast, the two GGO Speirops species sampled (*S. leucophaeus* and *S. lugubris*) are recovered into two independent coalescent clusters (P=0.992, respectively).

Northeast African and Arabian *Z. abyssinicus* forms that make up clade D (*Z. a. abyssinicus,* insular *Z. a. socotranus*, *Z. a. omoensis* and *Z. a. arabs*) are recovered as distinct from all other African Zosteropidae in lineage L5. Within this group GMYC analysis recovers two coalescence clusters and five singletons. While samples of *Z. a. omoensis* and insular *Z. a. socotranus* are supported as neutral coalescent clusters (P=0.99 and P=0.80, respectively) samples of the monophyletic subspecies *Z. a. arabs* and the paraphyletic subspecies *Z. a. abyssinicus* are recovered as independent singletons with intermediate coalescence estimates.

GMYC analysis recovers clade E (containing southern and East African *Zosterops* in addition to members of the IOM clade) into five distinct evolutionary lineages (L6-L10). The endemic montane *Z. p. mbuluensis* (Chyulu Hills, Kenya) and lowland *Z. a. flavilateralis* and *Z. a. jubaensis* (East Africa)

form the strongly supported lineage L6 (node 4; P=0.07). Within this group distinct neutral coalescence thresholds are clearly apparent. Samples of *Z. p. mbuluensis* are strongly supported as a single coalescence cluster (P=0.92), while *Z. a. flavilateralis* and *Z. a. jubaensis* form a further three clusters (P=0.92). Clustering within this yellow-bellied *Z. abyssinicus* clade (*Z. a. flavilateralis* and *Z. a. jubaensis*) is not concordant with sub-specific divisions, providing no evidence for the two sub-species acting as separately evolving meta-populations. The IOM white-eyes form the second lineage (L7) within clade E, with good support for Yule diversification of daughter lineages at node 5 (P=0.071). Although the GMYC analysis provides strongly support for the grouping of two *Z. m. maderaspatanus* samples as a single cluster, coalescence probability estimates for other members of this lineage are not supportive of either neutral coalescence or Yule diversification processes.

Lineage L8 is exclusive to samples of the *Z. p. silvanus* (node 6 P=0.126), rendering this montane endemic distinct from all other African forms. The analysis groups members of *Z. p. silvanus* into two coalescence clusters (P=0.822 and P=0.821), although there is no support for these clusters as independent evolutionary units (P=0.272).

The remaining taxa within clade E are recovered in lineages L9 and L10. The two *Z. pallidus* subspecies *Z. p. pallidus* and *Z. p. sundevalli* form lineage L9 (node 7= 0.150), while all other *Z. pallidus* subspecies (*Z. p. capensis, Z. p. virens Z. p. atmorii* and *Z. p. caniviridis*) are placed in a lineage L10 with *Z. p. winifredae* and all southern African *Z. senegalensis* subspecies (*Z. s. heinrichi, Z. s. anderssoni, Z. s. kasaicus, Z. s. stierlingi, Z. s. quanzae* and *Z. s. tongensis*). Lineage L10 contains a significant degree of structure with GMYC analysis recovering eight coalescence clusters and ten singletons

Within clade F, GMYC estimates provide strong support for four independent lineages (L11-L14). A geographically disparate clade containing *S. melanocephalus* and *Z. p. eurycricotus* form the first lineage with strong support for Yule diversification at node 8 (P=0.07). Within this lineage *S. melanocephalus* is recovered as a single coalescence clusters while samples of *Z. p. eurycricotus* are recovered as a cluster of two samples and a singleton. The second strongly supported (node 9 P=0.08) lineage within clade F (L12) contains central African *Z. senegalensis* samples (node 9 P=0.08). Although the GMYC analysis recovers the three *Z. s. toroensis* samples and two DRC

samples into two independent coalescence clusters, probability estimates for other members of this lineage are not supportive of either neutral coalescence or Yule diversification processes. The remaining taxa within clade F are recovered in lineages L13 and L14 (node 10 P=0.09). Lineage L13 contains *S. brunneus* and *Z. s. stenocricotus* from West Africa, while L14 contains northerly members of *Z. poliogaster* and *Z. senegalensis*. Analysis recovers a significant degree of genetic structure within L14 and recovers 14 coalescent clusters and 20 singletons.

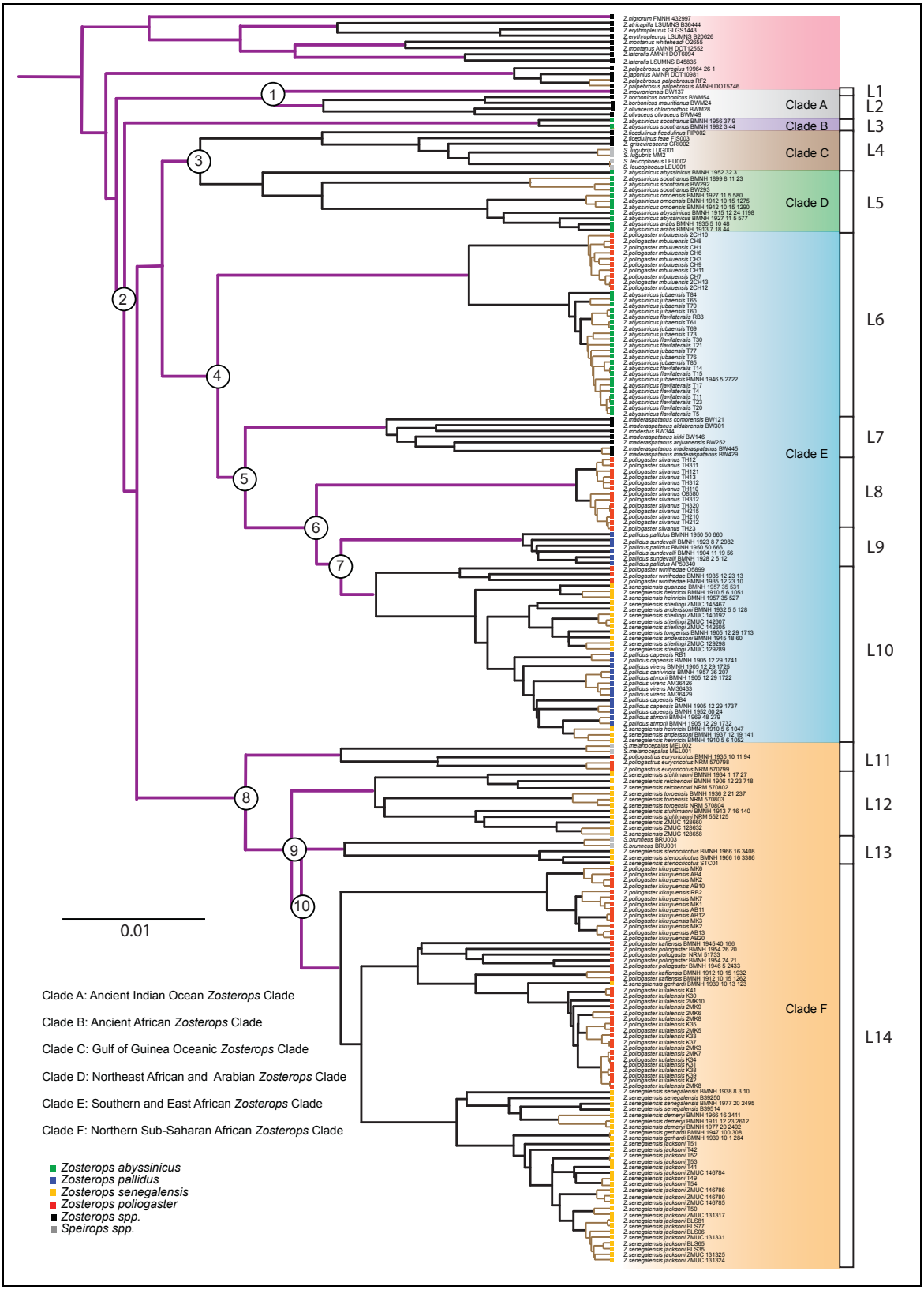


Figure 4.5 Coalescent probability estimates for African Zosteropidae. Generated using the GMYC multi-model inference and model averaging approach of Powell *et al.* (2012). Transition in branch colour from purple to brown across the tree indicates the transition from Yule diversification (purple) to neutral coalescent (brown) processes. Purple nodes indicates genetic cluster probabilities P≤0.15, black nodes indicates 0.15<P>0.80 and brown indicates P≥0.80. Key nodes are labelled 1-10. Key lineages are labelled L1-14 and corresponding nodes 1-10. Taxa are labelled with full trinomial nomenclature, following the taxonomy of Dickinson (2003), followed by either the museum specimen catalogue number, or collection code.

4.6.Discussion

4.6.1 Molecular phylogeny

Generating broad species-level phylogenies in highly speciose groups such as Zosteropidae can be particularly difficult, as broad geographic distributions makes extensive sampling for fresh material practically impossible. Unprecedented sampling of African Zosteropidae using both archive and fresh material has enabled the largest genetic assessment of mainland African Zosteropidae to date. The use of museum collections for DNA extraction has allowed for dense sampling of all four mainland African *Zosterops* species, including all 34 currently recognised subspecies. Although DNA obtained from fresh specimens is always preferred to DNA from museum specimens, this study demonstrate the advantages of using such material in the absence of recent DNA collections.

Phylogenetic results confirm the non-monophyly of all mainland African *Zosterops* species, rendering the current taxonomic framework invalid. Broad-scale relationships are largely congruent with the previous findings of Chapter 2, yet additional sampling for members of *Z. abyssinicus* from northeast African and the African Peninsula have allowed for the identification of a further two assemblages (Clades B and D). In total the results provide strong support for six key clades (Fig. 4.4; Clade A-F) within the African Zosteropidae system (Continental Africa, Gulf of Guinea, Indian Ocean, Gulf of Aden and the Arabian Peninsula), although relationships between these clades are less well resolved. Nevertheless, despite poor resolution of branching patterns between these larger clades, genetic results provide strong support for relationships within these six assemblages. These results provide the first comprehensive framework of genetic relationships, which can used to generate predictions regarding the probable placement of species boundaries.

4.6.2. Interpretation of GMYC clusters

While previous GMYC estimates have been observed to correspond with existing species limits in some higher taxa (Monaghan *et al.* 2009), questions remain regarding whether these clusters represent species or taxa at a different hierarchical level (Powell 2012). This study follows Barraclough *et al.* (2009) in defining these clusters as ‘evolutionary significant units’ owing to the fact that

they represent independently evolving lineages (at least at the level of the gene loci). As previously pointed out by Powell (2012), the fact that these clusters exist and that they often correspond with biological or ecological characteristics (Jouselin *et al.* 2009; Powell 2009) of the taxa in which they are found, suggests that they are representative of some fundamental evolutionary process. Nevertheless, divergence of taxa is a continuous process and consequently intermediate coalescence estimates are open to interpretation.

In this study it is the relative degree of divergence from a neutral coalescence model that has been considered as important to the interpretation of diversity and structure within clades. Divergence from a neutral coalescence model is interpreted as evidence of variation (or genetic variability) that makes daughter lineages distinct from one another, representing a biological phenomenon that needs to be explained even if they are not representative of species *per se*. However, in the absence of strong support for Yule diversification (>0.15) this study considers daughter lineage to not contain sufficient genetic variation to warrant them 'distinct evolutionary units' (Barraclough *et al.* 2009) and therefore looks to morphological and ecological characters to determine the nature of relationships.

4.6.3. Lowland northeast African and Arabian forms

Molecular investigation into the relationships and taxonomic validity of lowland northeast African and Arabian forms revealed *Z. abyssinicus* to be a polyphyletic species. In contrast to Moreau (1957), who grouped all lowland northeast African and Arabian forms based on a shared 'dingy' or 'dusty' plumage colouration, this molecular work provides no support for the monophyly of *Z. abyssinicus*, with members recovered into three independently well-supported clades that are polyphyletic with respect to each other (Fig 4.4). These results suggest that the 'dusty' or 'dingy' plumage colouration that has previous been use to group lowlands forms (Moreau 1957) more than likely represents shared ecological conditions of taxa, rather than close genetic affinities (a least at the level of sampled loci).

The phylogenetic division of yellow-bellied forms (Clade E: *jubaensis* and *flavilateralis*) from grey-bellied forms (Clade D: *abyssinicus*, *omoensis*, *arabs*, and insular *socotranus*) may seem to provide support for a previous taxonomic arrangement that separated lowland East African and Arabian forms into two

separate species based on belly colour (Sclater 1930). However, the non-monophyletic placement of a grey-bellied Somali form (Clade B: mainland *socotranus*) relative to all other lowland grey-bellied taxa (Clade D) suggests that belly colouration is an unstable character within this lowland group. Furthermore, the non-monophyly of *Z. a. socotranus* is interpreted as evidence that the black beak shared by members of *Z. a. socotranus* is not a sound taxonomic character for grouping populations. Instead, results place black-beaked forms from the island of Socotra within a clade containing birds with brown beaks (Clade D), while mainland black-beaked forms from northern Somalia are recovered at the base of all other mainland taxa. This is not to say that plumage or beak colouration is not useful for distinguishing between neighbouring species or subspecies, only that grouping geographically distinct lowland populations based solely on these morphological characters is not supported by these molecular results.

Yule diversification estimations provide strong support for clade B (mainland *socotranus*) and clade D (*abyssinicus*, *omoensis*, *arabs*, and insular *socotranus)* as 'evolutionary significant units', highlighting them as independently evolving lineages (Fig 4.5). For yellow-bellied forms *jubaensis* and *flavilateralis* that are recovered as sister to the endemic montane *Z. p. mbuluensis* in clade E, GMYC estimates are less clear-cut. Although phylogenetically distinct, support for these lowland forms as an evolution distinct lineage relatively to *Z. p. mbuluensis* is lacking. Nevertheless, results do not place them in a coalescence cluster with *Z. p. mbuluensis,* which suggests that these lineages are distinct from one another, but not to the extent of clades B and D. Coalescence estimates provide no support for the division of subspecies *Z. a. jubaensis* and *Z. a. flavilateralis*, which may be unsurprising given that plumage variation across the range of *jubaensis* and *flavilateralis* has previously been shown to vary in accordance with altitude and rainfall (Moreau 1957).

Within clade D, the monophyly of insular *socotranus* and *omoensis* is well supported (Fig 4.4) and GMYC analyses recover both forms as independent coalescence clusters (Fig 4.5). In contrast, samples of the monophyletic subspecies *arabs* and the paraphyletic subspecies *abyssinicus* are recovered as independent singletons, which is interpreted as evidence that members of *arabs* and *abyssinicus* demonstrate a higher degrees of genetic structure than members of *omoensis* and insular *socotranus*. However, further

sampling within the sub-specific ranges of *abyssinicus* and *arabs* is necessary in order to obtain a more complete assessment of subspecies validity.

4.6.4. Highland northeast African forms

While the relatively larger size and rich plumage colouration of montane endemics in northeast Africa has led to the various highland forms to be treated as subspecies of a wider species complex (Dickinson 2003), molecular results consistently recover this species as polyphyletic (Warren *et al.*2006; Melo *et al.* 2011; Chapter 2). Supportive of Moreau (1957) previous morphological conclusions, results are interpreted to suggest that the large size and rich plumage colouration exhibited in these highland forms is likely to be a product of the high elevation and moist forest habitat they occupy rather than shared genetic affinities.

The placement of *Z. poliogaster* samples is largely concordant with the topology generated in Chapter 2. However, extensive sampling including *Z. poliogaster* subspecies previously absence, recovers *Z. p. kulalensis* samples endemic to Mt Kulal in northern Kenya as conspecific with more northerly forms (*Z. p. poliogaster*, *Z. p. kaffensis*, *Z. p. omoensis*) that are distributed throughout highland areas of Eritrea, Ethiopia and South Sudan. This is in contrast to more southerly member of *Z. poliogaster* (*Z. p. mbuluensis, Z. p. silvanus, Z. p. winifredae, Z. p. eurycricotus* and *Z. p. kikuyuensis*) that continue to be supported as independent non-sister clades.

Variation in belly colouration within this northerly *Z. poliogaster* clade is clearly evident, matched only by the variation demonstrated between southern African forms. Nevertheless, in the absence of support for subspecies monophyly the use of belly colour as a taxonomic character within this clade is not supported. This result demonstrates the pace at which phenotypic divergence can be observed within Zosteropidae, giving rise to divergent phenotypes with minimal sequence divergence.

While the majority of endemic montane forms occurring south of Nairobi (Kenya) are recovered in clade E, samples of *Z. p. eurycricotus* from Mt Meru in northern Tanzania nests within a clade containing northern Sub-Saharan African samples (clade F), and are recovered as sister to *S. melanocephalus* from Mt Cameroon in West Africa. Members of *Speirops* have never been thought to be conspecific with *Z. poliogaster*, being both geographically

segregated and phenotypically distinct. Despite this, support for the two forms as independent evolutionary units is lacking with the node connecting *Z. p. eurycricotus* and *S. melanocephalus* yielding intermediate coalescent estimates.

Within clade E, GMYC estimates provide strong support for *silvanus* (Taita Hills, Kenya) as an independent evolutionary unit (Fig 4.5). In contrast intermediate GYMC estimates (0.33) on the nodes separating highland *mbuluensis* and *winifredae* from their respective lowland sister clades are not supportive of either neutral coalescence or Yule diversification processes and consequently are open to interpretation. In both instances weaker Yule diversification support may simply reflect the fact that *mbuluensis* and *winifredae* have diverged more recently relative to *silvanus.* However, in the absent of hybrids (Moreau 1957) and with ecological and phenotypic differentiation between highland and lowland forms clearly apparent, this study is more inclined to support these montane forms as distinct taxa.

A high degree of structure is noted within the *Z. p. mbuluensis* clade, which is unexpected given that all samples were taken from the same forest fragment. This structure may reflect high levels of diversity within the montane populations of South Pare (Tanzania), or more likely could be the result of missing data. Further sampling of montane forms within this region would allow for a more detailed assessment of genetic diversity within this clade.

4.6.5. Southern African forms

The taxonomic affinities of southern African *Zosterops* have long been a source of disagreement (Gill 1936; Moreau 1957; Clancey 1967; Hockey et al. 2005; van Balen 2008; Oatley 2011; 2012) and have led to numerous taxonomic revisions (Gill 1936; Moreau 1957; Clancey 1967). The phylogenetic placement of southern African forms is largely congruent with the recent southern African-centred phylogeny of Oatley *et al.* (2012) that placed *Z. pallidus* as sister to the other southern African taxa, with samples of *Z. senegalensis* recovered as sister to a clade comprising *Z. capensis* and *Z. virens*.

Additional sampling of the subspecies *Z. p. sundevalli*, places it as conspecific with *Z. p. pallidus* with no division of the two subspecies. This placement is concordant with the previous taxonomic arrangement of Moreau

(1957) who split the two ‘cinnamon’ flanked forms from all other southern African taxa based on vocalisation differences and plumage variation. GMYC analysis recovers this clade as independent from all other southern taxa (L9) providing strong support for this group as an independent taxonomic unit. In contrast despite falling into three well-marked phenotypic groupings (Oatley *et al.* 2011) relationships between all other southern Africa taxa are poorly resolved.

4.6.6. Northern and central Sub-Saharan forms

Molecular investigation into the relationships and taxonomic validity of broadly distributed yellow-bellied forms, revealed *Z. senegalensis* to be a cryptic species complex (Funk and Omland 2003). Despite being remarkably homogenous in appearance, members of *Z. senegalensis* are recovered in multiple independent lineages rendering *Z. senegalensis* polyphyletic. Northern and central members of *Z. senegalensis* are recovered into three distinct clades within the major clade E. The division of these clades is largely concordant with geography corresponding to central African forms, northern Sub-Saharan forms, and an isolate montane form from Mount Cameroon in West Africa.

Central African members of *Z. senegalensis (Z. s. stuhlmanni*, *Z. s. reichenowi* and *Z. s. toroensis*) are recovered as a distinct evolutionary lineage (L12) that contains a high degree of genetic structuring. While *Z. s. toroensis* and reichenowi are recovered as independent coalescent clusters, samples of *Z. s. stuhlmanni* are recovered as independent singletons that are non-monophyletic in their placement. Results are interpreted to suggest that *Z. s. stuhlmanni* may require further division, although further sampling throughout this central African region is required in order to obtain a better understanding of the geographic division of forms.

The placement of highland *Z. s. stenocricotus* (Mount Cameroon in West Africa) as sister to *S. brunneus* (Bioko, Gulf of Guinea) is concordant with the previous Gulf of Guinea-centred phylogeny of Melo *et al.* (2011). GMYC estimates provide strong support for these sister taxa as an independent taxonomic unit relative to all other African forms (L13). However, while phylogenetically distinct, support for *Z. s. stenocricotus* as an independent lineage relatively to *S. brunneus* is lacking. Nevertheless, results do not place

them in the same coalescent cluster suggesting that these lineages are representative of some genetic variability.

With the exception of a single sample of *Z. s. gerhardi* (BMNH 1939 10 13 123), all northern Sub-Saharan members of *Z. senegalensis* (*Z. s. senegalensis*, *Z. s. demeryi*, *Z. s. gerhardi*, *Z. s. jacksoni*) are recovered in a single clade. In contrast to the other *Z. s. gerhardi* samples that are recovered in a northern Sub-Saharan Z. senegalensis clade, sample BMNH 1939 10 13 123 is recovered within a clade of *Z. poliogaster*. Locality data associated with this specimen place it within the Imatong Mountains of South Sudan. At present the distribution of *Zosterops* forms within this region is limited to *Z. s. gerhardi*. However, given the topographical complexity of this region, the placement of sample BMNH 1939 10 13 123 within a clade containing northerly members of *Z. poliogaster* may indicate the presence of *Z. poliogaster* within this region. Other *Z. s. gerhardi* samples are recovered as sister to *Z. s. jacksoni* samples from northern Kenya. Lowland *Z. s. senegalensis* and *Z. s. demeryi* are recovered as a single clade. While samples of *Z. s. senegalensis* are recovered as independent singletons, there is strong support for *Z. p. demeryi* as a single coalescent cluster. This result might not be surprising given that *Z. senegalensis* (Senegal to northwest Ethiopia) occupies a much larger area relative to demeryi (Sierra Leone, Liberia, Ivory Coast).

According to the dated molecular phylogeny of Warren *et al.* (2006) colonisation of Africa occurred after an early wave expansion from Asia to the Indian Ocean. While the Ancient Indian Ocean white-eyes are consistently recovered at the base of the mainland African clade (Warren *et.al.* 2006; Melo *et.al.* 2011), poor support for relationships within the Ancient Indian Ocean clade has made resolving the exact origin of African white-eyes difficult (Warren *et.al.* 2006). The biogeographic disjunct between members of the Ancient Indian Ocean clade, coupled with the lack of branch support for relationships and high mtDNA divergences, suggest that related forms (which have now become extinct) may have once existed (Warren *et.al.* 2006). However, in the absence of molecular and geographic data for these putative extinct forms, very little is known regarding the spatial scales of this early expansion into the Indian Ocean region.

Extended sampling of lowland forms in northeast Africa has recovered two samples of mainland *Z. a. socotranus* as distinct from all other African taxa (Fig 4.4; Clade B). Results identify this group as an independently evolving evolutionary lineage and place samples outside the major African radiation between the Ancient Indian Ocean White-eyes and all other African Zosteropidae. These samples may represent a relic population of an early wave expansion into mainland Africa. However very little is known about the diversity and distribution of this mainland form and consequently examining the origin and evolutionary history of this group is not possible at present. These results highlight this region as focus for future sampling efforts, which may lead to a better understanding of the origin of African Zosteropidae.

4.7.Conclusion

The use of museum collections has enabled the dense sampling of all four mainland African *Zosterops* species, including all 34 currently recognised subspecies. This increased sampling has allowed for the resolution of relationships and patterns of diversity across mainland Africa. This work provides the first comprehensive framework of genetic relationships, which has been used to generate a series of predictions regarding the probable placement of species boundaries. Overall results provide strong support for six major clades within the African Zosteropidae system (Continental Africa, Gulf of Guinea, Indian Ocean, Gulf of Aden and the Arabian Peninsula). Despite poor resolution of branching patterns between these assemblages, genetic results provide strong support for relationships within these clades. GYMC results subdivide African Zosteropidae into 14 distinct evolutionary lineages, although whether these lineages represent species or taxa at different hierarchical levels still remains to be examined. The non-monophyly of mainland African taxa demonstrates that traditional morphological characters used to delineate species within Zosteropidae are not informative in an evolutionary context, with results indicating that the current taxonomic framework greatly underestimates *Zosterops* diversity within mainland Africa. While unparalleled sampling of African Zosteropidae using DNA extracted from archive and fresh material has allowed for the largest genetic assessment of mainland African Zosteropidae to date, extensive work is still required to resolve the systematics of this group.

4.8.References

Agapow P-M (2004) The impact of species concept on biodiversity studies. *The Quarterly Review of Biology*, **79**, 161-179.

Alfaro ME, Karns DR, Voris HK, Brock CD, Stuart BL (2008). Phylogeny, evolutionary history, and biogeography of Oriental-Australian rear-fanged water snakes (Colubroidea: Homalopsidae) inferred from mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics and Evolution*, **46**, 576–593.

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

Balakrishnan R (2005) Species concepts, species boundaries and species identification: A view from the tropics. *Systematic Biology*, **4**, 689-693.

Bantock TM, Prys-Jones RP, Lee PL (2008) New and improved molecular sexing methods for museum bird specimens. *Molecular Ecology Resources*, **8**, 519-528.

Barraclough TG, Hughes M, Ashford-Hodges N, Fujisawa T (2009) Inferring evolutionarily significant unit of bacterial diversity from broad environmental surveys of single-locus data. *Biology Letters*, **5**, 425-428.

Bello N, Francino O, Sánchez A (2001) Isolation of genomic DNA from feathers. *Journal of Veterinary Diagnostic Investigation*, **13**, 162-164.

Beltrán M, Jiggins CD, Brower AVZ, Bermingham E, Mallet J (2007). Do pollen feeding and pupal-mating have a single origin in *Heliconius*? Inferences from multilocus sequence data. *Biological Journal of the Linnaean Society*, **92**, 221–239.

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

BirdLife International. (2000) *Threatened birds of the world.* Barcelona and Cambridge, U.K.: Lynx Edicions and BirdLife International.

Borghesio L, Laiolo P (2004) Habitat use and feeding ecology of Kulal white-eye *Zosterops (poliogaster) kulalensis*. *Bird Conservation International*, **14**, 11-24.

Brito PH, Edwards SV (2008) Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica*, **135**, 439-455.

Carstens BC, Dewey TA (2010) Species delimitation using coalescent and information-theoretic approach: An example from North American *Myotis* Bats. *Systematic Biology*, **59**, 400-414.

Ceccareli FS, Sharkey MJ, Zaldívar-Riverón (2012) Species identification in the taxonomically neglected, highly diverse, neotropical parasitoid wasp genus *Notiospathius* (Braconidae: Doryctinae) based on an integrative molecular and morphological approach. *Molecular phylogenetics and Evolution*, **62**, 485-495.

Clancey PA (1967) Taxonomy of the southern African *Zosterops*. *Ibis*, **109**, 318-327.

Clegg SM, Degnan SM, Moritz C, Kikkawa J, Estoup A, Owens IP (2002) Microevolution in island forms: the roles of drift and directional selection in morphological divergence of a passerine bird. *Evolution*, **56**, 2090–2099.

Clegg SM, Frentiu FD, Kikkawa J, Tavecchia G, Owens IPF (2008) 4000 years of phenotypic change in an island bird: heterogeneity of selection over three micro-evolutionary timescales. *Evolution*, **69**, 2393-2410.

Collar NJ, Crosby M, Stattersfield AJ (1994) *Birds to watch 2: the world list of the threatened birds.* Cambridge, U.K.: BirdLife International. (BirdLife Conservation Series 4.)

David N, Gosselin M (2002) Gender agreement of avian species names. *Bulletin of the British Ornithologists' Club*, **122**, 14-49.

de Moraes-Barros N and Morgante JS (2007) A simple protocol for the extraction and sequence analysis of DNA from study skin of museum collection. *Genetics and Molecular Biology*, **30**, 1181-1185.

de Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology*, **56**, 879-886.

d'Horta FM, Cabanne GS, Meyer D, Miyaki CY (2011) The genetic effects of late quaternary climate changes over a tropical latitudinal gradient: diversification of an Atlantic Forest passerine. *Molecular Ecology*, **20**, 1923-1935.

Dickinson EC (ed.) (2003) The Howard & Moore Complete Checklist of the Birds of the World, 3rd Edition. Christopher Helm, London.

Dieffenbach CW, Lowe TM, Dveksler GS (1993) General concepts for PCR primer design. *Genome Research*, **3**, S30-S37.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Fry CH, Keith S, Urban EK (Eds) (2000) The birds of Africa, Vol VI. London: Academic Press.

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

Garcia-Moreno J, Sorenson MD, Mindell DP (2003) Congruent avian phylogenies inferred from mitochondrial and nuclear DNA sequences. *Journal of molecular ecology*, **57**, 27-37.

Gill EL (1936) A first guide to South African birds. Cape Town: Maskew Miller Limited.

Graves GR, Braun MJ (1992) Museums: store-houses of DNA? *Science*, **255**, 1335-1336.

Hall BP, Moreau RE (1970) *An atlas of speciation in African passerine birds*. British Museum (Natural History). London

Hamilton CA, Formanowicz DR, Bond JE (2011) Species delimitation and phylogeography of *Aphonopelma hentzi* (Araneae, Mygalomorphae, Theraphosidae): Cryptic diversity in North American Tarantulas. *PLoS ONE*, **6**, e26207.

Harrington RC, Near TJ (2012) Phylogenetic and coalescent strategies of species delimitation in Snubnose Darters (Percidae: *Etheostoma*). *Systematic Biology*, **61**, 63-79.

Hennache A, Rasmussen P, Lucchini V, Rimondi S, Randi E (2003) Hybrid origin of the imperial pheasant *Lophura imperialis* (Delacour and Jabouille, 1924) demonstrated by morphology, hybrid experiments, and DNA analyses. *Biological Journal of the Linnaean Society*, **80**, 573-600.

Herbert P D N, Cywinska A, Ball S H, deWaard J R (2003) Biological identification through DNA barcodes. *Proceeding of the Royal Society of London B*, **270**, 313-321.

Hockey PAR, Dean WRJ, Ryan PG. (Eds.) (2005) Roberts – Birds of Southern Africa, V11th ed. The Trustees of the John Voelker Bird Book Fund, Cape Town.

Houde P, Braun MJ (1988) Museum collections as a source of DNA for studies of Avian phylogeny. *The Auk*, **105**, 773-776.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics Applications Notes*, **17**, 754-755.

Hugall AF, Foster R, Hutchinson M, Lee MSY (2008) Phylogeny of Australasian agamid lizards based on nuclear and mitochondrial genes: implications for morphological evolution and biogeography. *Biological Journal of Linnean Society,* **93**, 343–358.

Irestedt M, Ohlson JI, Zuccon D, Kälersjö, Ericson PGP (2006) Nuclear DNA from old collections of Avian study skins reveals the evolutionary history of the Old World suboscines (Aves, Passeriformes). *Zoological Scripta*, **35**, 567-580.

Irwin DE, Alström P, Olsson U, Benowitz-Fredericks ZM (2001) Cryptic species in the genus *Phylloscopus* (Old World leaf warblers) *Ibis*, **143**, 233-247.

Jousselin E, Desdevises Y, Coeur d'acier A (2009) Fine-scale cospeciation between Brachycaudus and Buchnera aphidicola: bacterial genome helps define species and evolutionary relationships in aphids. *Proceedings of the Royal Society B*, **276**, 187–196.

Knowles LL, Carstens BC (2007) Delimiting species without monophyletic gene trees. *Systematic Biology*, **56**, 887-895.

Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695-1701.

Leaché AD, Fujita MK (2010) Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus). Proceeding of the Royal Society B*, **277**, 3071-3077.

Lee PL, Prys-Jones RP (2008) Extracting DNA from museum bird eggs, and whole genome amplification of archive DNA. *Molecular Ecology Resources*, **8**, 551-560.

Mandrioli M, Borsatti F, Mola L (2006) Factors affecting DNA preservation from museum-collected lepidopteran specimens. *Entomologia Experimentalis et Applicata*, **120**, 239-244.

Mann CF (2007) Use of skin from the feet of museum specimens as a source of DNA. *The Auk*, **124**, 339-339

Mayr (1965) Relationships among Indo-Australian Zosteropidae. *Breviora*, **228**, 1-6.

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, Warren BH, Jones PJ (2011) Rapid parallel evolution of aberrant traits in the diversification of the Gulf of Guinea white-eye (Aves, Zosteropidae). *Molecular Ecology*, **20**, 4953-4967.

Milá B, Warren BH, Heeb P, Thébaud C (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: *Zosterops borbonicus*). *BMC Evolutionary Biology*, **10**, 158-171.

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJG, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP (2009)

Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology*, **58**, 298-311.

Moreau RE (1957) Variation in the western Zosteropidae (Aves). *Bulletin of the British Museum,* **4**, 318–433.

Moritz C, Dowling T E, Brown W M (1987) Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. *Annual Review of Ecology and Systematics,* **18** 269-292.

Moum T, Arnason U, Árnason (2002) Mitochondrial DNA sequence evolution and phylogeny of the Atlantic Alcidae, including the extinct Great Auk (*Pinguinus impennis*). *Molecular Biology and evolution*, **19**, 1434-1439.

Moyle RG, Filardi CE, Smith CE, Diamond JC (2009) Explosive pleistocene diversification and hemi- spheric expansion of a ‘‘great speciator’’. *Proceedings of the National Academy of Science USA*, **106**, 1863–1868.

Mundy NI, Unitt P, Woodruff DS (1997) Skin from feet of museum specimens as a non-destructive source of DNA for Avian genotyping. *The Auk*, **114**, 126-129.

Nagai M, Yoshida A, Sato N (1998) Additive effect of bovine serum albumin, dithiothreitol, and glycerol on PCR. *Biochemistry and Molecular Biology International*, **44**, 157-163.

Nguembock B, Fjeldså J, Couloux A, Pasquet E (2008) Molecular phylogeny of Carduelinae (Aves, Passeriformes, Fringillidae) proves polyphyletic origin of the genera *Serinus* and *Carduelis* and suggests redefined generic limits. *Molecular Phylogenetics and Evolution*, **51**, 169-181.

Norman JA, Christidis L, Westerman M, Hill RFA (1998) Molecular data confirms the species status of the Christmas island Hawk-Owl *Ninox natalis. EMU Austral Ornithology*, **98**, 197-208.

Oatley G, Bowie RCK, Crowe TM (2011) The use of subspecies in the systematics of southern African white-eyes: historical entities or eco-geographic variants. *Journal of Zoology*, **284**, 21-30.

Oatley G, Voelker G, Crowe TM, Bowie RCK (2012) A multi-locus phylogeny reveals a complex pattern of diversification related to climate and habitat heterogeneity in southern African white-eyes. *Molecular Phylogenetics and Evolution*, **64**, 633-644.

Payne RB, Sorenson MD (2003) Museum collections as source of genetic data. *Bonner Zoologische Beiträge*, **51**, 97- 104.

Phillimore AB, Owens IPF, Black RA, Chittock J, Burke T, Clegg SM (2008) Complex patterns of genetic and phenotypic divergence in an island bird and the consequences for delimiting conservation units. *Molecular Ecology*, **17**, 2839-2853.

Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, **55**, 595-609.

Powell JR (2012) Accounting for uncertainty in species delineation during the analysis of environmental DNA sequence data. *Methods in Ecology and Evolution*, **3**, 1-11.

Prager M, Johansson AEI, Andersson S (2008) Molecular phylogeny of the African widowbirds and bishops, *Euplectes* spp. (Aves: Passeridae: Ploceinae) *Molecular Phylogenetics and Evolution*, **46**, 290-302.

Pratt DH, Bruner PL, Berrett DG (1987) A field guide to the birds of Hawaii and the tropical Pacific. Princeton, University Press, Princeton, USA.

Rach J, DeSalle R, Sarkar I, Schierwater B, Hadrys H (2008) Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata. *Proceedings of the Royal Society B*, **275**, 237–247.

Rohland N, Hofreiter M (2007) Comparison and optimization of ancient DNA extraction *BioTechniques*, **42**, 343-352.

Sclater WL (1930) *Systema avium Æthiopicarum*. British Ornithologists' Union, London. UK.

Sefc KM, Payne RB, Sorenson MD (2003) Microsatellites amplification from museum feather samples: effect of fragment size and template concentration on genotyping errors. *The Auk*, **120**, 982-989.

Singh VK, Govindarajan R, Naik S, Kumar A (2000) The effect of hairpin structure on PCR amplification efficiency. *Molecular Biology Today*, **1**, 67-69.

Slikas B, Jones IB, Derrickson SR, Fleischer RC (2000) Phylogenetic relationships of Micronesian White-eyes based on mitochondrial sequence data. *The Auk*, **117**, 355-365.

Slikas B, Olson SL, Fleischer RC (2002) Rapid, independent evolution of flightlessness in four species of Pacific island raills (Rallidae): an analysis based on mitochondrial sequence data. *Journal of Avian Biology*, **33**, 5-14.

Swofford DL (2003) PAUP*: phylogenetic analysis using parsimony (*and other methods). version 4 beta 10. Sinauer Associates, Sunderland, Massachusetts.

Töpfer T, Gamauf A, Haring E (2011) Utility of arsenic-treated bird skins for DNA extraction. *BMC Research Notes*, **4**, 1-8.

van Balen S (2008) Family Zosteropidae (White-eyes). In Handbook of the birds of the world, Vol. 13: 402–485. Del Hoyo, J., Elliott, A. & Christie, D.A. (Eds). Barcelona: Penduline Tits to Shrikes. Lynx Edicions.

van den Elzen R, König C (1983) Vögel des (Süd-) Sudan: taxonomische und Tiergeographische Bemerkungen. *Bonner Zoologische Beiträge*, **34**, 149-196.

Wandeler P, Hoeck PEA, Keller LF (2007) Back to the future: museum specimens in population genetics. *Trends in Ecology and Evolution*, **22**, 634-642.

Warren B, Bermingham E, Prys-Jones RP, Thebauds C (2006) Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on the Indian Ocean islands. *Molecular Ecology,* **15**, 3769-3786.

Wiens JJ, Penkrot TA (2002) Delimiting species using DNA and morphological variation and discordant species limits in Spiny Lizards (*Sceloporus*). *Systematic Biology*, **51**, 69-91.

Wiens JJ (2007) Species delimitation: New approaches for discovering diversity. *Systematic Biology*, **56**, 875-878.

Winker K (2005) Bird collections: Development and use of a scientific resource. *The Auk,* **122**, 966-971.

Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences USA*, **107**, 9264–9269.

Yeung CKL, Lin R-C, Lei F, Robson C, Hung, LM, Liang W, Zhou F, Han L, II, S-H, Yang X (2011) Beyond a morphological paradox: Complicated phylogenetic relationships of the parrotbills (Paradoxornithidae, Aves). *Molecular Phylogenetics and Evolution*, **61**, 192-202.

Zimmermann J, Hajibabaei M, Blackburn DC, Hanken J, Cantin E, Posfai J, Evans TC (2008) DNA damage in preserved specimens and tissue samples: a molecular assessment. *Frontiers in Zoology*, **5**, 18.

4.9 Appendix III

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3 1	ND3 2	Cyt-b 1	Cyt-b 2	Cyt-b 3	Cyt-b 4	Cyt-b 5	Cyt-b 6
Stachyris whiteheadi	Phillippines	KUNHM18001	Sequence	NCBI		JN826960.1		NJ827210.1					
Z. erythropleurus	Captive	LSUMNS B20625	Sequence	NCBI	FJ460937.1	FJ460868.1							
Z. erythropleurus	China	GLGS1443	Sequence	NCBI		DQ837492.1		DQ837524.1					
Z. japonicus	Japan	AMNH DOT10981	Sequence	NCBI	FJ460938.1	FJ460869.1							
Z. palpebrosus													
palpebrosus	Nepal	RF2	Sequence	NCBI		DQ328448.1		DQ328348.1					
palpebrosus	Nepal	AMNH DOT5746	Sequence	NCBI	FJ460810.1	FJ460947.1							
egregious	India	BMNH 1964.26.1	Sequence	NCBI		DQ328450.1		DQ328350.1					
Z. atricapilla	Malaysia	LSUMNS B36444	Sequence	NCBI	FJ460939.1	FJ460870.1		JN827239.1					
Z. nigrorum	Phillippines	FMNH 432997	Sequence	NCBI	FJ460945.1	FJ460876.1		JN827242.1					
Z. montanus	Sulwesi, Indonesia	AMNH DOT12552	Sequence	NCBI	FJ460947.1	FJ460878.1							
Z. montanus	Phillippines	O2655	Sequence	NCBI		DQ328461.1		DQ328361.1					
Z. lateralis	Vanuatu, South Pacific	LSUMNS B45835	Sequence	NCBI	FJ460941.1	FJ460872.1							
Z. lateralis	Australia	AMNH DOT6094	Sequence	NCBI		FJ460882.1					JN827242.1		
Z. abyssinicus													
abyssinicus	Asmara, Eritrea	BMNH 1952.32.3	Toe pad	BMHN									
abyssinicus	Lake Tana,Ethiopia	BMNH 1927.11.5.577	Toe pad	BMHN									
abyssinicus	Suakin, Sudan	BMNH 1915.12.24.1198	Toe pad	BMHN									
socotranus	Al Medo, Somalia	BMNH 1956.37.9	Toe pad	BMHN									
socotranus	Daloh, Somalia	BMNH 1982.3.44	Toe pad	BMHN									
socotranus	Socotra, Gulf of Aden	BMNH 1899.8.11.23	Toe pad	BMHN									
socotranus	Socotra, Gulf of Aden	BW292	Blood	WARREN									
socotranus	Socotra, Gulf of Aden	BW293	Blood	WARREN									
arabs	Taif, Saudi Arabia	BMNH 1935.5.10.48	Toe pad	BMHN									

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3		Cyt-b					
						1	2	1	2	3	4	5	6
arabs	Menacha, Yemen	BMNH 1913.7.18.44	Toe pad	BMHN									
omoensis	Lake Turkana, Ethiopia	BMNH 1912.10.15.1275	Toe pad	BMHN									
omoensis	Gofa, Ethiopia	BMNH 1912.10.15.1290	Toe pad	BMHN									
omoensis	Lake Tana, Ethiopia	BMNH 1927.11.5.580	Toe pad	BMHN									
jubaensis	Arero, Ethiopia	BMNH 1946.5.2722	Toe pad	BMHN									
jubaensis	South Horr, Kenya	T60	Blood	COX									
jubaensis	South Horr, Kenya	T61	Blood	COX									
jubaensis	South Horr, Kenya	T65	Blood	COX									
jubaensis	South Horr, Kenya	T69	Blood	COX									
jubaensis	South Horr, Kenya	T70	Blood	COX									
jubaensis	South Horr, Kenya	T73	Blood	COX									
jubaensis	South Horr, Kenya	T76	Blood	COX									
jubaensis	South Horr, Kenya	T77	Blood	COX									
jubaensis	South Horr, Kenya	T84	Blood	COX									
jubaensis	South Horr, Kenya	T85	Blood	COX									
jubaensis	S. Ethiopia	NRM 570800	Toe pad	NRM									
jubaensis	S. Ethiopia	NRM 570801	Toe pad	NRM									
flavilateralis	Umani spring, Kenya	T4	Blood	COX									
flavilateralis	Umani spring, Kenya	T5	Blood	COX									
flavilateralis	Umani spring, Kenya	T11	Blood	COX									
flavilateralis	Umani spring, Kenya	T14	Blood	COX									
flavilateralis	Umani spring, Kenya	T15	Blood	COX									
flavilateralis	Umani spring, Kenya	T17	Blood	COX									
flavilateralis	Umani spring, Kenya	T20	Blood	COX									
flavilateralis	Umani spring, Kenya	T21	Blood	COX									
flavilateralis	Umani spring, Kenya	T23	Blood	COX									

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3		Cyt-b					
						1	2	1	2	3	4	5	6
flavilateralis	Umani spring, Kenya	T30	Blood	COX									
flavilateralis	Nairobi, Kenya	RB3	Sequence	NCBI		DQ328439.1		DQ328339.1					
Z. pallidus													
pallidus	South Africa	AP50340	Blood	PFIAO									
pallidus	Great Namaqualand, Namibia	BMNH 1950.50.666	Toe pad	BMHN									
pallidus	Viools Drift, South Africa	BMNH 1950.50.660	Toe pad	BMHN									
sundevalli	Venterskroon, South Africa	BMNH 1928.2.5.12	Toe pad	BMHN									
sundevalli	Fourteen Streams, South Africa	BMNH 1904.11.19.56	Toe pad	BMHN									
sundevalli	Bloemfontein, South Africa	BMNH 1923.8.7.2982	Toe pad	BMHN									
caniviridis	Kanye, Botswana	BMNH 1957.36.207	Toe pad	BMHN									
capensis	Cape Town, South Africa	RB1	Sequence	NCBI									
capensis	Cape Town, South Africa	RB4	Sequence	NCBI									
capensis	Knysna, South Africa	BMNH 1905.12.29.1741	Toe pad	BMHN									
capensis	Table Mountain, South Africa	BMNH 1905.12.29.1737	Toe pad	BMHN									
capensis	Cape Town, South Africa	BMNH 1952.60.24	Toe pad	BMHN									
atmorii	Grahamstown, South Africa	BMNH 1969.48.279	Toe pad	BMHN									
atmorii	Litenhage, South Africa	BMNH 1930.12.20.35	Toe pad	BMHN									
atmorii	Ngoye Hills, South Africa	BMNH 1905.12.29.1722	Toe pad	BMHN									
atmorii	Legogot, South Africa	BMNH 1905.12.29.1732	Toe pad	BMHN									
virens	South Africa	AM36426	Blood	PFIAO									
virens	South Africa	AM36429	Blood	PFIAO									
virens	South Africa	AM36433	Blood	PFIAO									
virens	Zuurbron, South Africa	BMNH 1905.12.29.1725	Toe pad	BMHN									
Z. senegalensis													
senegalensis	Tunga, Nigera	BMNH 1938.8.3.10	Toe pad	BMHN									
senegalensis	Kagelu, Sudan	BMNH 1947.100.307	Toe pad	BMHN									

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3 1 2	Cyt-b 1 2 3 4 5 6
senegalensis	Mt Nimba, Libera	BMNH 1977.20.2495	Toe pad	BMHN			
senegalensis	N. Ghana	B39250	Sequence	NCBI		DQ328443.1	DQ328343.1
senegalensis	N. Ghana	B39514	Sequence	NCBI		DQ328442.1	DQ328342.1
demeryi	Enugu, Nigera	BMNH 1966.16.3411	Toe pad	BMHN			
demeryi	Kintampo, Ghana	BMNH 1911.12.23.2612	Toe pad	BMHN			
demeryi	Mt Nimba, Libera	BMNH 1977.20.2492	Toe pad	BMHN			
stenocricotus	Saxenhof, Cameroon	BMNH 1966.16.3408	Toe pad	BMHN			
stenocricotus	Mt Cameroon, Cameroon	BMNH 1966.16.3386	Toe pad	BMHN			
stenocricotus	Lake Oku, Cameroon	BMNH 1951.34.814	Toe pad	BMHN			
stenocricotus	Mt Cameroon, Cameroon	STC01	Blood	CIBIO			
gerhardi	Kipia, Imatongs Mts, Sudan	BMNH 1947.100.308	Toe pad	BMHN			
gerhardi	Didinga Hills, Sudan	BMNH 1939.10.1.284	Toe pad	BMHN			
gerhardi	Lomoling, Imatong Mts, Sudan	BMNH 1939.10.3.123	Toe pad	BMHN			
stuhlmanni	Mabira forest, Uganda	BMNH 1934.1.17.27	Toe pad	BMHN			
stuhlmanni	Mpumuru, Uganda	BMNH 1913.7.16.140	Toe pad	BMHN			
stuhlmanni	Bukoba, Tanzania	NRM 552125	Toe pad	NRM			
reichenowi	Mubuku Valley,Uganda	BMNH 1906.12.23.718	Toe pad	BMHN			
reichenowi	Sake, DRCongo	NRM 570802	Toe pad	NRM			
toroensis	Campi ya Wambutti, DRCongo	NRM 570803	Toe pad	NRM			
toroensis	Sembiliki, DRCongo	NRM 570804	Toe pad	NRM			
toroensis	N.E. Congo	BMNH 1936.2.21.237	Toe pad	BMHN			
jacksoni	Mathews Range, Kenya	BLS06	Blood	BORGHESIO			
jacksoni	Mt Nyiru, Kenya	BLS35	Blood	BORGHESIO			
jacksoni	Mt Nyiru, Kenya	BLS65	Blood	BORGHESIO			
jacksoni	Mathews Range, Kenya	BLS77	Blood	BORGHESIO			
jacksoni	Mathews Range, Kenya	BLS81	Blood	BORGHESIO			

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3		Cyt-b					
						1	2	1	2	3	4	5	6
jacksoni	Mt Nyiru, Kenya	ZMUC 131317	Blood	ZMUC									
jacksoni	Mt Nyiru, Kenya	ZMUC 131324	Blood	ZMUC									
jacksoni	Mt Nyiru, Kenya	ZMUC 131325	Blood	ZMUC									
jacksoni	Mt Nyiru, Kenya	ZMUC 131331	Blood	ZMUC									
jacksoni	Aberdares Range, Kenya	ZMUC 146780	Blood	ZMUC									
jacksoni	Aberdares Range, Kenya	ZMUC 146784	Blood	ZMUC									
jacksoni	Aberdares Range, Kenya	ZMUC 146785	Blood	ZMUC									
jacksoni	Aberdares Range, Kenya	ZMUC 146786	Blood	ZMUC									
jacksoni	Kakamega, Kenya	T41	Blood	COX									
jacksoni	Kakamega, Kenya	T42	Blood	COX									
jacksoni	Kakamega, Kenya	T49	Blood	COX									
jacksoni	Kakamega, Kenya	T50	Blood	COX									
jacksoni	Kakamega, Kenya	T51	Blood	COX									
jacksoni	Kakamega, Kenya	T52	Blood	COX									
jacksoni	Kakamega, Kenya	T53	Blood	COX									
jacksoni	Kakamega, Kenya	T54	Blood	COX									
kasaicus	N'dalatando, Angola	BMNH 1910.5.6.1050	Toe pad	BMHN									
kasaicus	N'dalatando, Angola	BMNH 1910.5.6.1050	Toe pad	BMHN									
heinrichi	N'dalatando, Angola	BMNH 1910.5.6.1047	Toe pad	BMHN									
heinrichi	N'dalatando, Angola	BMNH 1910.5.6.1052	Toe pad	BMHN									
heinrichi	N'dalatando, Angola	BMNH 1910.5.6.1051	Toe pad	BMHN									
heinrichi	Luena, Angola	BMNH 1957.35.527	Toe pad	BMHN									
quanzae	Mt Moco, Angola	BMNH 1957.35.531	Toe pad	BMHN									
anderssoni	Livingstone, Zambia	BMNH 1945.18.60	Toe pad	BMHN									
anderssoni	Mangochi, Malawi	BMNH 1937.12.19.141	Toe pad	BMHN									
anderssoni	Kalulabula, Angola	BMNH 1932.5.5.125	Toe pad	BMHN									

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3		Cyt-b					
						1	2	1	2	3	4	5	6
anderssoni	Kalulabula, Angola	BMNH 1932.5.5.128	Toe pad	BMHN									
tongensis	Zululand, Mozambique	BMNH 1956 25 115	Toe pad	BMHN									
tongensis	Coguno, Mozambique	BMNH 1905.12.29.1713	Toe pad	BMHN									
stierlingi	Tabora, Tanzania	ZMUC 145467	Blood	ZMUC									
stierlingi	Udzungwa Mts, Tanzania	ZMUC 140192	Blood	ZMUC									
stierlingi	Poroto Mts,Tanzania	ZMUC 142605	Blood	ZMUC									
stierlingi	Poroto Mts,Tanzania	ZMUC 142607	Blood	ZMUC									
stierlingi	Usambara Mts, Tanzania	ZMUC 129298	Blood	ZMUC									
stierlingi	Usambara Mts, Tanzania	ZMUC 129289	Blood	ZMUC									
unknown	DRCongo	ZMUC 128660	Blood	ZMUC									
unknown	DRCongo	ZMUC 128632	Blood	ZMUC									
unknown	DRCongo	ZMUC 128658	Blood	ZMUC									
Z. poliogaster													
poliogaster	Faghena, Eritrea	BMNH 1954.26.20	Toe pad	BMHN									
poliogaster	Faghena, Eritrea	BMNH 1954.26.21	Toe pad	BMHN									
poliogaster	Borena, Ethiopia	BMNH 1946.5.2433	Toe pad	BMHN									
poliogaster	Addis-Abeba, Eritea	NRM 51733	Toe pad	NRM									
kaffensis	Maji, Ethiopia	BMNH 1945.40.166	Toe pad	BMHN									
kaffensis	Jimma, Ethiopia	BMNH 1912.10.15.1932	Toe pad	BMHN									
kaffensis	Charada forest, Ethiopia	BMNH 1912.10.15.1262	Toe pad	BMHN									
kulalensis	Mt Kulal, Kenya	K30	Blood	COX									
kulalensis	Mt Kulal, Kenya	K31	Blood	COX									
kulalensis	Mt Kulal, Kenya	K33	Blood	COX									
kulalensis	Mt Kulal, Kenya	K34	Blood	COX									
kulalensis	Mt Kulal, Kenya	K35	Blood	COX									
kulalensis	Mt Kulal, Kenya	K37	Blood	COX									

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3		Cyt-b					
						1	2	1	2	3	4	5	6
kulalensis	Mt Kulal, Kenya	K38	Blood	COX									
kulalensis	Mt Kulal, Kenya	K39	Blood	COX									
kulalensis	Mt Kulal, Kenya	K41	Blood	COX									
kulalensis	Mt Kulal, Kenya	K42	Blood	COX									
kulalensis	Mt Kulal, Kenya	2MK3	Blood	HABEL									
kulalensis	Mt Kulal, Kenya	2MK5	Blood	HABEL									
kulalensis	Mt Kulal, Kenya	2MK6	Blood	HABEL									
kulalensis	Mt Kulal, Kenya	2MK7	Blood	HABEL									
kulalensis	Mt Kulal, Kenya	2MK8	Blood	HABEL									
kulalensis	Mt Kulal, Kenya	2MK9	Blood	HABEL									
kulalensis	Mt Kulal, Kenya	2MK10	Blood	HABEL									
kulalensis	Mt Kulal, Kenya	2MK11	Blood	HABEL									
kikuyuensis	Mt Kenya, Kenya	MK1	Blood	HABEL									
kikuyuensis	Mt Kenya, Kenya	MK2	Blood	HABEL									
kikuyuensis	Mt Kenya, Kenya	MK3	Blood	HABEL									
kikuyuensis	Mt Kenya, Kenya	MK6	Blood	HABEL									
kikuyuensis	Mt Kenya, Kenya	MK7	Blood	HABEL									
kikuyuensis	Aberdares Range, Kenya	AB2	Blood	HABEL									
kikuyuensis	Aberdares Range, Kenya	AB4	Blood	HABEL									
kikuyuensis	Aberdares Range, Kenya	AB10	Blood	HABEL									
kikuyuensis	Aberdares Range, Kenya	AB11	Blood	HABEL									
kikuyuensis	Aberdares Range, Kenya	AB12	Blood	HABEL									
kikuyuensis	Aberdares Range, Kenya	AB13	Blood	HABEL									
kikuyuensis	Aberdares Range, Kenya	AB20	Blood	HABEL									
kikuyuensis	Mt Kenya, Kenya	RB2	Sequence	NCBI		DQ328447.1		DQ328247.1					
silvanus	Mbololo, Taita, Kenya	O8580	Sequence	NCBI		DQ328437.1		DQ328337.1					

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3		Cyt-b					
						1	2	1	2	3	4	5	6
silvanus	Taita Hills, Kenya	TH 1 2	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 1 3	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 1 10	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 1 12	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 1 21	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 2 3	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 2 10	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 2 12	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 2 15	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 3 11	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 3 12	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 3 20	Blood	HABEL									
silvanus	Taita Hills, Kenya	TH 3 21	Blood	HABEL									
eurycricotus	Mt Meru, Tanzania	BMNH 1937.1.28.37	Toe pad	BMHN									
eurycricotus	Mt Meru, Tanzania	BMNH 1939.10.11.94	Toe pad	BMHN									
eurycricotus	Mt Meru, Tanzania	NRM 570798	Toe pad	NRM									
eurycricotus	Mt Meru, Tanzania	NRM 570799	Toe pad	NRM									
mbuluensis	Chyulu Hills, Kenya	CH1	Blood	HABEL									
mbuluensis	Chyulu Hills, Kenya	CH3	Blood	HABEL									
mbuluensis	Chyulu Hills, Kenya	CH6	Blood	HABEL									
mbuluensis	Chyulu Hills, Kenya	CH7	Blood	HABEL									
mbuluensis	Chyulu Hills, Kenya	CH8	Blood	HABEL									
mbuluensis	Chyulu Hills, Kenya	CH9	Blood	HABEL									
mbuluensis	Chyulu Hills, Kenya	CH11	Blood	HABEL									
mbuluensis	Chyulu Hills, Kenya	2CH10	Blood	HABEL									
mbuluensis	Chyulu Hills, Kenya	2CH12	Blood	HABEL									

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3		Cyt-b					
						1	2	1	2	3	4	5	6
mbuluensis	Chyulu Hills, Kenya	2CH13	Blood	HABEL									
winifredae	S. Pare Mts, Tanzania	BMNH 1935.12.23.13	Toe pad	BMHN									
winifredae	S. Pare Mts, Tanzania	BMNH 1935.12.23.10	Toe pad	BMHN									
winifredae	S. Pare Mts, Tanzania	O5899	Sequence	NCBI		DQ328436.1		DQ328336.1					
Z. borbonicus													
borbonicus	La Réunion, Mascarenes	BWM54	Blood	WARREN									
mauritianus	Mauritus, Mascarenes	BWM24	Blood	WARREN									
Z. ficedulinus													
ficedulinus	Annobón, Gulf of Guinea	FIP002	Sequence	NCBI		GU827278.1		GU827230.1					
feae	São Tomé, Gulf of Guinea	FIS003	Sequence	NCBI		GU827280.1		GU827232.1					
Z. griseovirescens	Príncipe, Gulf of Guinea	GRI002	Sequence	NCBI		GU827282.1		GU827234.1					
Z. maderaspatanus													
aldabrensis	Aldabra, Seychelles	BW301	Sequence	NCBI		DQ328487.1		DQ328387.1					
maderaspatanus	Mt Ankaratre, Madagascar	BW429	Blood	WARREN									
maderaspatanus	Mt Ankaratre, Madagascar	BW445	Blood	WARREN									
kirki	Grande Comore, Comoros	BW146	Sequence	NCBI		DQ328476.1		DQ328376.1					
anjuanensis	Anjouan, Comoros	BW252	Sequence	NCBI		DQ328482.1		DQ328382.1					
comorensis	Moheli, Comoros	BW121	Sequence	NCBI		DQ328479.1		DQ328379.1					
Z. modestus	Conception, Seychelles	BW344	Sequence	NCBI		DQ328463.1		DQ328363.1					
Z. mouroniensis	Grande Comore, Comoros	BW137	Sequence	NCBI		DQ328459.1		DQ328359.1					
Z. olivaceus													
olivaceus	La Réunion, Mascarenes	BWM49	Sequence	NCBI		DQ328454.1		DQ328454.1					
chloronothos	Maurtius, Mascarenes	BWM28	Sequence	NCBI		DQ328494.1		DQ328394.1					
S. brunneus	Bioko, Gulf of Guinea	BRU001	Sequence	NCBI		GU827267.1		GU827219.1					
S. brunneus	Bioko, Gulf of Guinea	BRU003	Sequence	NCBI		GU827268.1		GU827220.1					
S. melanocephalus	Mt Cameroon, Cameroon	MEL002	Sequence	NCBI		GU827266.1		GU827218.1					

Taxon	Locality	Sample code	Sample type	Source	TGFß2	ND3		Cyt-b					
						1	2	1	2	3	4	5	6
S. melanocephalus	Mt Cameroon, Cameroon	MEL001	Sequence	NCBI		GU827265.1		GU827217.1					
S. lugubris	São Tomé, Gulf of Guinea	LUG001	Blood	CIBIO									
S. lugubris	São Tomé, Gulf of Guinea	MM2	Sequence	NCBI		DQ328503.1		DQ328403.1					
S. leucophaeus	Príncipe, Gulf of Guinea	LEU001	Sequence	NCBI		GU827269.1		GU827221.1					
S. leucophaeus	Príncipe, Gulf of Guinea	LEU002	Sequence	NCBI		GU827270.1		GU827222.1					

Chapter 5

Conclusions on the systematics and diversification of African Zosteropidae.

Table of Contents

5.1. Conclusions

Examining the diversity and systematics of highly speciose groups such as Zosteropidae can be particularly problematic. Broad geographic ranges can often limit sampling, which is required for phylogenetic resolution of taxa across broad spatial scales. While the last decade has seen an abundance of research focusing on colonisation abilities and speciation patterns of Zosteropidae with the insular systems surrounding Africa (Warren *et al.* 2006; Melo *et al.* 2011; Milá *et al.* 2012), a lack of sampling for continental forms has hindered assessments of relationships and patterns of divergence between mainland forms. The aim of this thesis was to produce the first comprehensive molecular phylogeny for western Zosteropidae, and use it to address a number of questions regarding the relationships and patterns of diversification of mainland African forms.

Extensive sampling of East African *Zosterops* made it possible to explore one of the most geographically complex areas within the African system to examine how past climate has shaped the fragmented distribution of montane endemics in northeast Africa. The phylogeny generated revealed several poor taxonomic groupings, indicated by non-monophyly of species. Results revealed that in many cases endemic montane populations are more closely related to taxa with divergent habitat types, elevation distributions and dispersal abilities than they are to populations of restricted endemics that occur in neighbouring montane forest fragments.

Addressing the validity of current taxonomic groupings was made possible by comprehensive sampling across Africa and the Arabian Peninsula using DNA extracted from museum collections. This extended phylogeny allowed for the resolution of relationships across the African Zosteropidae complex, which includes the Arabian Peninsula, the Gulf of Guinea, the Gulf of Aden and Indian Ocean island systems, and resulted in the first comprehensive molecular assessment of patterns of systematic relationships across the range of western Zosteropidae.

5.1.1 Museum collections and the use of ‘archive’ DNA

Although DNA obtained from fresh specimens is always preferred to DNA from museum specimens, in the absence of fresh collections, sampling across broad geographic ranges is often unfeasible. The use of museum collections for DNA extraction has allowed for dense genetic sampling across mainland Africa, which has enabled the first robust molecular assessment of genetic relationships. This study

demonstrates the utility of 'archive DNA' in the absence of fresh material and highlights museum collections as an important, yet often unvalued, genetic resource. Nevertheless, it is important to recognise the destructive nature of sampling (Mundy *et al.* 1997; Payne and Sorenson 2002; Mann 2007), which conflicts with the need to maintain collections for future research. Museum collections, while vast, are not replaceable and consequently damage to samples needs to be justifiable (Payne and Sorenson 2002).

5.1.2. Molecular phylogeny and taxonomic implications

In his review of African Zosteropidae, Moreau (1967) identified that the features used to determine relationships among African taxa may be problematic and advised others to 'deal with current taxonomy with great caution'. Molecular investigation with comprehensive sampling throughout the western Zosteropidae system confirms that Moreau's (1957) caution was not unjustified. The non-monophyly of mainland African taxa demonstrates that traditional morphological characters used to delineate species within Zosteropidae are not informative in an evolutionary context, with results indicating that the current taxonomic framework greatly underestimates *Zosterops* diversity within mainland Africa. Results obtained using a combination of fresh and archive samples (Chapter 4), that give an exceptional coverage of Zosteropidae diversity across mainland Africa (including Arabian Peninsula) and its associated island systems (Gulf of Guinea, Gulf of Aden and Indian Ocean), provide the first comprehensive molecular framework for this group, that will undoubtedly form the foundation for a complete systematic review.

5.1.3 Niche divergence as a driver of speciation

Phylogenetic analysis indicates that divergence leading to the current distribution of East African *Zosterops* is more complex than previously anticipated. The non-monophyly of *Z. poliogaster* indicates that the endemic montane populations of East Africa are not relics of a previously widespread population, as shown in African Bulbuls (Roy 1997), Akalats (Roy *et al* 2001) and Forest Robins (Voelker *et al.* 2010). Consequently, the postulated montane speciation model was rejected in favour of the vanishing refuge model to explain lineage diversification of montane endemic in East Africa. In testing alternative models of speciation, results identified that niche divergence rather than niche conservatism has played a key role in the diversification of mainland African forms. The East African-centred phylogeny of Chapter 2 identifies three key biotic diversification events within the African Zosteropidae system, where

niche divergence has led to aggregates of taxa with divergent habitat types, elevational distributions and dispersal abilities.

5.1.4. Role of Plio-Pleistocene climatic fluctuations as a driver of speciation

Divergence estimates recover African Zosteropidae as a very recently diverged group (<5Ma). Results indicate that diversification of African *Zosterops* occurred during a period of climatic instability associated with the Plio-Pleistocene. It has been widely postulated that the climatic fluctuations associated with this period of climatic variability would have had a profound effect on the vegetation of Africa (deMenocal 1995; Plana 2004; Trauth *et al.* 2007), causing widespread shifts in Afro-tropical forests and leading to the intermittent fragmentation on forest areas. Results are interpreted to suggest that the effect of climatic history on ancestral divergence within African *Zosterops* is not limited to divergence between montane endemics. Instead the unstable Plio-Pleistocene African climate may have provided the primary driver for lineage diversification in all mainland African *Zosterops* lineages.

The work conducted in this thesis has dramatically changed our understanding of the relationships between African Zosteropidae. In assessing the phylogenetic placement of all 34 currently recognised mainland African subspecies; this work provided the first extensive molecular assessment for the African Zosteropidae that will undoubtedly be used as a molecular framework for a taxonomic review of this group. This work nullifies the current taxonomic framework rendering all four mainland African *Zosterops* species invalid. The widespread non-monophyly of all mainland Africa taxa demonstrates that the tradition morphological characters used to delineate species within Zosteropidae are not informative in an evolutionary context.

This work also provides an important stepping stone in our understanding of the process of diversification in mainland African Zosteropidae. Divergence estimates demonstrate that divergence within African Zosteropidae is very recent (<5Ma) coinciding with periods of climatic instability during the Plio-Pleistocene. In contrast to other avian groups studied (African Bulbul: Roy 1997; Akalats: Roy *et al.* 2001; Forest Robins: Voelker *et al.* 2010) the non-monophyly of *Z. poliogaster* indicates that the endemic montane populations of East Africa are not relics of a previously widespread population. In contrast to the widely postulated Montane speciation model, these results

provide support for ancestral populations being adaptive rather than non-adaptive with divergence events leading to aggregates of taxa with divergent habitat types, elevational distributions and dispersal abilities.

Furthermore, in generating a more broad-scale assessment of relationships within African Zosteropidae this work has identified areas of taxonomic instability within the group and indicates areas for future research. The molecular framework generated will enable future researchers to take a more informed and systematic approach to future research and sampling efforts, which will allow for more detailed and fine grain assessment of relationships within mainland Africa.

5.2. Future direction

5.2.1. Origin and evolutionary history of mainland African Zosteropidae

According to the dated molecular phylogeny of Warren *et al.* (2006), colonisation of Africa occurred after an early wave of expansion from Asia to the Indian Ocean. While the Ancient Indian Ocean white-eyes are consistently recovered at the base of a mainland African clade (Warren *et al.* 2006; Melo *et al.* 2011; Chapter 2-4), poor support for relationships within the Ancient Indian Ocean clade has made resolving the exact origin of African white-eyes difficult (Warren *et al* 2006). The biogeographic disjunction between members of the Ancient Indian Ocean clade, coupled with the lack of branch support for relationships and high mtDNA divergences, suggest that related forms (which have now become extinct) may have once existed (Warren *et al.* 2006). However, in the absence of molecular and geographic data for these putative extinct forms, very little is known regarding the spatial scale of this early expansion into the Indian Ocean region.

Extended sampling of lowland forms in northeast Africa has recovered two samples of mainland *Z. a. socotranus* as distinct from all other African taxa. Results identify this group as an independently evolving evolutionary lineage and places samples outside the major African radiation between the Ancient Indian Ocean white-eyes and all other African Zosteropidae. These samples may represent a relic population of an early expansion into mainland Africa. However, with very little known about the diversity and distribution of this form in mainland Africa. Consequently, examining the origin and evolutionary history of this group is not possible at present. These results highlight this region as focus for future sampling efforts, which may lead to a better understanding of origin of African Zosteropidae.

5.3 References

deMenocal PB (1995) Plio-Pleistocene African Climate. *Science*, **270**, 53-59.

Mann CF (2007) Use of skin from the feet of museum specimens as a source of DNA. *The Auk*, **124**, 339-339

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

Milá B, Warren BH, Heeb P, Thébaud C (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: *Zosterops borbonicus*). *BMC Evolutionary Biology*, **10**, 158-171.

Moreau RE (1957) Variation in the western Zosteropidae (Aves). *Bulletin of the British Museum,* **4**, 318–433.

Mundy NI, Unitt P, Woodruff DS (1997) Skin from feet of museum specimens as a non-destructive source of DNA for Avian genotyping. *The Auk*, **114**, 126-129.

Payne RB, Sorenson MD (2003) Museum collections as source of genetic data. *Bonner Zoologische Beiträge*, **51**, 97- 104.

Plana V (2004) Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. *Philosophical Transactions of the Royal Society B*, **359**, 1585-1594.

Roy MS (1997) Recent diversification in African greenbuls (Pycnonotidae: *Andropadus*) supports a montane speciation model. *Proceedings of the Royal Society B*, **264**, 1337-1344.

Roy MS, Sponer R, Fjeldså J (2001) Molecular systematics and evolutionary history of Akalats (Genus *Sheppardia*); A pre-Pleistocene radiation in a group of African forest birds. *Molecular Phylogenetics and Evolution*, **18**, 74-83.

Trauth MH, Maslin MA, Deino AL, Strecker MR, Bergner AGN, Dühnforth M (2007) High- and Low- latitude forcing of Plio-Plistocene East African climate and human evolution. *Journal of Human Evolution*, **53**, 475-486.

Voelker G, Outlaw RK, Bowie RCK (2010) Pliocene forest dynamics as a primary driver of African bird speciation. *Global Ecology and Biogeography*, **19**, 111-121.

Warren B, Bermingham E, Prys-Jones RP, Thebauds C (2006) Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on the Indian Ocean islands. *Molecular Ecology,* **15**, 3769-3786.