CLIMATIC PERTURBATIONS AND SPECIATION OF SOUTHERN AND EASTERN AFRICAN GREENBULS (PASSERIFOMES, PYCNONOTIDAE)

L. G. SAMPATH LOKUGALAPPATTI

Dissertation presented for the Degree of Doctor of Philosophy
at Stellenbosch University

Supervisors:
Prof. Rauri C. K. Bowie
Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, CA 94720, USA

Prof. Bettine van Vuuren
Evolutionary Genomics Group, Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa

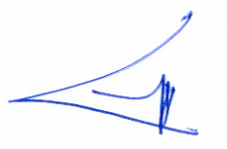
Prof. Jon Fjeldså
Zoological Museum, University of Copenhagen, Copenhagen, Denmark

May 2011

Declaration

I, the undersigned, hereby declare that the entirety of work contained in this dissertation is my own original work (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining a degree at any academic institution in the Republic of South Africa or any other country. The experimental work has been conducted in the Department of Botany and Zoology, University of Stellenbosch, South Africa and the Museum of Vertebrate Zoology, University of California Berkeley, USA.

24th May 2011
Date: ……..



……
L G Sampath Lokugalappatti

Dedicated to my parents

K.B. Nanadawathi and late Sani Lokugalappatti

Abstract

African greenbuls are an ideal model system with which to explore different hypothesis that may underlie observed patterns of lineage diversification. Among those few hypotheses amenable to falsification by phylogenetic methods concerning the diversification of the southern and east African biota, three can be singled out because of their verifiable predictions: the Plio-Pleistocene refugia, montane speciation and the gradient hypotheses. I used phylogenetic and population genetics methods to reconstruct the diversification history of three African greenbul species/species complexes (Pycnonotidae) in southern and east Africa.

My study established that most of the greenbul diversification took place in Plio-Pleistocene and the primary mechanism appears to be climatic cycling, yet dispersal and vicarince too have shaped the population genetic structure. The pattern of diversification observed in the three study taxa/species complexes differs substantially and can mostly be explained by the Pleistocene refuge hypothesis. The study did not support the montane speciation hypothesis as articulated by Roy (1997) for some of the montane *Andropadus* taxa.

Phylogeographic and population genetics analyses on the grey-olive greenbul found a close association between palaeodrainage systems and swampy areas (seeps) as an important habitat configuration for diversification of lineages restructured to these patchy habitats. Historical demographic analysis on Sombre greenbul revealved evidence for putative eastern and southern coastal forest refugia. Further, this study revealed the complex nature of East African biogeography, and two possible routes of dispersal from Albertine Rift refugia to the Eastern Arc and East African coastal forest in the Yellow-streaked greenbul complex.

Opsomming

Afrika greenbuls is 'n ideale modelstelsel wat gebruik kan word om verskillende hipotese rakend die waargenome patrone van lyndiversifisering te verken. Onder dié paar hipoteses met betrekking tot die diversifisering van die suidelike en oos Afrika biota wat ontvanklik is vir vervalsing deur filogenetiese metodes, kan drie uitgesonder word as gevolg van hulle verifieerbare voorspellings: die Plio-Pleistoseen toevlugs-, bergspesiasie- en die gradiënt hipoteses. Ek gebruik filogenetiese en bevolkingsgenetiese metodes om die diversifikasiegeskiedenis van die drie Afrikaanse greenbul spesies / spesiekomplekse (Pycnonotidae) te rekonstrueer in die suidelike en oos Afrika.

My studie bevind dat meeste van die greenbul diversifikasie plaasgevind het in Plio-Pleistoseen en die primêre meganisme blyk klimaatsverandering te wees alhoewel verspreiding en vikariansie ook bygedra het tot die algemene bevolkingstruktuur. Die waargenome patroon van diversifikasie in die drie studie taxa / spesiekomplekse verskil aansienlik van mekaar maar kan meestal verklaar word deur die Pleistoseen toevlugshipotese. Die studie het egter geen ondersteuning gevind vir die bergspesiasiehipotese soos verwoord deur Roy (1997) vir 'n paar van die berg *Andropadus* taxa nie.

Filogeografiese- en bevolkingsgenetiese ontledings op die grys-olyf greenbul het 'n noue verband gevind tussen palaeodreineringstelsels en moerasgebiede (syferwater) as' n belangrike habitatdrywer vir diversifisering van lyne wat gestruktureer is rondom verbreekte habitatte. Historiese demografiese analises op die Sombre greenbul het moontlike oostelike en suidelike kuswoud heenkomes uitgewys. Hierdie studie het ook die komplekse aard van Oos-Afrika biogeografie ontbloot, asook twee moontlike roetes van verspreiding vanaf die Albertine Rif na die Oostelike Ark asook Oos-Afrika kuswoude in die geel-gestreepte greenbul-kompleks.

Acknowledgements

I would like to express my sincere appreciation to the following people and institutions:

i) My supervisors, Profs. Rauri C. K. Bowie, Bettine van Vuuren, and Jon Fjeldså. I am greatly indebted to Prof. Bowie for initially accepting me as a doctoral student at the evolutionary genomics group (EGG), Stellenbosch University and for continuing to be my primary advisor and mentor throughout the course of my study even after he left Stellenbosch University. The opportunity he gave me to visit the Museum of Vertebrate Zoology (MVZ) at the University of California, Berkeley is greatly appreciated. Prof. Fjeldså is immensely thanked for being my co-advisor throughout the study period and for providing me with nearly half of the tissue samples used in this study. Prof. Van Vuuren is greatly thanked for accepting me as one of her doctoral students when Prof. Bowie left Stellenbosch University and for helping me to finish up my doctoral study. I am greatly indebted to all three advisors for their guidance, sustained throughout the course of study, as well as their constructive appraisal of the draft manuscripts. Their enthusiasm for research and this project in particular has greatly inspired me.

ii) The World Bank funded Improvement of Relevance and Quality of Undergraduate Education (IRQUE) project of the University of Peradeniya, Sri Lanka for providing me with sufficient funds to cover my tuition, living, and medical expenses throughout the study period. The coordinator, Veterinary science study program and the staff at the Local Technical Secretariat/IRQUE project are greatly appreciated for their continued assistance.

iii) The South African National Research Foundation, the Centre of Excellence at the Percy FitzPatrick Institute and UC-Berkeley for providing Prof. Bowie with research grants, which were extensively used for my laboratory and field research during this study.

iv) Colleagues from the EGG: Thomas Lado, Prince Kaleme, Woody Cotterill, Clement Gilbert, Liezl Mortimer, Sandi Willows-Munro, Jane Deuve, Hanneline Smit, Shelly Edwards, Ulf Johansson, Sophie von der Heyden, Conrad Matthee, Savel Daniels, Victor Rambau and Terry Robinson.

v) Colleagues from MVZ, UC Berkeley: Anna Sellas, Rebbeca Chong, Ryna Bell, Jerome Fuchs, Angela Ribeiro, Ricardo Pereira, Adam Leache, Matthew Fujitha, Sean Rovito, Anna Carnaval, Bryan Stuart, Lydia Smith, Monica Albe, Jayson MacKenzie, Jay McEntee, Andrew Rush, Norah Sarmaan. Craig Moritz, Jimmy McGuire and David Wake are thanked for their valuable comments on the methodologies used in the study. Verna Bowie is specially thanked for her continuous support during my 14 months stay in Berkeley, California.

vi) Colleagues from the Percy FitzPatrick Institute of African Ornithology and the Animal Demographic Unit of the University of Cape Town: Tim Crowe, Ian Little, Graeme Ortley, Potiphar Kaliba and Marta de Ponte Machado.

vii) Friends, colleagues and birders for numerous memorable field trips: Ian Little, Graeme Ortley, Potiphar Kaliba, David Willard, Thomas Gnoske. Special thanks to the owners of “Landela” farm and Weavers Nature Reserve, Dreadnaught Nature Reserve as well as

the staff of the South African National Parks, Cape Nature , Northern Cape Conservation and the staff of the Museums of Malawi.

viii) Curators, collection mangers and their institutions for the loan of blood/tissue and museum samples: John Bates, Shannon Hackett, David Willard and Tom Gnoske from the Field Museum of Natural History, Chicago, Illinois / Knud Jonson and Jacob Kiure from the Zoological Museum, Copenhagen, Denmark / Potiphar Kaliba and the staff from Museums of Malawi / David Allen at the Durban Museum of Natural History / Michel Louette from the Royal Museum of Central Africa, Teruuven, Belgium.

ix) For specimens and samples collected: David Willard, Tom Gnoske, Gary Voelker, Rauri Bowie, David Moyer, Jacob Kuire, Ben Marks, Norbert Kruger, Jerome Fuchs, Angela Ribeiro, Ian Little, Graeme Ortely and Potiphar Kaliba.

x) Assistance with data analyses, software packages and running programs: Jerome Fuchs, Jason Makenzie, Anna Sellas, Anna Carnaval, Ryna Bell, Adam Leache; and also Jimmy McGuire for allowing me to use the MVZ ganglia cluster; CBSU clusters at Connell University for running some of my analysis.

xi) To the faculty and instructors at the molecular evolution workshop 2008, and specially to Michael Cummings, David Swofford, Derrik Zwickel, Fredrik Ronquist, Peter Beerli and Susana Rapo

xii) I am also thankful for the opportunities to have attended the Molecular Evolution workshop at Marine Biological Laboratory, Woods Hole, MA, USA in August 2008 as well as the 125th stated meeting of the American Ornithological Union held at the University of Wyoming Laramie USA in July 2007.

xiii) The Department of Botany and Zoology of Stellenbosch University, with sincere thanks to Nico Solomons, Lydia Willems, Janine Basson, Mari Sauerman and Fawsia for technical assistance.

xiv) International office, University of Stellenbosch with special thanks to Lynda Uys, Melinda Rust and Carmien Snyman.

xv) Administrative and support staff at the Museum of Vertebrate Zoology, University of California Berkeley, Specially to Jeanne Kassis and Billy Dunaway.

xvi) Vice Chancellor, University of Peradeniya / Dean, Faculty of Veterinary Medicine and Animal Science / Head and the staff of the Department of Basic Veterinary Sciences, University of Peradeniya, Sri Lanka.

xvii) To my family, especially my dearest wife Nawodani, daughter Thilini and son Nujitha for their loving support throughout my career and the patience during this study. My mother, two brothers and the entire family-in-law are also thanked for their encouragement and particularly for taking care of my wife and two kids during the course of this study.

Contents

List of Tables

List of Figures

Figure 5.1 - Distribution range of *Phyllastrephus flavostriatus* complex in southern and eastern Africa. Distribution pattern of two colour forms are shown in the area demarcated by blue lines and the current distribution of each putative subspecies is shown in grey and brown highlighted areas in the map. White dots indicate the sample localities. Picture of grey-headed Yellow-streaked greenbul is modified from Keith (1992).

Figure 5.2 - Majority rule consensus tree calculated from 100 bootstrap trees generated using RAXML (Stamatakis 2006; Stamataskis *et al.* 2008). * indicates the nodes which received more than 75% bootstrap support and * indicates nodes which did not receive considerable support. The tree on the right indicates the phylogenetic relationship of the brown headed and two grey-headed Yellow-streaked greenbul (three basal nodes) clades and the tree on the left depict the subspecies/populations recovered as distinct terminal clusters in the analysis.

Figure 5.3 - Network produced from a NeighborNet analysis of genetic distances among individuals of the *Phyllastrephus flavostriatus* complex for the combined mtDNA data (n=248). Terminal clusters are coloured to represent different subspecies/populations. Bands of parallel edges represent incompatible splits and the value at each major split indicates the bootstrap support.

Figure 5.4 - Unrooted statistical parsimony network constructed from the combined mtDNA dataset, which resulted in eight unlinked haplotype clusters. Circle diameter is proportional to the number of individuals possessing this haplotype. Extinct or unsampled haplotypes are indicated by black dots. Numbers within the circled areas indicate the number of mutational steps connecting the adjacent haplotypes.

Figure 5.5 - Mismatch distributions for the main mtDNA lineages of the Yellow-streaked greenbul. Coloured bars in the graphs represent the observed data and red lines represent the simulated data used to test the goodness-of-fit to the sudden expansion model. Black lines represent the simulated data used to test the goodness-of-fit to the spatial expansion model. The x-axis represents the number of base pair differences and the y-axis represent the number of pairwise comparisons made. A – lineage *P. f. olivaceogriseus* from DRC, B – lineage *P. f. olivaceogriseus* from Burundi and Uganda, C- *P. f. uzungwensis*, D- *P. f. vincenti* excluding the three individuals which showed sister relationship with nominal spp from Ngoye Forest South Africa. E- *P. f. tenuirostris*, F- nominal spp from Zimbabwe, G- lineage *P. f. alfredi* from Tanzania, and H-lineage *P. f. alfredi* from Malawi and Zambia.

Figure 6.1 - Distribution range of *Phyllastrephus flavostriatus* complex in southern and eastern Africa. Distribution pattern of two color forms are shown in the area demarcated by black lines and the current distribution of each putative subspecies is shown in colored areas in the map. White color dots indicate the sample localities.

Figure 6.2 – A: Maximum likelihood majority rule consensus tree from the eight nuclear gene concatenated dataset produced from the RAxML. The color of the terminal branches correspond to the subspecies. Arrows indicate the nodes which received >75% bootstrap support and >0.95 Bayesian posterior probability. B - Species tree estimate for the *Phyllastrephus flavostriatus* complex group based on the BEST analysis of the 8 nuclear loci (mean Θ = 0.015). Values at each node indicate the estimated posterior probability.

Figure 6.3 - Statistical parsimony networks constructed with phased data using TCS v 3.1(Clement *et al.* 2000) for each nuclear intron. Circled areas in the haplotype networks are proportional to the number of individuals possessing this haplotypes (drawn to scale). Extinct or unsampled haplotypes are indicated by black dots. Black lines indicate the number of connections between haplotypes. Brown color corresponds to the brown–headed subspecies *alfredi*. Dark grey corresponds to the individual haplotypes representing the Eastern Arc, east African coastal forest and southern grey-headed form whereas, the light grey represent the individual haplotypes from the Albertine Rift grey-headed clade.

Figure 6.4 - Multilocus network produced from a Neighbor-Net analysis of genetic distances among individuals of the *Phyllastrephus flavostriatus* complex as implemented in program SplitsTrees (Huson & Bryant 2006*)* for the eight nuclear intron loci (n=70 individuals). Terminal clusters are colored to represent different subspecies/populations. Three main basal color forms are also shown. Bands of parallel edges represent incompatible splits.

Figure 6.5 - Assignment of individuals to genetic clusters using the STRUCTURE algorithm for K=5 (mean LogLikelihood across five runs, -ln=2025.1,). The blue color corresponds to the brown-headed individuals (*P. f. alfredi*) sampled in Tanzania and Malawi. Light blue corresponds to individuals sampled in the Albertine Rift and on Mt. Mahale (*P. f. grauri, P. f. olivaceogriseus* from Burundi and *P. f. kungwensis* from Mt. Mahale). Pink corresponds to *P. f. olivaceogriseus* from DRC and red represent individuals from subspecies *P. f. flavostriatus* (Ngoye Forest), *P. f. vincenti* (Mt. Mulanje). Yellow corresponds to P*. f. tenuirostris* from the Eastern Arc and East African coastal forest. Light green corresponds to the individuals from *P. f. uzungwensis*. Multicolored bars represent admixture individuals. Population one corresponds to brown-headed Yellow-streaked greenbuls, two to Albertine rft and Mahale grey–headed forms and the three correspond to the Eastern Arc, east African coastal forest and southern grey-headed forms.

List of Appendices

List of Abbreviations

Ma = Million years

Mya = Million years ago

EGG = Evolutionary Genomics Group

MVZ = Museum of Vertebrate Zoology

FMNH = Field Museum of Natural History

MOM = Museums of Malawi

ZMUC = Zoological Museum of University of Copenhagen

SE = Southeast

Mt. = Mount

PCR = Polymerase Chain Reaction

AMOVA = Analysis of Molecular Variance

BI = Bayesian Inference

ML = Maximum Likelihood

bp = base pairs

ATP6 = ATP synthase subunit 6

ND2 = NADH dehydrogenase subunit 2

ND3 = NADH dehydrogenase subunit 3

TGF = Transforming growth factor

MYO = Myoglobin gene

ODC = Ornithine decarboxylase gene

BRM15 = Fifteenth intron of the Brahma Protein gene

CHDZ = Chromosome Helicase gene

GAPD = Eleventh intron of the

mtDNA = mitochondrial Deoxyribo Nucleic Acid

DNA = Deoxyribo Nucleic Acid

nDNA = nuclear Deoxyribo Nucleic Acid

Fib5 = Fifth intron of the ß fibrinogen gene

Fib7 = Seventh intron of the ß fibrinogen gene

IM = Isolation with Migration

M = Migration

T = Divergence Time

TMRCA = Time to most recent common ancestor

ul = Microlitre

E P = Eastern Province

NP = National Park

CHAPTER 1

Introduction

1.1: African birds and molecules: a conservation perspective

The African bird fauna consists of approximately 2260 species including vagrants and regular summer visitors (Sibley & Monroe 1996) making the continent second only to South America in terms of species richness (Sibley & Monroe 1996). Africa is home to two endemic bird orders, ten endemic families (with two more only reaching Madagascar or Arabia), and of the more than 2100 species in sub-Saharan Africa (below 20^0N), almost 1400 (67%) are restricted to this region (Sinclair & Ryan 2003). These numbers have increased gradually over the years as an outcome of new discoveries (e.g. Beresford *et al*. 2004; Fjeldså *et al.* 2006; Bowie *et al.* 2009; Voelker *et al.* 2010), as well as from recent taxonomic revisions (e.g. Ryan & Bloomer 1999; Bowie *et al.* 2004a, 2005).

The extensive role that modern taxonomic revisions of many bird groups is playing in increasing the number of species delineated stems from many taxa, referred to with binominals in the 19th and early 20th centuries, having been lumped as subspecies under a broad Biological Species Concept in various volumes of Peters' Checklist and in the 'Speciation Atlases' of African birds (Hall & Moreau 1970; Snow 1978 – see Fjeldså 2003 for discussion). Taxonomy of birds based on the Biological Species Concept has been subject of much debate, primarily because subsuming of multiple taxa into a single species, each of which has its own history of speciation could disguise informative patterns of distribution and speciation (Barrowclough 1992). Further, underestimates of the species richness in such taxonomic treatments may misalign conservation efforts of unique taxa/ecosystems with distinct evolutionary histories (de Klerk 1998). However, despite the potential conservation importance, it is essential that an objective taxonomy based on sound data collection and analyses, without the use of non-biological criteria, be used when assigning species rank to taxa, if meaningful insights into evolutionary history are to be gained, and effective conservation strategies are to be developed and implemented (Bowie & Fjeldså 2005).

Recent advances in molecular genetics have facilitated the use of phylogenetic and phylogeographic approaches to help objectively resolve the species status of morphologically distinct subspecies, as well as enabled the detection and delimitation of cryptic taxa. Phylogenetic studies have been widely used to solve taxonomic uncertainties and for the delimitation of range boundaries, whereas phylogeographic methodologies have been used for studying species' evolutionary histories in the context of palaeoenvironmental change (e.g. Hewitt 1996, 2004; Taberlet *et al.* 1998; Tribsch & Schonswettter 2003). Specifically, phylogeography involves the investigation of geographical distributions and relationships in terms of genealogy and demographic history of individual species or a group of related taxa (Sgariglia & Burns 2003). The relationships that exist among lineages and their distributions today likely arose as a result of dynamic processes along both temporal and spatial scales (e.g. Bowie *et al.* 2004a), which makes a phylogeographic approach necessary for adequate interpretation of pattern and process (Templeton 1998). The phylogenetic analysis of DNA sequences, when superimposed over a geographical area may help explain whether the history of populations has been one of isolation, panmixia or a combination of both (Pavlova *et al.* 2003).

Insights from molecular systematics and population genetics, when combined with analyses of species distributions, phenotypic variation, and landscape history provide a context with which to improve strategies for conservation (Moritz *et al.* 2000; Moritz 2002; Kahindo *et al.* 2007). Prioritization of areas for conservation ideally requires that data from these various components be integrated. Such integrated molecular genetic approaches are particularly important for understanding the reality of morphologically cryptic species with localized molecular endemism within widely distributed species complexes (Fuchs *et al.* 2008). Erwin (1991) and Smith *et al.* (1993) have pointed out that for conservation planning to be successful, steps need to be taken to preserve not only the pattern of biodiversity, but also the evolutionary processes that generate and maintain it. In this context, the use of genetic markers to infer evolutionary histories at

the intra- or inter-specific level, to help prioritize areas for conservation, is widely recognized (Avise 1994; Humphries *et al.* 1995; Smith *et al. 1997*; Moritz & Faith 1998; Kahindo *et al.* 2007). One useful approach is to compare intra-specific phylogeographies of co-distributed species (i.e. comparative phylogeography) to identify important geographical areas in which populations have undergone independent evolution (Avise 1992; Faith 1992; Crozier & Kusmierski 1994; Moritz & Faith 1998).

1.2: Historical geography, palaeoclimate, and avian evolution in southern and eastern Africa

The historical geographical processes that can shape the genetic structure of a species include population division, range expansion, isolation-by-distance (i.e. clinal signatures in genetic variation) and long distance colonization, which are expected under a neutral model of evolution to produce distinct patterns and relationships among haplotypes/alleles (Templeton *et al.* 1995). According to Partridge *et al.* (1997) changes in regional topography caused by the "African super swell" (see Nyblade & Robinson 1994 for details) an extensive period of uplift before 2 Myr BP caused the eastern lowland forest block (coastal) to become isolated from the main Guinea-Congolian rainforest block (Lovett 1993; Coppes 1994), which in turn had profound effects on the palaeoclimate of Africa. Another significant geological change that affected the palaeoclimate and hence the biota of Africa during the Tertiary was the northward drift of Africa and eventual "collision" with Asia in the early Miocene (Axelrod & Raven 1978), which resulted in the gradual aridification of Africa and the formation of an `arid corridor' isolating the eastern forests from the much larger central and western forest blocks (Axelrod & Raven 1978; Lovett 1993). Development of rift valley systems (see Rosendahl 1987; Chorowicz, 2005 for details) in the eastern and southern parts of Africa during the middle to late tertiary have been postulated to have had a tremendous influence on the climatology of the region and to have resulted in vicariance for many

taxa (Partridge *et al.*1997). From the mid-Cretaceous to mid-Tertiary there is no evidence of a major tectonic disturbance in Africa (King 1978) suggesting that for many taxa, range dynamics have been largely confined to the Tertiary.

Africa has undergone several major episodes of climatic change since the Pliocene/Pleistocene (van Zinderen Bakker 1978; deMenocal 1995, 2004), that are likely to have caused shifts in vegetation pattern (Smith *et al.* 2004). These Plio-Pleistocene climate changes in Africa are probably one of the main causes of speciation in African bird species (Voelker 1999; Bowie *et al.* 2004a, 2004b, 2006; Fjeldså & Bowie 2008; Voelker *et al.* 2010). According to deMenocal (2004) large climate changes alter landscape ecology in such a way that it leads to particular faunal adaptations or speciation pressures. The basic mechanism is centered on climatic changes leading to fragmentation of the habitat occupied by these birds, which then leads to the isolation of individuals and hence genetic drift or selective adaptation to localized environments.

Taken together, the African faunal (human and other mammal fossil records of east Africa) and palaeoclimatic (e.g. eolian dust from deep sea core drills, pollen cores from rock hyrax middens, cave sediments, charcoal deposited etc.) records suggest three restricted intervals (2.9-2.4 Ma, 1.8-1.6 Ma and 1.2-0.8 Ma) when shifts toward increasingly variable, drier African conditions were accompanied by changes in African faunal assemblages and perhaps speciation (see deMenocal 2004 for details). These same datasets have revealed a long-term trend toward greater abundance of arid-adapted taxa during the Pliocene-Pleistocene, with indications of faunal change between 2.9 and 2.4 Ma and after 1.8 Ma. Further, the current literature on palaeoclimatic records does not support unidirectional shifts to permanently drier conditions in contrast to many habitat specific evolutionary hypotheses (e.g. Savanna hypothesis: Dart 1925; Bartholomew & Birdsell 1953). Current interpretations of the savanna hypothesis state that the evolution of African mammalian fauna, including early hominids, was primarily attributable to the step-like development of cooler and drier and more open conditions that occurred since the late Miocene. The mid-

Pliocene aridification shift near 3.2 - 2.6 Ma (de Menocal 1995), in particular, is viewed to have favored the evolution of arid-adapted fauna and to have influenced early hominid evolution and behavior (de Menocal 1995, 2004). Denys *et al.* (1985, 1986) have found close correlations between the Mio-pliocene evolution of rodents and the tectonic history of the eastern branch of the east African rift. Isolation of savanna basins seems to have played a major role in forcing rapid evolution of the endemic rodent populations. This suggests that the same structural and climatic history has favored the birth of savanna-adapted hominoids and *Homo* species. Since 2 Ma, disappearing of barriers in the eastern branch of the rift has facilitated dispersal, which is reflected in the palaeoclimatic record as an increasingly more even distribution of rodent populations in the savannas, as well as the migration of *Homo* species out of east Africa. There is also evidence for periodic aridity alternating with periods of raised lake levels in the Pleistocene (Livingstone 1975, 1993; Hamilton 1976, 1982; Bonnefille 1983; Dupont & Hooghiemstra 1989; Maley 1991; deMenocal 1995) and Holocene (Nicholson 1994), and it has been suggested that these changes were the most important determinants of the biogeography of African forest biota (Moreau 1966; Schiùtz 1967; Laurent 1973).

1.3: An overview of models of speciation

Several models have been proposed and tested to reveal the effect of vegetational shifts and landscape modifications during palaeoclimatic changes on vertebrate faunal assemblages through space and time. Many researchers have emphasized allopatric speciation for forest birds either in refugia (Haffer 1969; Mayr & O'Hara, 1986) on isolated mountains (Moreau 1966; Fjeldså & Lovett 1997; Roy 1997; Graica-Moreno & Fjeldså 2000; Bowie *et al.* 2004a), in scrap and lowland forests (Diamond & Hamilton 1980; Mayr & O'Hare 1986; Prigogine 1987,1988; Lawes 1990), or across major rivers and other topographic barriers (Wallace 1852; Haffer 1992; Moritz *et al.* 2000) as likely possibilities.

Of the above models, the refuge model has been the most widely discussed (Moreau 1966; Fjeldså & Lovett 1997; Roy 1997; Graica-Moreno & Fjeldså 2000; Bowie *et al.* 2004a, 2004b; Fjeldså & Bowie 2008; Voelker *et al.* 2010) and rests on the premise that climatic change caused rainforests to contract to refugia, that were separated by dry arid scrublands and savanna, and that this isolation promoted speciation. Initial discussions of the refuge model were focused on Pleistocene events, particularly those from the last glacial cycle or two (Haffer 1969; Diamond & Hamilton 1980; Crowe & Crowe 1982; Mayr & O'Hara 1986). The model has been extended to Tertiary events on the assumption that climatic oscillations driven by Milankovitch cycles throughout this period were of sufficient amplitude and duration to promote speciation (Haffer 1993, 1997). This has been partly in response to studies of African birds which have demonstrated that many speciation events among sister taxa predate the Pleistocene (e.g. Roy 1997; Roy *et al.* 2000; Bowie *et al*. 2004a, 2005; Fjeldså & Bowie 2008; Voelker *et al.* 2010).

"The Montane Speciation Model" has been suggested as an alternative to the Pleistocene Refuge Model (Fjeldså & Lovett 1997; Roy 1997; Roy *et al.* 1998). This model hypothesizes, that speciation events primarily take place in topographically complex places, where orographic moderation of the weather creates high eco-climatic stability on certain slopes and basins. Once sufficient population sizes and genetic diversity have accumulated, these 'younger species' are redistributed within the lowland rainforest, thereby erasing much of the historical biogeographic signal of these habitats. Thus, under the Montane Speciation Model, the lowland forests acts as a 'museum' for old and relict species (Fjeldså & Lovett 1997; Roy 1997; Roy *et al.* 1997; Fjeldså *et al.* 2007; Fjeldså & Bowie 2008). However, the recent work by Marks (2010) on Green Hylia (Aves: Sylviidae: *Hylia prasina*) in western Africa has challenged this view of lowland rainforests as evolutionary 'museums' where biotic diversity is maintained over evolutionary time, and additional diversity is accrued from peripheral areas, but where there has been little recent diversification. In another study conducted on lowland forest African Illadopsis species (Ngumbeck *et al.* 2009)

suggested that most splits in *Illadopsis* occurred synchronously around the Plio-Pleistocene transition, suggesting that some diversification events in African forest taxa took place before the onset of the large-amplitude climatic cycles of the Pleistocene epoch. Thus, the diversification of African taxa in time and space is thought to be more complex than being restricted to the Pleistocene traditionally associated with the diversification of African forest taxa. The results also suggest that deep genetic divergences do exist among species complexes of lowland and montane African birds, which differ only slightly in morphological characters (Ngumbeck *et al.* 2009).

The riverine model posits that substantial river systems act as barriers to gene flow, such that populations on either side gradually diverge to form separate species (Wallace 1852). The "riverine barrier" model has been demonstrated using various bird species (Haffer 1978, 1992 & 1997), monkeys (Hershkovitz 1977) and Amazonian lizards (Avila-Pires 1995) across the major rivers of Amazonia. However, Patton *et al.* 1994 on the genealogical differentiation of arboreal spiny rats across Amazon and a recent work by Gascon *et al*. 2000 on the phylogeography of frogs and small mammals in Amazonia have indicated that a putative riverine barrier does not relate to present-day patterns of community similarity and species richness, thus significantly weaken the postulated role of rivers as the major driver of Amazonian vertebrate diversification.

According to Moritz *et al.* (2000) the gradient model suggests that strong environmental (e.g. habitat) gradients result in rapid adaptive divergence in morphology often with reinforcement, thus leading to speciation. This is expected to result in sister species adapted to adjacent but distinct environments (e.g. rainforest versus dry forest/savanna). Gradient models differ fundamentally from allopatric/vicariance models in that complete suppression of gene flow is not a prerequisite for phenotypic divergence and speciation (Endler 1977; Rice & Hostert 1993; Smith *et al.* 1997). Models that emphasize the role of natural selection along ecological gradients between parapatric populations (Endler 1982a, b; Smith *et al.* 1997; Schneider & Moritz 1999; Schneider *et al.* 1999) have also proved popular. For example, ecotones have recently been suggested by

several researchers (Alexio 2002; Smith *et al*. 2004) to be an important habitat configuration for parapatric speciation of tropical forest species, birds in particular. As mentioned above, these gradient or parapatric models of speciation contrast sharply with the generally accepted models of speciation which suggest that the accentuated global climatic cycles since the Plio-Pleistocene (2.8 Ma) have caused the intermittent fragmentation of forest regions into isolated refugia thereby providing a mechanism for speciation of the tropical forest biota contained within them (Roy 1997).

1.4: Avian habitats in southern and eastern Africa

The montane regions in eastern Africa, associated with the geological events described above, occur along the transition between the Congo Basin and east Africa (the Albertine Rift mountains of the eastern DRC, Rwanda, Burundi and western Uganda), but otherwise form a discontinuous circle of inselbergs following clockwise from:-the Albertine Rift to include - the Kenya Highlands, the Eastern Arc crystalline fault-blocks in Tanzania and southeastern Kenya, the Malawi Rift Mountains, and uplifted escarpments locally along the Ufipa plateau and Lake Tanganyika (Figure 1.1). This circle is partly connected with the mountain scarps of the Ethiopian highlands to the north and the highlands of eastern southern Africa. It is widely recognized that this 5000 km mountain chain provided a colonization route for many taxa from east African forest refugia to southern Africa (Lawes *et al.* 2007 and references there in).

The Albertine Rift is the most species rich region for vertebrates on the African continent (Brooks *et al.* 2001; Plumptre *et al*. 2003, 2007). This part of Africa contains the 'Mountains of the Moon' or Ruwenzori Massif that includes Africa's third highest peak, the Virunga Volcanoes made famous by its mountain gorillas, active volcanoes in the Virunga National Park, and Lake Tanganyika – Africa's deepest lake. The Albertine Rift encompasses much of the western Rift valley, which extents to southern Tanzania and northern Zambia.

The Eastern Arc Mountains of Tanzania and southern Kenya are also well known for their high concentrations of endemic species of animals and plants. Thirteen separate mountain blocks comprise the Eastern Arc, supporting around 3300 km^2 of sub-montane, montane and upper montane forest, less than 30% of the estimated original forested area. At least 96 vertebrate species are endemic: 10 mammal, 19 bird, 29 reptile and 38 amphibian species (Burgess *et al.* 2007). Afromontane forests are found in most parts of the montane circle and usually occur about 1800m above sea-level. As one moves south, latitude starts to compensate for elevation, enabling montane taxa to occur at sea-level in South Africa. In addition to montane forests, this area also supports coastal forest, savanna-woodland, thicket, broadleaf woodland forest, mangroves, grassland and extensive tracks of habitat modification for farmland (see White 1983; Lovett & Wasser 1993).

Coastal forests are found along the 4000 km stretch of the coastal belt extending from SE Somalia to the eastern Cape Province in South Africa (see Stattersfield *et al.* 1998; Lincoln *et al.* 2001 for details) are often described as regions of mixed vegetational composition, most commonly named the Zanzibar-Inhambane floral mosaic and Tongoland-Pondoland regional mosaic (White 1983) owing to the floral heterogeneity. Coastal habitats generally occur below 700 m. In between montane and coastal forest habitat, various types of lowland habitats such as savanna and miombo woodland forests persist (White 1983). In montane habitats of east Africa it is often possible to recognize stratification of vegetation in association with the altitude (see White 1983; Hedberg 1986 for details). Within these diverse habitat categories lives an evolutionary complex forest biota including many endemic bird species (Fjeldså & Louvett 1997; Burgess *et al.* 1998; Fjeldså & Burgess 2008).

1.5: Phylogeographic studies on African birds

Of the reported 2260 species of African birds, phylogeographical studies using DNA markers have been conducted on only a very few species.

Open Habitats: Freitag and Robinson (1993) have studied the phylogeographical structure of the Ostrich (*Struthio camelus*) using mtDNA markers. They established that there was little genetic divergence among individuals from southern Africa, whereas the individuals from eastern and northern Africa had much higher diversity in their mtDNA. A subsequent study by Robinson and Matthee (1999) which also included the extinct Arabian ostrich, *Struthio camelus syriacus* revealed recent common ancestry for the southern African *S. c. australis* and the east African *S. c. massaicus*. The same study suggested that gene flow between the two geographic forms may have taken place in the recent evolutionary past, probably along the Egyptian–Sinai–Israel passageway.

A mitochondrial DNA based study of the arid-adapted Long-billed Lark (*Certhilauda curvirostris*) complex in Southwestern Africa (Ryan & Bloomer 1999) discovered substantial sequence divergence and revealed several cryptic taxa. In addition, a mitochondrial DNA and morphometrics based study on the Karoo Lark (*Certhiladua albescens*) complex resulted in the recognition of a new species, Barlow's Lark (*Certhilauda barlowi*), in southwestern Africa (Ryan *et al.* 1999). Studies on species of pipits and wagtails conducted at the intercontinental level (e.g Voelker 1999a and 1999b; Voelker 2002) have revealed valuable insights into how the diversification of African birds has taken place in open habitats. Pipits colonized Africa around five to six million years ago and the present distribution and diversity of the genus *Anthus* in Africa is mainly due to multiple dispersal events between continents rather than being solely the result of within-continent speciation. Similarly, wagtails colonized Africa around 4.5 Mya, and accomplished a second colonization about 2.9 Mya.

Forested Habitats: Roy (1997) investigated the evolutionary history of the African greenbuls (genus *Andropadus*), using a molecular approach. This analysis revealed that montane species are a recently derived monophyletic group when compared to lowland species of the same genus and that recent speciation events (within the Plio-Pleistocene) have exclusively occurred in montane regions supporting the view that montane regions have acted as centers of speciation during recent climatic instability. Roy described this as the "montane speciation hypothesis" (see above). Further details about the radiation of African bulbuls are included in Chapter 2 and form the primary focus of this thesis.

The molecular systematics and evolutionary history of African forest robins or Akalats (Genus *Sheppardia*) was initially described by Roy *et al.* (2000). This study suggested that a rapid Pre-Pleistocene radiation occurred, with more recent intraspecific population expansion in the upper Pleistocene being concordant with major climatic variation and vegetational changes during the last three to 5 Ma. Further studies on Eastern Arc Akalats has revealed reciprocal monophyly of several populations including one form that is related to *Sheppardia lowei* and *Sheppardia montana* (Beresford *et al*. 2004). These data were used together with morphology to describe a new species of Akalat, *Sheppardia aurantiithorax* and demonstrated relatively high levels of sequence divergence among populations of the different species isolated on small inselbergs within the Eastern Arc. A recent study by Voelker *et al.* (2010) based on part of the African forest robin assemblage with robust estimates of phylogenetic relationships and clock-based divergences rejected Miocene tectonic uplift and Pleistocene forest refugia as the primary drivers of speciation in forest robins. Instead, the data suggest that most speciation took place in the Late Pliocene, from 3.2 to 2.2 Ma. The same data set also suggested that most montane speciation resulted from the rapid isolation of populations in montane areas, rather than montane areas themselves being drivers of speciation, thus providing additional evidence that Pliocene climate change was a major driver of speciation in broadly distributed African animal lineages. They further show that lowland

forest robins are no older than their montane relatives, suggesting that some lowland areas may not solely have acted as "museums" housing more relict (deeper phylogenetic branches) taxa.

The influence of climatic cycles on forest fragmentation and subsequent population expansion as forests have expanded under more favourable climatic conditions have been demonstrated in African greenbuls (Roy 1997), Sunbirds (Bowie *et al.* 2004 a,b) and Thrushes (Bowie *et al.* 2005). Results of a study by Bowie *et al.* (2004a) on the Olive sunbird (*Nectarinia olivacea/obscura*) clearly indicated that this species had a vast range expansion during the Pleistocene. Bowie *et al.* (2004b) concluded that modern distributions of Double-collared sunbirds (*Nectarinia moreaui/loveridgei/mediocris*) occurring in the Eastern Arc Mountains reflect allopatric speciation following loss of forests along dispersal corridors. This was caused by severe aridity. Thus, populations were effectively isolated in climatically more favourable and stable areas. It is generally accepted that climatic cycles beginning in the Plio-Pleistocene (2.8 Mya) have caused the intermittent fragmentation of forest into isolated refugia in east Africa, thereby providing a mechanism for speciation. However, for many bird species that exhibit phylogeographic structure, this hypothesis is still not certain, although several studies have suggested that a number of bird taxa diversified before the onset of the Pleistocene (Klicka & Zink 1999; Garcia-Moreno & Fjeldså 2000: Bowie *et. al*. 2004b). Thus, although a forest refuge pattern appears to be operating within Africa, the timing of refuge formation is not constrained to the Pleistocene (Voelker *et al.* 2010).

A study focused on the Starred robin (*Pogonocichla stellata;* Bowie *et al.* 2006) was the first to investigation how the indirect aridification of Africa caused by global cooling in response to glacial cycles at higher latitudes has influenced the evolutionary history of an African montane bird. At least two major vicariant events, one that separated the Albertine Rift from all but the Kenyan Highlands around 1.3–1.2 Mya, and another that separated the Kenyan Highlands from the northern Eastern Arc, and the northern Eastern Arc from the south-central Eastern Arc between 0.9 and 0.8 Mya appear to underlie much of the observed genetic structure among Starred robin

populations. Bowie *et al.* (2006) used the same study to demonstrate that the maintenance of gene flow between montane populations could occur through lowland coastal forest and that this has likely been taking place for a long time, suggesting that for some species each habitat type should be considered independently.

Another well-documented phylogeographic pattern observed in montane bird species in eastern African are phylogeographic breaks to gene flow among geographically close populations. For the Eastern Arc Mountains, at least four bird datasets, namely Sunbirds (Bowie *et al.* 2004a), Thrushes (Bowie *et al.* 2005), Akalats (Beresford *et al.* 2004) and the Green barbet (Thomasset *et al*. *unpublished data*) reveal a congruent phylogenetic break between the northern and central Eastern Arc, which suggests that many of these lineages have been evolving separately for several million years.

Faunal turnover between eastern and southern Africa has been tested using phylogeographic studies of montane taxa (e.g. for birds Bowie *et al.* 2005). Bowie *et al.* (2005) have shown that the central and southern Eastern Arc, and northern Malawi Rift taxon *Turdus abyssinicus nyikae* is closely related to the Albertine Rift taxa *T. a. baraka* and *T. a. bambusicola* and Kenyan highland *T. a. abyssinicus*. These taxa are both genetically and morphologically distinguishable from the Eastern Arc endemics *T. helleri* and *T. roehli*, as well as the southern African clade *T. olivaceus*. The sister relationship of these clades suggests that southwestern East Africa, central Africa and the Kenyan highlands were biogeographically linked. However, little phylogeographic data are currently available on the geographic position of places where faunal turnover occurs between eastern and southern African vertebrate species, with regard to any of the following habitat communities: coastal-lowland species, montane-coastal species and broad-leafed woodland taxa. At least three phylogeographical breaks have been proposed in the Malawi Rift according to a recent study conducted on two forest bird species and a small mammal (Kaliba 2006).

1.6: Scope of the dissertation

One of the remarkable features of the African avifauna is the interspecific and intraspecific heterogeneity among its species assemblage. Intraspecific morphological polymorphism has been described in many African birds, including for example in the endemic greenbuls (genus *Andropadus*, Roy 1997; Roy *et al.* 1998), and Double-collard sunbirds (genus *Nectarinia*, Bowie *et al*. 2004b). These differentiated forms are sometimes distributed over the entire range of the species, or limited to a narrow geographic area. Further, the same species' is sometimes found in different habitats such as costal forest (at sea-level) and tropical montane rainforest (over 2000 m). Such variation has resulted in considerable dispute over the taxonomic status and delineation of range boundaries of described races/subspecies, which occupy different habitats, or occur in allopatry. Taxonomists frequently use the term "species complex" to describe such complex species/subspecies groups as they comprise related but distinctive populations.

Another noteworthy feature of the African avifauna is the highly disjunct present-day distribution patterns of some bird species, especially among montane rainforest species (Bowie *et al.* 2004a. This disjunct pattern and often associated morphological polymorphism among bird species has led to much debate over taxonomic status and range boundaries, as well as speculation on how these patterns may have evolved. Hence this area has become of key research interest in recent African bird studies. This is especially true of phylogenetic and phylogeographical studies, which offer sophisticated methodologies with which to separate past and present determinants of species limits (e.g. Bowie *et al.* 2004a, 2004b, 2006). It is important to investigate what evolutionary mechanisms underlie the present distribution pattern and pattern of genetic relationships of taxonomically uncertain members of species complexes in order to resolve these disputes and to identify important geographical areas in which populations have undergone independent evolution (Avise 1992; Faith 1992; Crozier & Kusmierski 1994; Moritz & Faith 1998).

One of the most morphologically complex groups of passerine birds are the bulbuls and

greenbuls (Family: Pycnonotidae). Greenbuls show a high degree of morphological variation, as well as a wide array of distribution patterns and habitat preferences in sub-Saharan Africa (see Chapter 2 for details). Hence, greenbuls are an ideal clade with which to test different hypothesis centered on the role that climatic and geological perturbations have had on the speciation of birds in eastern and southern Africa (e.g. Roy 1997; Smith *et al.* 2004). Greenbuls form the focal taxa of this study with chapters centered on the following three taxa/species complexes: Sombre greenbul (*Andropadus importunus*), Grey-olive greenbul (*Phyllastrephus cerviniventris),* and Yellow-streaked greenbul *(Phyllastrephus flavostriatus*).

The aims of this dissertation are of two-fold:

First, to determine the evolutionary relationship among and within eastern and southern African bulbul/greenbul species complexes (Passeriformes, Pycnonotidae), with special emphasis on the three species/subspecies complexes: Grey-olive greenbul (*Phyllastrephus cerviniventris)* – Chapter *3,* Sombre greenbul (*Andropadus importunus*) – Chapter 4*,* and Yellow-streaked greenbul *(Phyllastrephus flavostriatus*) – Chapters 5 and 6. The data generated during the study will also be included in collaborative research on the higher-level phylogeny of African greenbuls and bulbuls (Chapter 2).

Secondly, to determine the phylogeographic population structure and historical biogeographic patterns that have led to the current distribution patterns of the three species/species complexes mentioned above in eastern and southern Africa, and to elucidate the most likely hypothesis (e.g. Pleistocene Refugia Model, Haffer 1997; Mayr & O'Hara, 1986; Montane speciation hypothesis, Fjeldså & Lovett 1997, Roy 1997, Roy *et al.*1997), that explain the observed spatial structure.

The specific objectives are as follows.

• To investigate the historical biogeographical processes that have led to current distribution patterns of the Sombre greenbul (*Andropadus importunus),* Grey-olive greenbul (*Phyllastrephus cerviniventris)* and Yellow-streaked greenbul *(Phyllastrephus flavostriatus).*

• To investigate the roles of gene flow and climate change on the phylogeographic and demographic history of the three species/species complexes detailed above.

Figure 1.1 - Map of the east African montane circle of mountains. Blackened areas indicate highland regions with montane forest. Light grey areas indicate lowland rainforest (map modified from Roy 1997).

CHAPTER 2

A systematic review of the African greenbuls (Pycnonotidae)

2.1: Introduction

A thorough understanding of phylogenetic relationships among species and clear delineation of monophyletic groups is a fundamental step if appropriate biogeographical inferences are to be made about those groups. Bulbuls and greenbuls (family: Pycnonotidae) represent one of the most taxonomically complex groups of passerine birds (Keith 1992) partly owing to their morphological similarity. Despite the fact that studies of the relationships among members of the Pycnonotidae have a long history, the phyolgenetic relationships and inferred patterns of speciation among the bulbuls continue to engender debate. Therefore, to gain insight into the historical biogeography of African greenbuls, as well as following the established tradition of thesis preparation, I present an up-to-date review of the literature on greenbul phylogeny and diversification in this chapter. The review begins with a summary of the research conducted in mid-1950s through to the 1990s, and concludes with the most recent contemporary studies, including selected sections from two recent publications, which made use of data generated from this thesis – primarily outgroup taxa with respect to my core data chapters (3-6). The references for the aforementioned publications are as follows.

Johansson, U.S., Fjeldså, J., **Lokugalappatti, L.G.S.** and Bowie, R.C.K. (2007) A nuclear DNA phylogeny and proposed taxonomic revision of African greenbuls (Aves, Passeriformes, Pycnonotidae). *Zoologica Scripta* 36: 417-427.

Fjeldså, J., Johansson, U.S., **Lokugalappatti, L.G.S.** and Bowie, R.C.K. (2007) Diversification of African greenbuls in space and time: linking ecological and historical processes. *Journal of Ornithology* 148: S359-367.

2.1: Bulbuls and greenbuls (Passeriformes, Pycnonotidae)

Phylogeny: Bulbuls and greenbuls represent an interesting group of medium sized, short-necked songbirds with short round wings, a fluffy plumage. Some species have crests on their heads. Bulbuls are distributed in forests and wooded patches of the Afrotropical and Indo-Malayan biogeographic regions and are one of the few avian lineages, that are restricted exclusively to the Paleotropics. Only the genus, *Pycnonotus*, is shared between Africa and Asian (Pasquet *et al*. 2001; Moyle & Marks 2006; Johansson *et al.* 2007), and the genus *Hypsipetes*, colonized the western Indian Ocean islands (Madagascar, Comoros, Mascarenes, Seychelles) from Indo-Malaya (Warren *et al*. 2005). Bulbuls inhabit a wide array of habitats from sea-level to elevations of 3000+ meters (Keith 1992) and exhibit great variation in range size. Some of the bulbuls have very wide distribution ranges (e.g., *Pycnonotus barbatus*) others are narrow endemics (e.g., *Pycnonotus capensis*). Some have highly disjunct distribution patterns (e.g., *Phyllastrephus flavostriatus),* or have narrow but spatially continuous distributions (e.g., *Andropadus importunus*).

The placement of the Pycnonotidae within Passeriformes is uncertain, as are the relationships of the genera to each other (Keith 1992). Bulbuls have been allied with a number of families as diverse as the Campephagidae (cuckoo-shrikes), Dicruridae (drongos), Sturnidae (starlings) and Oriolidae (orioles) depending on the types of characters compared (Sibley 1970). DNA-DNA hybridization data (Sibley & Ahlquist 1990) suggests that bulbuls belong to the superfamily Sylvioidea, between kinglets (Regulidae) and the African warblers (Cisticolidae). However, a more recent DNA-DNA hybridization study by Sheldon and Gill (1996) has suggested that bulbuls belong to the superfamily Sylvioidea, situated between old world warblers (Sylviidae) and swallows (Hirundinae), contrary to Sibley and Ahlquist's (1990) placement. Many subsequent studies based on nuclear and mtDNA sequence data have indicated that bulbuls are more closely related to old world warblers (Sylviidae), swallows (Hirundinae) and babblers (Timaliidae) than to kinglets or the Cisticolidae (e.g., Chikuni *et al*. 1996; Honda & Yamagishi 2000; Ericson &

Johansson 2003; Barker *et al*. 2002, 2004; Beresford *et al*. 2005; Spicer & Dunipace 2004; Per Alström *et al.* 2006; Johansson *et al.* 2008). As hypothesized for the Passerida, bulbuls may be of African origin (Fuchs *et al.* 2006).

According to Jean Delacour (1943) the bulbuls constitute one of the most clearly defined groups of perching birds (Passerines). Despite Delacour's view, the family has been subjected to major taxonomic revisions (Dowsett & Dowsett- Lamarie, 1980; Roy *et al*. 1998; Pasquet *et al*. 2001, Moyle & Marks 2006; Johansson *et al*. 2007).

Keith (1992) in the *Birds of Africa compendium* considered the family to consists of 15 genera and more than 123 species worldwide. Of the roughly 123 species, about 56 (in 13 genera) are African (Keith 1992); widely distributed across the continent, being absent only from desert regions. Subsequent, DNA analyses have suggest that the African and Asiatic genus *Criniger*, united chiefly by the conspicuously colored throat, is not a natural group; it has been suggested that the Asian *Criniger* taxa take the genus name *Alophoixus*, and the African species remain in *Criniger* (Pasquet *et al*. 2001). Before 1989, the genus *Phyllastrephus* was considered to be distributed in both Africa and Madagascar. Owing to their distinct evolutionary history as revealed by Olson (1989) and subsequently by Fjeldså *et al*. (1999), *Phyllastrephus* species found in Madagascar were removed from the Pycnonotidae and are now placed in their own family the Bernieridae (Cibois *et al.* 2001; Moyle & Marks 2006; Johanson *et al.* 2008).

Dickinson (2003) recognizes 23 genera and 118 species of bulbuls (Passeriformes: Pycnonotidae), including *Alophoixus* (Pasquet *et al.* 2001; Moyle & Marks 2006). With the suggested revisions of Johansson *et al.* 2007 for the African greenbuls (genus *Andropadus*), as well as the recent work on Philippine bulbuls by Oliveros and Moyle (2010) the total number of species may well exceed 135. These recent studies highlighted two main points regarding the evolutionary history of the Pycnonotidae: 1) several genera are para- or polyphyletic (e.g., *Andropadus*, *Criniger*, *Pycnonotus*), and 2) two or three main lineages contain a strong

biogeographic component where two of these clades are endemic to Africa, while the third includes all Indo-Malayan genera as well as *Pycnonotus* and *Hypsipetes*.

The endemic and enigmatic African genera *Neolestes* and *Nicator* have been subject to much taxonomic dispute (Keith 1998; Beresford *et al*. 2005). Because the three *Nicator* species and *Neolestes* lack a thin sheet of nostril-covering bone that is present in the rest of the bulbuls (Olson 1989), they have been placed elsewhere. According to Dowsett *et al*. (1999) the striking plumage of the Black-collared bulbul (*Neolestes torquatus*) allies it with the shrikes (Malaconotidae, Laniidae or Prionopidae). DNA-DNA hybridization data allied *Neolestes* with other bulbuls (Sibley & Ahlquist 1990), but its placement within the family could not be determined. Using an extensive DNA dataset, Zuccon and Ericson (2010) recently confirmed that *Neolestes* is not just a basal branch of, or the sister group to the bulbuls; but instead is embedded within the African radiation of bulbuls.

Similarly, *Nicator* has also been allied with the shrikes, but feather protein and early DNA evidence suggested affinities with bulbuls (Hanotte *et al.* 1987). More DNA studies using various nuclear genes (Beresford *et al*. 2005; Johansson *et al.* 2008) consistently places *Nicator* as a deep branch in the superfamly Sylviodea, and in time this genus may warrant recognition at the family level.

Alström *et al.* (2006) recently assigned Kretschmer's longbill (*Macrosphenus kretschmeri)* to the Pycnonotidae. This species had been placed in the Acrocephalinae according to Sibley and Monroe (1990). With respect to *Macrosphenus kretschmeri*, Urban *et al.* (1997) has stated that "it was originally thought to be a bulbul and placed in the monotypic genus *Suaheliornis*," and that "any placement of *Macrosphenus* remains speculative." Beresford *et al*. (2005) found *Macrosphenus flavicans* to belong in the "*Sphenoeacus* group", well removed from *M. kretschmeri*. Hence, further studies are required to evaluate the relationships of the other taxa currently in the genus *Macrosphenus (*Alström *et al.* 2006).

In summary, based on more recent DNA-based studies, some species previously considered bulbuls, e.g., the three species of African nicators (*Nicator*), the Malagasy tetrakas (formerly *Phyllastrephus*, but now *Bernieria* and *Xanthomixis*), the Blackcap Mountain-babbler (*Lioptilus nigricapillus*), Tylas vanga (*Tylas eduardi*) and the Dapple-throat (*Arcanator orostruthus*), have now been placed outside the Pycnonotidae (Schulenberg 1993; Cibois *et al.* 2001; Yamagishi *et al.* 2001; Barker *et al.* 2004; Beresford *et al.* 2005; Johansson *et al.* 2008). Recent molecular studies have suggested that the Pycnonotidae can be divided into an African (greenbul) and a primarily Asian (bulbul) radiation (Pasquet *et al.* 2001; Moyle & Marks 2006; Johanssson *et al.* 2007), with the latter clade also including African members of the genus *Pycnonotus* (*P. barbatus, P. nigricans* and *P. capensis*) that presumably represent a recent colonization from the Middle East (Moyel & Marks 2006). The bulbuls on Madagascar and on the islands of the Indian Ocean (*Hypsipetes*) are part of the Asian radiation (Warren *et al.* 2005) and this region appears to have been colonized from Asia rather than the geographically closer African continent.

Phylogeography: Some species of bulbuls have been intensively studied at the population or among members of a closely related species complex. Several of these studies have served as models systems for investigating patterns and processes of diversification in forest birds of Africa and Madagascar (Roy *et al.* 1997; Smith *et al.* 2005; Warren *et al.* 2005). Roy (1997) studied the phylogeny and phylogeography of the genus *Andropadus.* His analyses revealed that montane species are a recently derived monophyletic group when compared to lowland species of the same genus and that recent speciation events (within the Plio-Pleistocene) have exclusively occurred in montane regions supporting the view that montane regions have acted as centers of speciation during recent climatic instability. However, recent research has revealed that the genera *Andropadus* and *Phyllastrephus* are not monophyletic (Johansson *et al.* 2007), making several of Roy's conclusions questionable.

2.2: A nuclear DNA phylogeny and proposed taxonomic revision of African greenbuls

Using three nuclear markers (myoglobin intron 2, ODC introns 6 and 7 along with intervening exon 7, and β-fibrinogen intron 5) Johansson *et al.* (2007) investigated the phylogenetic relationships within the African clade of bulbuls. The authors included most of the African greenbul species that belong to the three large genera *Andropadus*, *Phyllastrephus* and *Chlorocichla* (approximately 50 spp. excluding three *Pycnonotus* species), as well as taxa belonging to several monotypic genera (e.g., *Caliptocichla*, *Pyrrhurus, Ixonotus etc.*). The resulting Bayesian consensus tree (Figure 2.1) of the combined, mixed-model analysis of the three nuclear markers is generally well-supported and indicates that none of the three largest genera as generally delimited are monophyletic. For instance, the species included in *Andropadus* represent three different clades that are not each other's closest relatives. The montane species currently placed in the genus *Andropadus* form a strongly supported clade, which is sister to *Ixonotus*, *Thescelocichla*, *Baeopogon* and *Chlorocichla*, although within this clade the genus *Chlorocichla* is polyphyletic. The remaining *Andropadus* species fall into two groups, one of these with *A. importunus* and *A. gracilirostris*, which along with *Calyptocichla serina* form a basal branch in the African greenbul radiation. In support of some previous studies, the Leaf-love (*Pyrrhurus scandens*) is placed within *Phyllastrephus* in these analyses. Based on the results Johansson *et al.* (2007) also proposed a new classification that reflects the phylogenetic relationships among African greenbuls. In the proposed classification, the authors included and organized the clades recovered with high support and proposed several genera: *Arizelochichla* (Oberhosler 1905), *Eurillas* (Oberholser 1900), *Stelligidillasa* (Oberholser 1899) and *Atimastillas* (Oberholser 1905), in addition to the currently recognized genera. Phylogeographic inference and the population genetic analysis included throughout my thesis are referenced (i.e. outgroup choice) to the findings the authors made in this study and I have adopted the taxonomic revision presented above and highlighted on Figure 2.1.

2.3: Diversification of African greenbuls in space and time: linking ecological and historical processes

Unraveling the mechanisms underlying the origins of the enormous species richness of tropical rainforests has long been constrained by a deep schism between ecological and historical biogeography (Wiens & Donoghue 2004). Contemporary environmental factors ignoring the species history (phylogeny) have been used to explain the large-scale patterns of variation in species richness (e.g., Brown & Lomolino 1998; Jetz & Rahbek 2002). Other researchers have been more concerned with historical area relationships (Nelson & Platnick 1981; Wiley 1988; Ronquist 1997; Linder 2001). A number of new analytical approaches have been proposed to analyze the interaction between evolutionary and environmental factors in this context (Kerr & Currie 1999; Hawkins *et al.* 2005, 2006; Fjeldså & Rahbek 2006). Following one such approach Fjeldså *et al.* (2007) illustrated the process of diversification of African forest birds since the upper Tertiary by linking together a well-resolved phylogeny and species distribution data for the 65 species of African greenbuls (Pycnonotidae, exculsive of the bulbuls, genus *Pycnonotus*). The geographical pattern of species richness and of branch-length sums resembles the general forest biodiversity pattern, with richness peaks corresponding to postulated Pleistocene forest refuges (Figure 2.2). However, closer examination reveals that this is caused by the disproportionately high contribution of data points representing old and widespread species. Most of the recently evolved species have tiny geographical ranges, mostly in montane areas, and therefore contribute little to the overall picture. It is suggested that speciation may generally take place in topographically well-structured areas, but that species distributions are gradually shifted towards productive lowland areas as a consequence of long-term range dynamics. Old greenbul species are well represented in the coastal forest mosaics of eastern Africa and in the savannah, mainly in tangled thickets, secondary growth and gallery forest, which is similar to the majority of Asian bulbuls, suggesting that this niche was the ancestral habitat for the family. However, the subsequent radiation,

presumably from the early Miocene (Roy *et al.* 1998; Fjeldså *et al.* 2005), was mainly in tropical rainforest.

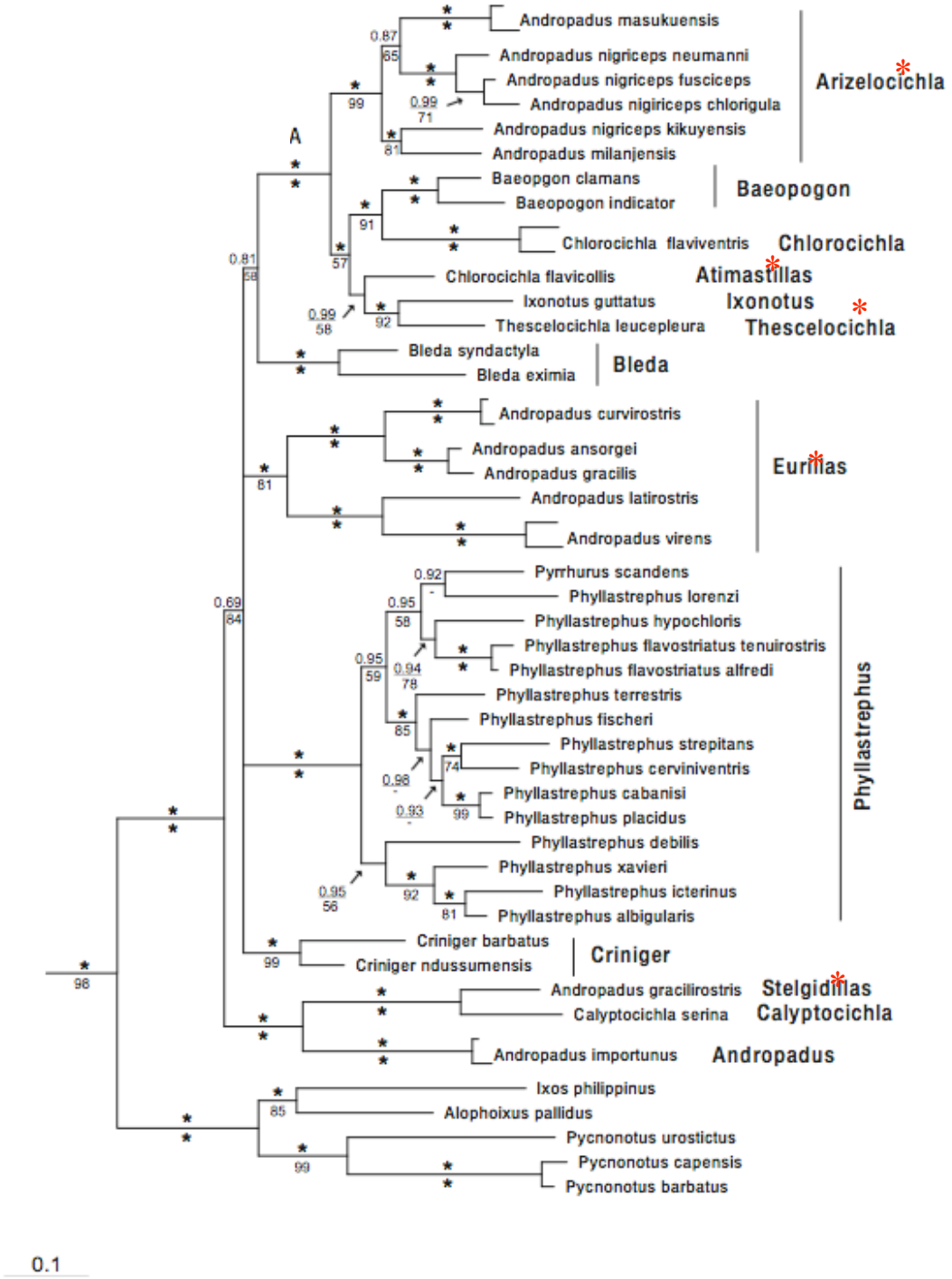


Figure 2.1 - Bayesian consensus tree of the combined, mixed-model analysis of the three nuclear markers myoglobin intron 2, ODC introns 6 and 7, along with exon 7, and β-fibrinogen intron 5 (modified from Johansson *et al.* 2007). Numbers above branches indicate Bayesian posterior probabilities and numbers below bootstrap support under the parsimony criterion. * indicates a posterior probability of 1.0 or 100% bootstrap support. Names in bold to the right of the tree are the revised genera proposed and * denote new taxonomic clades.

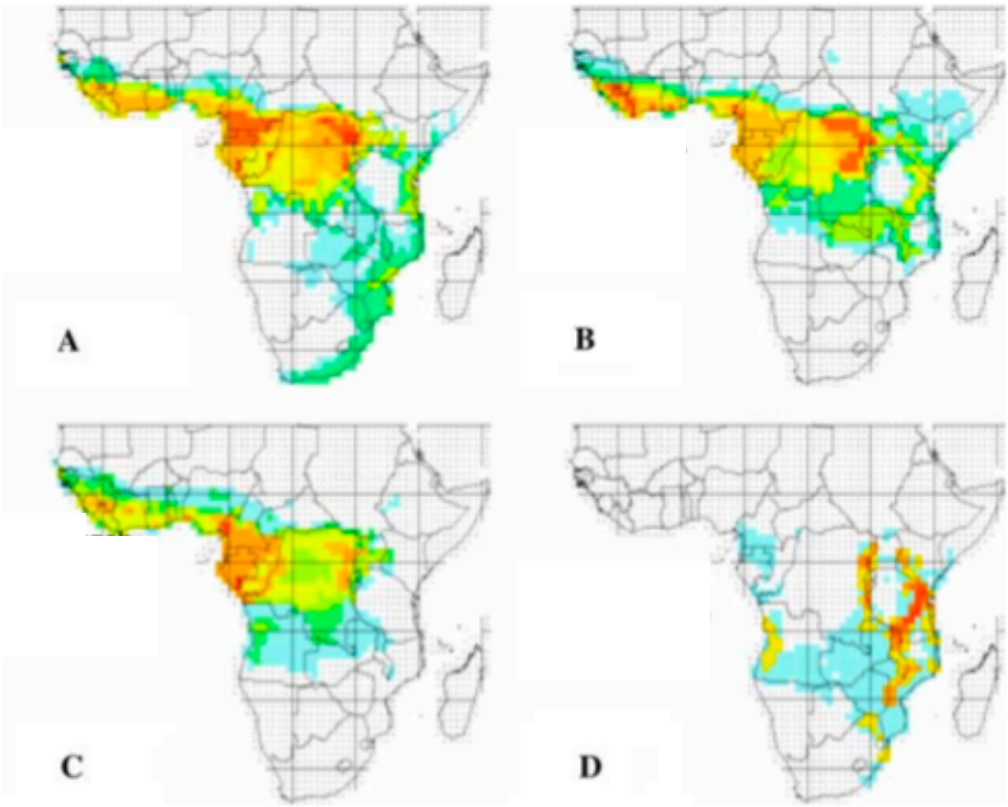


Figure 2.2 - Species richness of greenbuls in terms of branch-length quartiles (modified from Fjeldså *et al.* 2007). A,B,C and D represents first (38.4%), second (30.2%), third (28%) and fourth (3.4%) branch length quartiles respectively. First quartile represents the shortest branches (oldest lineages) in the phylogeny and fourth quartile represent longest branches (youngest lineages). Red and orange indicate high values, light blue the lowest values, white absence of species.

CHAPTER 3

Phylogeography of the Grey-olive greenbul (*Phyllastrephus cerviniventris*): Molecular evidence for paleo-drainage dynamics shaping population genetic structure

3.1: Introduction

Global climate change that resulted in high latitude glaciations both restricted and promoted divergence in African terrestrial biomes (deMenocal 1995, 2004). Studies using pollen and sedimentation data (Livingstone 1993; Partridge *et al*. 1995; deMenocal 1995, 2004; Maley, 1996, 2001; Gasse *et al.* 2008) together with eolion dust from deep-sea drill cores (deMenocal 1995) have documented how aridity increased during the Pleistocene. In sub-Saharan Africa, savanna and semi-arid biomes expanded while montane and coastal forests contracted (Bowie *et al.* 2004a, 2006; Fjeldså & Bowie 2008 and refs therein). Yet, how the semi-evergreen habitats in places with high groundwater level around streams and along the margins of swampy plains (watershed basins) responded to climatic perturbations is at present poorly understood.

The Zambezi and Rufiji river basins (Figure 3.1) are believed to represent remnants of larger paleo-drainage system in central, southern and eastern Africa (Stankiewicz & de Wit 2006). These two river basins support and sustain much of the semi-evergreen habitats and the unique fauna associated with them (Cotterill 2004, 2005 & 2006). As water is essential for sustaining of life, the evolving geometry of river basins often have influence on local speciation. Thus the evolution of these basins had a major influence on the distribution of riverine plant and animal species (e.g. Linder 2001 for plants; Cotterill 2003, 2005 for mammals and fish; Daniels *et al.* 2006 for freshwater crabs; Dijkstra 2007 for dragonflies). Many of the patterns seen today in the distribution of species within the unique habitat configurations of south-central Africa are the result of historical processes that have occurred over the last few million years (see Timberlake 1998). Various vicariant events especially the development of East African rift system (EARS) about 5 Ma had a profound effect on the landscape ecology and the species distribution (Partridge *et al*. 1997; Lovett 1993; Coppes 1994; Axelrod & Raven 1978) in sub-Saharan Africa. With the formation of EARS, marked rearrangement of paleorivers took place by connecting and sundering which lead to formation of new drainage basins and lakes. In addition to altering the drainage topology, The EARS effectively

separated the faunal assemblages in to several isolated populations; thus, altering the extent and connectivity of habitats. Cotterill (2006), suggested that an ancient drainage system, with connectivity from Zambia to the swampy plains below the Udzungwa/Rubeho mountains of Tanzania was interrupted by the Rukwa-Malawi rifting. The Rukwa-Malwi rift may have effectively divided lineages that previously were distributed widely in the region. In addition to formation of the EARS, Pleistocene climatic oscillations also have drastically challenged the connectivity of the habitats among such isolated populations. Thus, river captures have not only been important in facilitating species dispersions but, particularly in the Plio-Pleistocene, have also led to the disruption of formerly continuous populations, providing a driving force for speciation (Moore *et al.* 2007).

Understanding how species responded to such long-term fluctuation in the extent and connectivity of habitat is key to reveal the evolutionary history of tropical forest fauna (Dolman & Moritz 2006). Alteration of landscape ecology and isolation of populations often leads to particular faunal adaptation or selection pressure and/or genetic drift resulting in allopatric speciation. To gain insight in to how Pleistocene climatic perturbations and Pliocene tectonic events associated with the formation of the EARS may have shaped the spatial distribution of a species adapted to seeps and semi-evergreen habitats in eastern and central Africa, I focus on the Grey-olive greenbul (*Phyllasterphus cerviniventris*) that is distributed across the Malawi-Rukwa rift basins. The habitat specificity of this species also allowed me to test how palaeodrainage systems particularly the Zambezi and Rufiji river basins shaped the population genetic structure of this taxon.

Grey-olive greenbul: Phyllastrephus cerviniventris (Shelly 1894)

The Grey-olive greenbul is relatively small pycnonotid (18-21cm, 30g) and inhabits a wide array of habitats with a preference for riverine vegetation (gallery forest, gullies in open woodland, lake littoral, ground water forest, bamboo thickest etc.) in western, eastern and southern Africa (Keith *et al.* 1992). The species typically inhabits broadleaf woodland forest (400m-1900m), skulking close to the forest floor (Dowsett-Lemaire 1989). Its present distributional range encompasses scattered localities in southern Kenya (Meru, Tharaka, Thika, Lilterish, Mzima Spring, Kitouv Forest and Bura), northern and eastern Tanzania (Lake Manyara, around the base of Mt. Meru and Mt. Kilimanjaro, west and east Usambara Mountains and the Uluguru Mountains extending to Luwegu River and Songea), Zambia (relatively widespread in Miombo, but local, absent only from the north between Lakes Tanganyika and Mweru and areas close to the Tanzanian border, as well as from Barotse and most of the southern provinces; Dowsett *et al.* 2009), and is present in the extreme southeastern corner of the DRC, and widespread in Malawi (Dowsett-Lemaire & Dowsett 2005) and northern Mozambique (Figure 3.2). It feeds mainly on insects including grasshopper and beetles, and rarely on fruits.

As the name suggests the Grey-olive greenbul has grey-olive tinged brownish upper-parts, a white or very pale buffy chin and throat, grayish olive-brown breast and flank, with a tawny wash on the lower flank, a few buffy streaks on the center of the breast, creamy buff on the belly with tawny under tail coverts, a pale underside to the bill and whitish or pinkish legs. The species has some morphological resemblance to larger sympatric congeners: Northern brownbul (*P. stripitans*), Terrestrial brownbul (*P. terrestris*) and Cabanis's greenbul (*P. cabanisi*). Pale to purple eye color and whitish or pinkish legs are the main characters by which the species can be distinguished from the other sympatric congeners in the field.

Based on plumage variation, Prigogine (1969) described a subspecies *Phyllastrephus cerviniventris schoutedeni* from four specimens collected in the Kisanga River region of the DRC (27^0 35'E, 11^0 43'S). This taxon has a darker and greyer throat and breast, darker rufous under tail coverts and belly, and dark brown tail feathers with less rufous than the nominate race. However, an intermediate morphological form between these two forms does exist in Kakanda near the Lualaba River (Prigogrine 1969). As a consequence Keith *et al.* (1992) consider "*schoutedeni*" to represent part of a cline (Keith *et al.* 1992) and do not recognize it as a distinct subspecies.

3.2: Materials and Methods

Population sampling

A total of 46 samples covering the distribution range of the Grey-olive greenbul were included in the study (Figure 3.2; Table 3.1). In addition to the fresh samples (either blood or tissue), I also included eight toe-pad samples from museums skins collected between 1935-1970 to augment our field sampling. One of the toe-pad samples included was from the presumed subspecies *Phyllasterphus cerviniventris schoutedeni* (Prigogine 1969). A closely related greenbul species, *Phyllastrephus flavostriatus* was used as the outgroup taxon (Johansson *et al.* 2007).

Laboratory procedures

Genomic DNA was isolated using either a Puregene DNA isolation kit (Gentra) or a DNeasy Tissue Kit (Qiagen) following the manufacturer's protocols. In the case of museum skins, protein digestion time was expanded to 24h and twice the volume of tissue lysis buffer and Protinase K was added (DNeasy Tissue Kit) along with 10mM DTT (dithiothreitol) solution to facilitate tissue digestion. Extraction and amplification of museum skin samples were carried out in a dedicated biosafety level two laboratory at the Museum of Vertebrate Zoology, University of California Berkeley USA, following stringent precautions to avoid contamination.

I amplified and sequenced the complete mitochondrial ND2 gene (1041 bp), most of ATPase

6 (ca. 690 bp) including the flanking tRNA, as well as two nuclear introns: Ornithine Decarboxylase(ODC) and *ß*- fibrinogen (Fib5). The portion of ODC sequenced (700 bp) included introns 6 and 7, along with the intervening exon 7, as well as portions of the flanking exons. The sequenced portion of β-fibrinogen (558 bp) included the complete intron 5 and small pieces of flanking exons. PCR-amplification conditions follow the protocols in Fuchs *et al.* (2004) for ND2 and FIB5, and Allen and Omland (2003) for ODC. PCR protocols were slightly modified (e.g. by the use of Amplitaq Gold (Roche) DNA polymerase and Bovine Serum Albumin) for the amplification and sequencing of DNA from toepads. In addition, I developed several internal primers (Table 3.2) to amplify and sequence DNA isolated from museum specimens as DNA was highly fragmented in these isolations.

Following ethidium-bromide visualization on a 1% agarose gel, PCR products were cleaned by using either an Agarase enzyme (GELase™, EPICENTRE® Biotechnologies) or a combination of Shrimp Alkaline Phosphatase and Exo-nuclease I enzymes (ExoSap-IT, USB Corporation). Sequencing reactions were performed using BigDye (Applied Biosystems) terminator chemistry, and run on an ABI 3730 (Applied Biosystems) DNA sequencer. Sequences were obtained from both strands of DNA for each individual, and some individuals were sequenced several times if any base ambiguity was encountered. Length polymorphism was detected in some individuals for Fib5 and ODC and these PCR products were cloned into pGEM-T-Easy-Vector (Promega) and transformed into JM109 Competent Cells following the manufacturers recommendations (pGEM®-T Vector Systems, Promega). Eight positive colonies were selected from each plate and used as a template for PCR-amplification to check the size of the insert. Colonies with inserts of the anticipated size were then purified using the standard Miniprep protocol (Sambrook *et al*. 1989) and sequenced in both directions using the M13 forward and reverse primers or with the internal sequencing primers developed specifically for this study (Table 3.2).

Sequence alignment

All mitochondrial DNA (mtDNA) sequences were checked using the program Sequencher 4.7 (Gene Codes Corp) to test for the presence of any insertions or deletions, as well as to check that no stop codons were present. Nuclear intron sequences were edited and assembled using Sequencher 4.7 (Gene Codes Corp). For the two intron datasets allelic phase was determined using PHASE 2.1 (Stephens *et al.* 2001; Stephens & Donnelly 2003). I performed two independent phase estimates using the recombination model as implemented in PHASE 2.1 with different starting seeds to check for consistency across runs.

Phylogenetic reconstruction and networks

Phylogenetic trees were estimated using the model-based methods of maximum likelihood (ML) and Bayesian inference (BI). Maximum likelihood (ML) analysis was performed on the combined mtDNA dataset as implemented in GARLI v.95 (Zwickl 2006). Clade support for the ML analysis was estimated using 100 nonparametric bootstrap replicates (Felsenstein 1985). MrBayes 3.1 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003) was used to conduct a Bayesian approach to phylogenetic inference. The GTR + G model was selected as the best fit model to the combined mtDNA data set with the assistance of MrModeltest (Nylander 2004) and the Akaike information criterion (Akaike 1973). Four Metropolis-coupled Markov chains (one cold and three heated) were run for five million iterations, with trees sampled every 100 iterations. The first 250,000 iterations (2500 trees) were discarded ('burn-in' period), and the posterior probabilities were estimated from the remaining sampled iterations. Two independent Bayesian runs initiated from random starting trees were performed, and the log-likelihood values, posterior probabilities and average deviation of split frequencies were checked to ascertain whether the chains had reached convergence. I also used Tracer v1.4 (Rambaut & Drummond 2007) to check that stationarity had been reached by graphically monitoring the fluctuating value of the likelihood scores.

Genealogical patterns for the combined mtDNA dataset and the phased intron data sets were visualized as a statistical parsimony network using the program TCSv1.18 (Clement *et al*. 2000). Gaps in the phased intron data were treated as fifth state when analyzing data to construct the parsimony network. A statistical parsimony network provided a better graphical representation of genealogical pattern for the phased intron data sets compared to a neighbor-joining split network (Huson & Bryant 2006), as this data consisted of only of few haplotypes with most of them being shared between the mtDNA clades.

Population structure and historical demography

Analysis of molecular variance (AMOVA) using Tamura and Nei's pairwise distances (1000 permutations) (Excoffier *et al.* 1992) as implemented in Arlequin 3.1 (Schneider *et al.* 2000) was used to determine how genetic variability was partitioned within and between the populations east and west of the Rukwa-Malawi rift for all three data sets: combined mtDNA, ODC and FIB5. I used Arlequin (Schneider *et al.* 2000) to calculate the number of haplotypes, haplotype diversity (H), nucleotide diversity (p), Tajima's D and Fu's Fs tests of selective neutrality (1000 replicates), as well as to estimate mismatch distributions for the populations east and west of the Rukwa-Malawi rift. I also used a Mantel test (Mantel 1967) as implemented in Arlequin 3.1 (Schneider *et al.* 2000) to test for isolation-by-distance for both populations on either side of Rukwa-Malawi rift basin.

Estimation of divergence times

I used Beast v1.4.6 (Drummond *et al.* 2002, 2006; Drummond & Rambaut 2007), to estimate divergence dates within *P. cerviniventris* using mtDNA sequence data. I assigned the best-fitting model, as estimated by MrModeltest2 to each of the partitions. I used two alternate mutation rates, the classical rate of $1.05*10^{-8}$ substitutions/site/year (s/s/y) for mtDNA (2.1% per million years, Weir & Schluter 2008), as well as the faster rate of $3.05*10^{-8}$ (6.1% per million years, Arbogast *et al.* 2006, see also Warren *et al.* 2003), which translate into per locus (mtDNA: ND2 & ATP6 combined) rates of $1.82*10^{-5}$ and $5.29*10^{-5}$ substitutions per year, respectively. This range of rates carries with

it a significant margin of error and thus I emphasize the importance of thinking of the time estimates only as a rough attempt to place diversification events of *P. cerviniventris* in a historical context. I assumed a Yule Speciation Process for the tree prior and an uncorrelated lognormal distribution for the molecular clock model (Drummond *et al.* 2006; Ho 2007). I used default prior distributions for all other parameters and ran MCMC chains for 50 million generations. The analysis was repeated twice. I used the program Tracer (Rambaut & Drummond 2007) to assess convergence diagnostics.

Gene flow versus ancestral polymorphism

I used the Markov chain Monte Carlo method implemented in the program IMa (Hey & Nielsen 2007) to fit the data to a model that included both isolation and migration. IMa estimates six parameters scaled to the neutral mutation rate (μ): θ_{pop1} ($4Ne_{pop1}\mu$), θ_{pop2} ($4Ne_{pop2}\mu$), θ_{popA} ($4Ne_{popA}\mu$), t (T/μ, where T is the time since population divergence in years before present), m1 ($2M/\theta_{pop1}$, where M is the effective number of migrants moving into population 1) and m2 ($2M/\theta_{pop2}$, where M is the effective number of migrants moving into population 2). I defined inheritance scales to reflect the difference in modes of inheritance among the loci used: 0.25 for the mtDNA locus and 1.0 for the two autosomal loci. I used an HKY model of nucleotide substitution for all three data partitions. Parameters and genealogies were sampled every 100 steps for both the 10 and 20 million step runs. To assess convergence I monitored the extent of autocorrelation, parameter trend lines and the effective sample size (ESS) throughout the run and I also compared the results between three independent runs. I used the Difference of Sums of Squares (DSS) test (McGuire & Wright 2000) as implemented in the program TOPALI v 2.5 (McGuire & Wright 2000) to test for evidence of recombination in both intron datasets. Based on the result of the DSS test, I carried out a Hudson-Kreitman-Aguade (HKA) test (Hudson *et al.* 1987) to test for neutrality across the two nuclear loci with 100,000 simulations using the program HKA (Jody Hey, Rutgers University).

I used a generation time of 1.7 years, which reflects the average for several passerine species (Sæther *et al.* 2005). I used the same two alternate mutation rates, described previously for the dating for the mtDNA partitions. For the autosomal loci I selected a rate of $3.61*10^{-9}$ s/s/y (Axelsson *et al.* 2006). This translated into per locus rates of: FGB $2.04*10^{-6}$ s/l/y, and ODC $2.44*10^{-6}$ s/l/y. The geometric mean of the combined mtDNA and ncDNA loci was $4.49*10^{-6}$ s/l/y (mtDNA 2.1%/ Ma) or $6.41*10^{-6}$ s/l/y (6.1%/ Ma), and for the two nuclear loci $2.23*10^{-6}$ s/l/y.

3.3: Results

mtDNA population structure and genetic diversity

The mitochondrial DNA matrix contained 30 ingroup haplotypes from the 46 individuals sampled and consisted of 1734 bp (1041 ND2, and 693 ATP6) of which 32 (1.85%) sites were parsimony informative. The 50% majority-rule consensus tree (-ln 3577.5716) resulting from the Bayesian analysis was largely unresolved; only a few nodes received posterior probabilities equal to or greater than 0.90 (Figure 3.3). A similar pattern resulted with the maximum likelihood analysis. Uncorrected-p distances varied between 0% and 2.1%. The number of haplotypes, gene and nucleotide diversities, and the results of the tests of selective neutrality for the two populations across Malawi and Rukwa rift and correspond to the two watershed basins Zambezi and Rufiji populations (Table 3.3). The Mantel test detected a correlation between genetic distance and geographic distance among the samples from the Tanzania east of Rukwa-Malawi rift basin whereas there was no correlation for samples west of the rift basin.

A statistical parsimony network of the combined mtDNA data set revealed two main haplotype clusters that are differentiated from each other by at least nine substitutions (Figure 3.4). The haplotype clusters have a biogeographical component revealing differentiation across the Rukwa-Malawi rift. The haplotype cluster west of the Rukwa-Malawi rift (hereafter referred to as Zambezi population) consists of all samples from Malawi, Zambia, Mozambique and the DRC,

whereas the haplotype cluster east of Rukwa-Malawi rift (Rufiji population) consists exclusively of the samples from Tanzania. As a whole, individuals sampled at the same locality or geographically proximate localities, have identical or nearly identical haplotypes. No haplotypes were shared between Zambezi and Rufiji populations (Fig. 3.4B). The subspecies. *P. cerviniventris schoutedeni* from the DRC grouped with a haplotype from Zambia within the Zambezi population.

Zambezi clade has significantly negative Tajima's D and Fu's Fs statistics whereas Rufiji clades does not (Table 3.3). Mismatch profile follow a Poisson distribution (Figure 3.5), suggesting that these populations may have undergone recent population growth following range expansion. In the AMOVA, significant structure was detected among populations (Fst = 0.71, P < 0.001) compared to within population genetic diversity 28.27%. This pattern corroborates the results from the statistical parsimony network (Figure 3.4b) and the strong geographic structuring across the Rukwa-Malawi Rift.

Nuclear gene population structure and diversity

The intron alignments, after pruning both 3' and 5' ends to remove missing data, consisted of 566 bp and 679 bp for FIB5 and ODC respectively. Merging identical sequences in the FIB5 and ODC alignments (after estimating haplotypic phase) yielded eight and nine unique haplotypes respectively. Both loci revealed little genetic structure relative to the mtDNA (Table 3.3). Further, the Rufiji and Zambesi clades were not recovered in the ML analysis of either intron data set, or when the two were combined. Instead, an unresolved comb was recovered.

When partitioning the nuclear alleles for the two loci into the Rufiji and Zambezi mtDNA clades, AMOVA revealed that for both nuclear loci almost all the genetic variability was found to be within populations, rather than among populations as in the mtDNA. For FIB5, 11% (Fst = 0.11, d.f. = 91, *P* < 0.02) of the total genetic variability was attributed to among population, and 89% was attributed to within population variation. For ODC, no proportion (Fst = -.008, d.f. = 90, *P* = 0.63) of

the genetic variability was partitioned among populations distributed Rufiji versus Zambezi basins across Rukwa-Malawi rift (100% within population).

Neither the statistical parsimony network constructed from the allelic variation for FIB5 (Figure 3.6) nor ODC (Figure 3.7) recovered the split of lineages on either side of the Rukwa-Malawi Rift as suggested by the mtDNA. Populations on either side of the rift shared several alleles In FIB5, three out of five alleles from Zambezi population are shared with Rufiji populations east of the rift. Two private alleles were restricted to the west and three alleles were restricted to populations east of the rift. The statistical parsimony network for FIB5 did not show a clear signature of recent population expansion with alleles not forming the typical star like pattern. In the ODC network, the Zambezi clade had seven unique haplotypes and the Rufiji population had one unique haplotype, with the central and most common haplotype shared by both mtDNA clades. Thus, the network recovered for ODC was more representative of the expected star like pattern indicative of recent population expansion. Mismatch distributions for both nuclear introns show similar trends by following the expected poison distribution under a model of population expansion (Figures 3.6 and 3.7); both raggedness indices were insignificant.

Dating analyses

Divergence date estimates obtained using a molecular clock hypothesis for the two (6.1% and 2.1%) clock rates implemented in BEAST suggest that the lineages east and west of the Rukwa basin diverged between 0.218 Mya (95% HPD: 0.1361-0.3062 Mya) and 0.667 Mya (95% HPD: 0.3163-1.1052 Mya) depending on which clock rate was used (Table 3.4). The 95% HPD interval does not overlap between these two alternate rates of divergence. Under both divergence rate estimates the TMRCA of the Zambezi clade was slightly older than the northern clade (Table 3.4).

Evidence for a model of isolation-with-migration

The DSS algorithms implemented in Topali v 2.5 (McGuire & Wright 2000) did not detect any evidence of recombination within the two nuclear loci. Hence I used the complete sequences for the IMa analyses. Appropriate mixing and satisfactory effective sampled sizes were achieved using 12 chains and a geometric heating scheme (g1=0.15 and g2=0.70). The sharp peak of the mode of the marginal posterior distribution suggested that the population migration rates between the Zambezi population and Rufiji population were low although the confidence interval was large (effective rate at which genes come into the Zambezi population 0.01 (90%HPD_{adj}: 0.01-14.82), and into the Rufiji population 0.01 (0.00-33.45)). Log-likelihood ratio tests suggested that neither a model of symmetrical nor a model of zero recurrent gene flow could be rejected (symmetrical 2LLR = 0.0154, df =1, $P < 0.9$; zero 2LLR = 0.0167, df = 2, $P < 0.99$). Similarly, although the respective modes of the marginal distribution of theta (effective population size scaled by mutation rate) differed among the two extant populations (south 5.559, north 3.54) as well as relative to the ancestral population size (1.736), likelihood-ratio tests suggested that equal effective population sizes could not be rejected (2LLR = 5.72. df = 2, $P < 0.1$). However, any model that assumed equal population size and migration rates could be rejected. (2LLR = 10.04, df = 3, $P < 0.025$). Further inspection suggests that this may be due to a larger population size in the Zambezi basin, which would be consistent with the greater availability of seepy and swamp areas.

3.4: Discussion

Using a combination of mtDNA and two nuclear loci the objective of the present study was to assess for the first time, how palaeodrainage systems particularly the Zambezi and Rufiji River basins have shaped the population genetic structure of bird species and the importance of these overlooked habitat configurations as centers of speciation. Further, any influence of the Rukwa-Malawi rift on the diversification of a riverine habitat specialist such as Grey-olive greenbul were also tested. Below, I specifically attempt to place the molecular results in the context of known Pleistocene climatic perturbations in the region and derive new insight into how watershed basin and specifically taxa restricted to riverine semi-evergreen forest may have responded to high amplitude glacial cycles.

Genetic differentiation across the Rukwa-Malawi Rift

Our analysis of mtDNA revealed that the grey-olive greenbul is geographically structured across the Rukwa-Malawi Rift, in the two watershed basin areas with the detection of significant genotypic diversity among populations in the AMOVA (d.f. = 45, Fst = 0.71, $P < 0.001$) and fairly low within population genetic variation (28.27%) corroborating the structure depicted within the statistical parsimony network of the combined mtDNA dataset (Figure 3.4). In contrast, AMOVA revealed very limited structure among populations for FIB5 and none for ODC, with several alleles being shared between the east and western mtDNA lineages. However, results from the coalescent analyses suggest that a model of zero recurrent gene flow cannot be reject and this together with the mtDNA results suggest that the lack of sorting among the nuclear alleles is most likely a consequence of the retention of ancestral polymorphism and insufficient time for lineage sorting to have been completed, as would be expected given the limited (ca. 2%) and recent (most likely with past 0.35 Ma, Table 3.4) divergence estimates (see also Zink & Barrowclough 2008).

Historical demography

The genetic structure and genetic diversity observed in the Rufiji clade, as inferred from the mean genetic distance among haplotypes and nucleotide diversity (Table 3.3) can be explained by isolation-by-distance. For instance, some haplotypes are found more than 300 km from each other in isolated patches of riverine vegetation within the interior of Tanzania. This is in agreement with correlation found among the sample of Rufiji clade in Mantel test. There is evidence in this data set for a signature of population expansion within the Zambezi clade, with individuals sampled from localities in Malawi primarily having haplotypes shared or derived from Zambian populations.

Biogeography

My estimates of divergence times suggest that the diversification of the Rufiji and Zambezi Grey-olive greenbul populations across the Rukwa-Malawi rift is relatively recently, and certainly occurred within the Pleistocene. This is in general concordance with the Pleistocene estimates of divergence of several African savanna/woodland ungulates (Lorenzen *et al.* 2006a,b, 2008, 2010), but is younger than observed for most divergences among allopatric montane bird populations (e.g. Bowie *et al.* 2004, 2006; Fjeldså & Bowie 2008; Voelker *et al.* 2010); no comparable data exists for other watershed basin and broadleaf woodland birds. These all suggest that the region comprises permanent refuges, most probably forest associated with seeps around the large permanent swamps of the Zambian Plateau and the Kilombero Valley, although these habitats are only likely to have been connected during the most humid periods. Further, these permanent swamps experienced cycles of expansion and contraction throughout the Pleistocene (Hamilton 1976, 1982; Thackeray & Reynolds 1997). Episodic cold or cool intervals, occurring at intervals of 100 000 years or less would have led to repeated episodes of fragmentation of riverine woodland allowing for episodic gene flow to occur between various taxa. Expansion of watershed basins would have facilitated the expansion and range dynamics of the watershed-adapted species such as Grey-olive

greenbul. Yet it is important to recognize that boundaries between taxa may be blurred on account of episodic expansion and contraction of habitats, associated with repeated episodes of gene flow between populations of taxa that were only temporally isolated by habitat change in miombo woodlands (Lorenzen *et al.* 2010).

On the other hand the increased tectonic activities around lakes Rukwa and Malawi during Pleistocene have been documented as evident from the radiometric dating on lavas, detailed stratigraphic field evidence combined with whole-rock major and trace element analyses of tephra samples in the Rukwa-Malawi rift basin specially in Rugwe Volcanic Province (RVP)(Harkin 1960; Ebinger *et al.* 1989, 1993; Delvaux *et al.* 1992; Ivanov *et al.* 1999; Fontijn *et al.* 2010). Though an initial episode of RVP volcanic activity took place as early as the Early Miocene, between 19 and 17 Ma (Ivanov *et al.* 1999; Rasskazov *et al.* 2003), the main phase of RVP volcanism started ca. 9 Ma ago and has been divided into three stages based on radiometric dating on lavas (Ebinger *et al.* 1989, 1993; Delvaux *et al.* 1992; Ivanov *et al.* 1999): (1) LateMiocene: ~8.6 to 5.4 Ma, (2) Late Pliocene – Early Pleistocene: ~3 to 1.6 Ma, (Harkin 1960; Fontijn *et al.* 2010), and (3) Mid-Pleistocene – Recent: ~0.6 Ma to present (Harkin 1960; Fontijn *et al.* 2010). Further, Delvaux (1995) documented high lake level of lake Malawi from 0.45 Ma to 0.25 Ma based on high-resolution seismic stratigraphy. This allows for consideration of both climatic and tectonic events such as intense volcanic activities, associated with the mid Pleistocene to early Holocene in the region (Delvaux *et al.* 1998).

Therefore the observed genetic differentiation of the Grey-olive greenbul across Rukwa–Malawi basin can be explained by both high amplitude climatic perturbations in mid Pleistocene to early Holocene and the associated tectonic activities of the region specially intense volcanic activities during the mid Pleistocene to Holocene. Perhaps climatically induced lake level fluctuations too would have contributed for the isolation of the lineages on either side of Rukwa-Malawi rift basin. The timing of intense volcanic activities as well as climatically induced high lake

level correlate with the estimated divergence time of this study. Combined effect of both probably had lead to the complete reproductive isolation of the Rufiji and Zambezi clades as revealed by my analysis (no shared mtDNA hplotypes). Within Zambezi and Rufiji basin regions periodical isolation of populations and subsequent range overlap, most likely triggered by climatic changes, might have led to a genetic differentiation and diversification of the Grey-olive greenbul observed today.

Taxonomic status of Phyllasterphus cerviniventris schoutedenti (Prigogine 1969)

I managed to obtain one sample of this taxon (a paratype) from a museum specimen in the collection of the Royal Museum for Central Africa (Tervuren, Belgium). This sample had an identical mtDNA haplotype to birds from nearby Zambia and Malawi and given the apparent slight plumage differences (Urban *et al.* 1992) I do not consider *Phyllasterphus cerviniventris schoutedenti* to be valid, and hence no subspecies can be recognized for the Grey-olive greenbul.

Table 3.1 - Sample origins and the number of individuals included in the study

Country of origin	*Number of samples used*	*Extraction numbers**
Tanzania	08	LGS140 to LGS144, LGS246 to LGS248
Malawi	25	LGS030-LGS031, LGS033 to LGS038, LGS042 to LGS044, LGS047 to LGS049, LGS056, LGS407 to LGS411, LGS902 to LGS906
Zambia	10	LGS040-LGS041, LGS755 - LGS756, LGS758 to LGS763,
Mozambique	02	LGS045-LGS046
DRC	01	LGS673

* Detailed information (museum catalog number, source, country of origin, locality etc.) of the samples can be found in the Appendix 3.1).

Table 3.2 - Novel PCR primers developed for this study to amplify three gene regions from museum skin samples.

Primer name	*Gene region*	*Primer location within the gene*	*Primer sequence -5'-3'*
1. ND2PcvR1	NADH subunit 2	368-388 bp	GAG TAG TCC GGT RGT GAG G
2. ND2PcvR2	NADH subunit 2	701-718 bp	TGA CGG TGC TTT CGC TC
3. ND2PcvF2	NADH subunit 2	341-358 bp	GAT TCC CYG AAG TCC TGC
4. ND2PcvF3	NADH subunit 2	665-685 bp	CGA GRC CCC RAA ACT ATC C
5. ATP6PcvF	ATP6	340-361 bp	TCC CTA RGC CAC CTY TTA CC
6. ATP6PcvR	ATP6	380-399 bp	GGA TGA GTG CTG GGA TTA GG
7. ODCPcvR1	ODC	290-308 bp	ACA CAC AGC GGG CAT CAG
8. ODCPcvF1	ODC	384-410 bp	GAA GTG TGT ACA ACT GTT ATG GGGTT
9. ODCPcvR2	ODC	490-527 bp	GTA ACA GTC ATT TGA GTT TGA GCT GCC
10.ODCPcvF2	ODC	240-264 bp	AGT GGA TGT ACA GAC CCA GAG ACC
11.ODCPcvF3	ODC	390-414 bp	AAG CTA GCT AAG TCA GTG ACT GCA

Table 3.3 - The number of haplotypes, gene and nucleotide diversities, and the results of the tests of selective neutrality for the two populations of *Phyllastrephus cerviniventris* distributed across Rukwa-Malawi Rift for mitochondrial DNA and two nuclear intron loci.

Gene	*Population*	*Number of haplotypes*	*Number of polymorphic sites*	*Haplotype diversity (H)*	*Nucleotide diversity (π)*	*Tajima's D test*	*Fu's Fs test*
mtDNA	Rufiji (n=8)	7	15	0.9643 +/- 0.0017	0.0028 +/- 0.0017	-0.7575 (P>0.25)	-1.9422 (P<0.1)
	Zambezi (n=38)	23	38	0.9543 +/- 0.0011	0.0018 +/- 0.0011	-2.1287 (P<0.01)	-16.8577 (P=0.00)
FIB5	Rufiji (n=16)	6	4	0.6833 +/- 0.1200	0.0013 +/- 0.0012	-0.4673 (P>0.7)	-2.7031 (P<0.01)
	Zambezi (n=76)	7	5	0.7607 +/- 0.0186	0.0017 +/- 0.0013	0.4400 (P<0.4)	-0.3970 (P>0.45)
ODC	Rufiji (n=16)	2	1	0.1333 +/- 0.1123	0.0002 +/- 0.0003	-1.1594 (P>0.15)	-0.6489 (P=0.1)
	Zambezi (n=76)	8	9	0.3330 +/- 0.0698	0.0004 +/- 0.0004	-1.8525 (P<0.005)	-4.9504 (P<0.01)

Table 3.4 - Estimated time-to-most-recent-common-ancestry (TMRCA) of *P. cerviniventris* and the two mtDNA lineages under two alternate clock rates. Upper and lower 95% confidence levels for estimates are presented in parenthesis.

Taxon group	*TMRCA estimated at 2% clock rate*	*TMRCA estimated at 6.1% clock rate*
P. cerviniventris	0.667 (0.3163-1.1052)	0.218 (0.1361-0.3062)
Zambezi clade	0.451 (0.2199-0.7540)	0.146 (0.0943-0.2069)
Rufiji clade	0.396 (0.1458-0.6821)	0.129 (0.0636-0.1965)

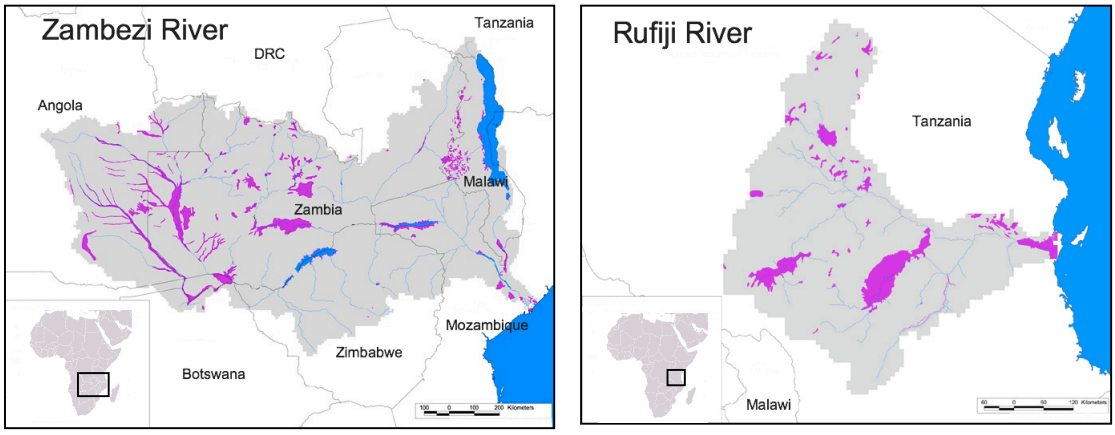


Figure 3.1 - The Zambezi and Rufiji river systems showing respective drainage basins (areas highlighted in grey) and associated major wetlands/swampy areas (highlighted in purple). Map modified from the water resources of the world eAtlas (2003).

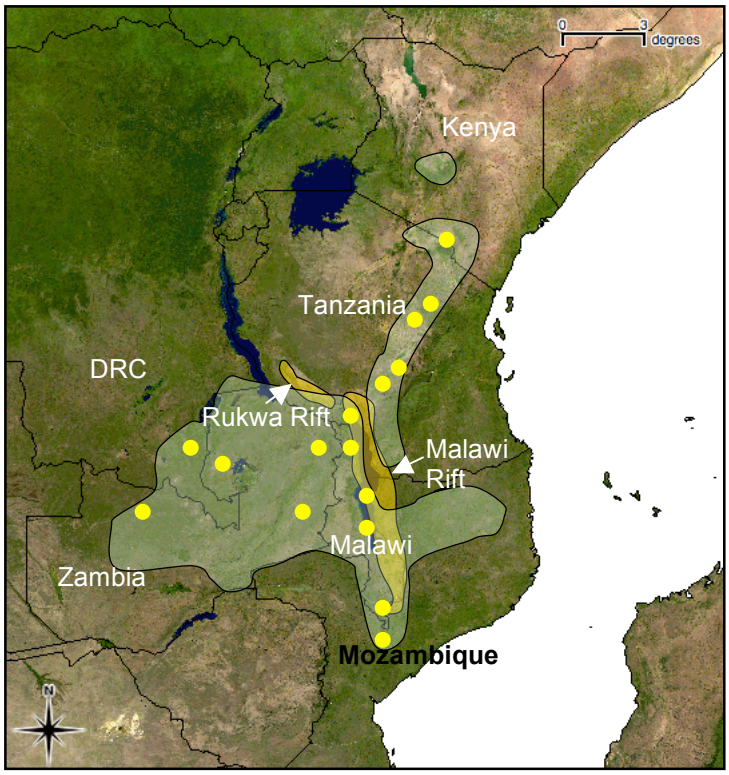


Figure 3.2 - Geographical distribution of the grey-olive greenbul *Phyllatrephus cerviniventris*. Areas in light green depict the current distribution range of the species and the dots in yellow highlight the sampling localities for genetic analyses. Areas highlighted in light orange depicts the location of the Rukwa and Malawi rifts.

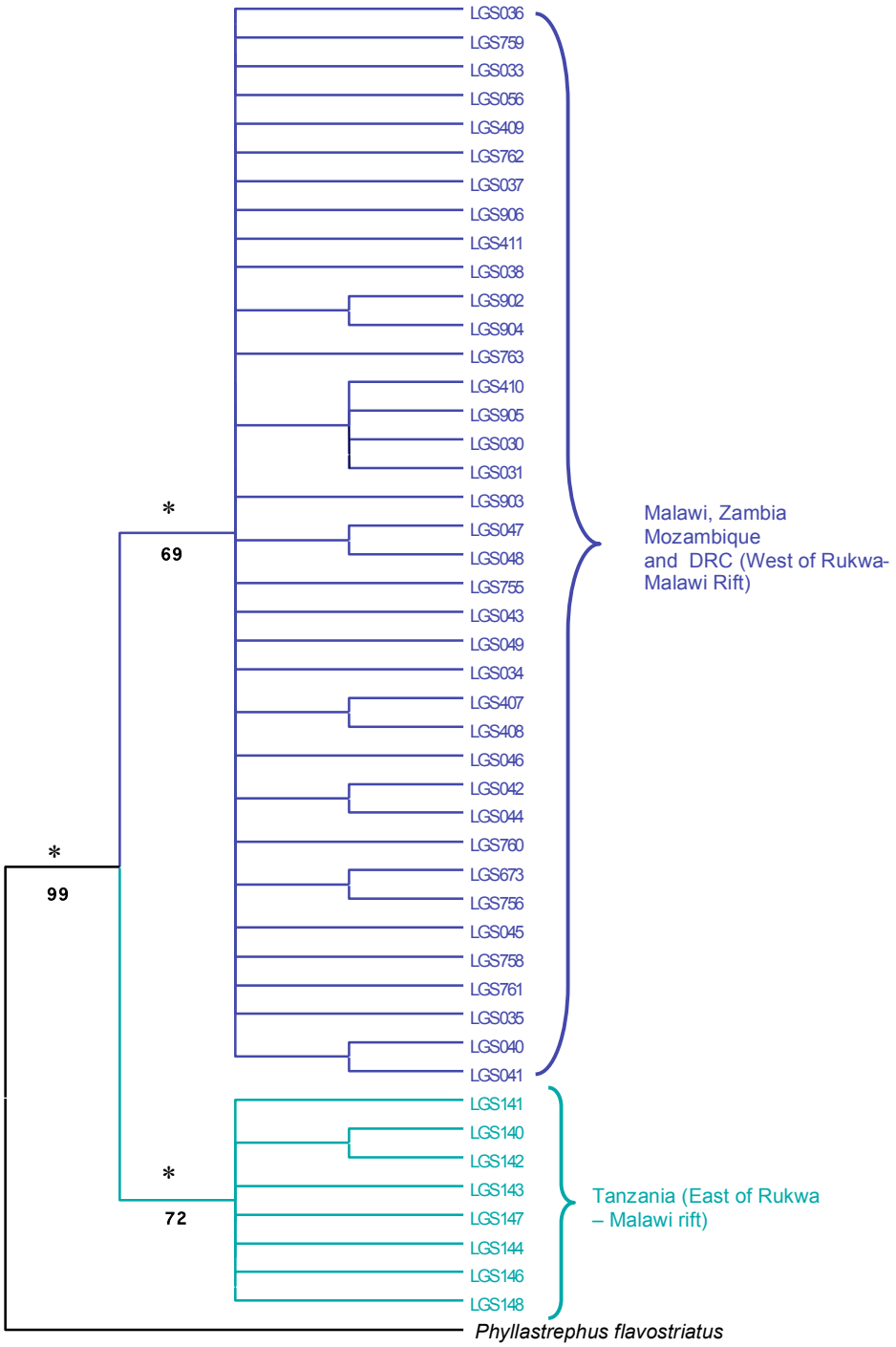


Figure 3.3 Majority rule consensus of the posterior distribution of trees obtained in the Bayesian analyses of the combined mtDNA data. Asterisks (*) at nodes indicated a posterior probability of greater than 0.90. Numbers at nodes indicate ML bootstrap support obtained with GARLI. In both analysis trees were rooted with *Phyllastrephus flavostriatus.*

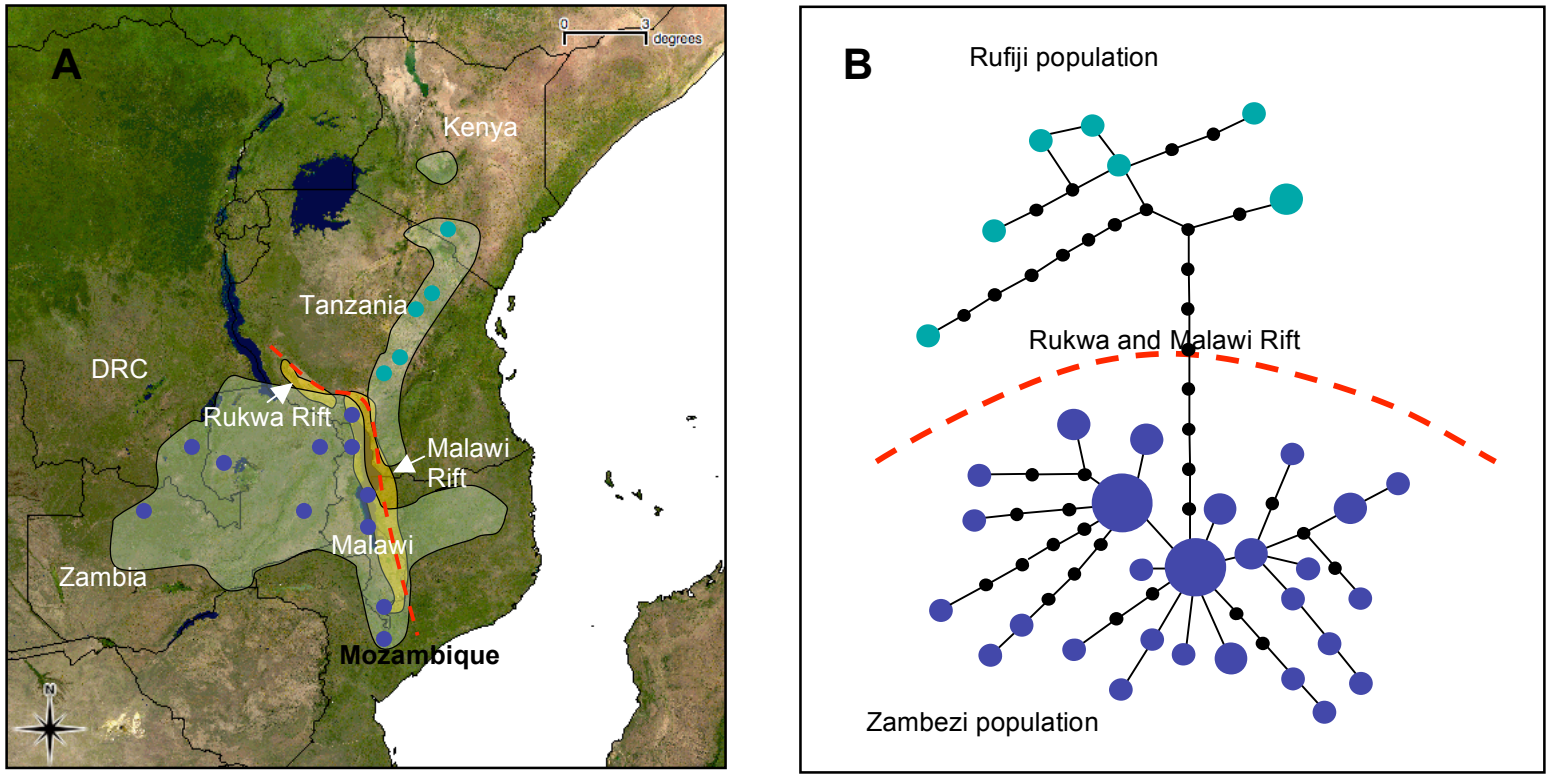


Figure 3.4 - Statistical parsimony networks for the combined mitochondrial DNA data set showing the geographical separation of haplotypes. (A) depicts the geographical separation of haplotypres across the Rukwa-Malawi Rift. Green coloured dots indicate samples from the Rufiji (Tanzania) population and the dots in blue indicate samples from the Zambezi (Malawi, Zambia, DRC and Mozambique) population. (B) shows the corresponding statistical parsimony network as computed using TCS 1.18 (Clement *et al.* 2000). Small black color dots in B indicate extinct or unsampled haplotypes and the circles in blue and green represent the Zambezi and Rufiji haplotypes respectively. Size is proportional to the number of individuals representative of each haplotype.

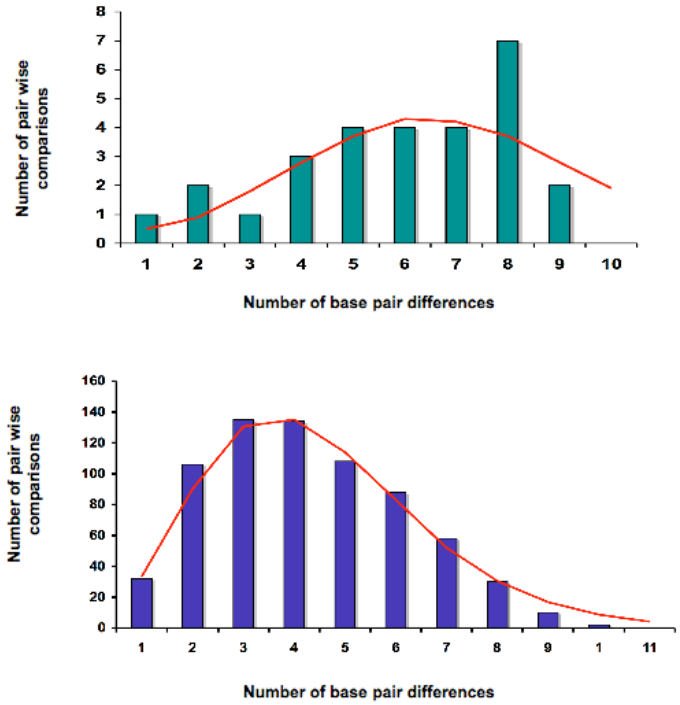


Figure 3.5 - Mismatch distribution of Rufiji and Zambezi populations of *Phyllastrephus cerviniventris* obtained with the program Arlequin version 3.0 (Schneider *et al.* 2000). Solid bars in green and blue represent observed data, and the red lines represent the simulated data used to test the goodness-of-fit to a sudden expansion model.

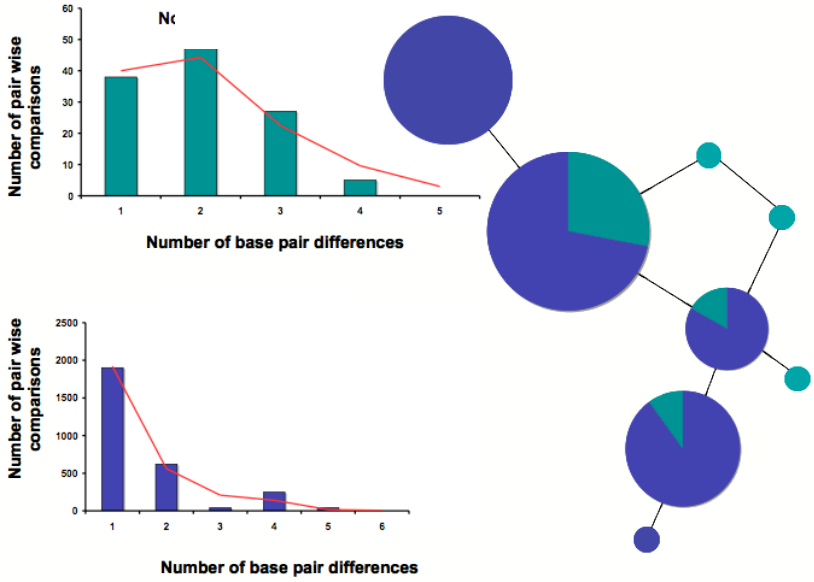


Figure 3.6 - Statistical parsimony network obtained with TS 1.18 (Clement *et al*. 2000) for the phased haplotypic data for the nuclear intron FIB5. Circles are drawn to scale and are proportionate to the number of individuals with a specific allele. Blue indicates the Zambezi population and green indicates the Rufiji population. Corresponding mismatch distributions are also presented

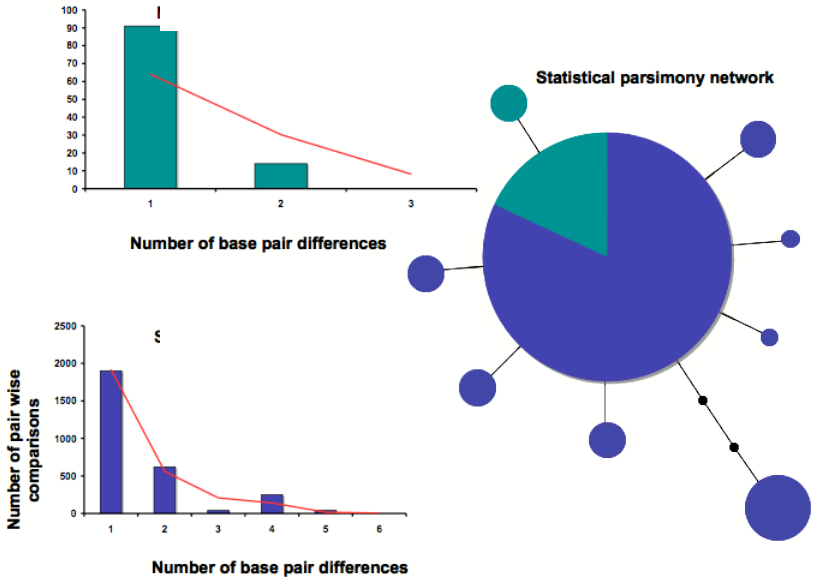


Figure 3.7 - Statistical parsimony network obtained with TS 1.18 (Clement *et al.* 2000) for the phased haplotypic data for the nuclear intron ODC. Circles are drawn to the scale and are propionate to the number of individuals with a specific allele. Black color dots indicate missing or unsampled haplotypes. Blue indicates the Zambezi population and green indicates the Rufiji population. Corresponding mismatch distributions are also presented.

CHAPTER 4

Phylogeography of an ancient endemic of the southern and eastern African coastal forest; the Sombre Greenbul (*Andropadus importunus*): evidence for palaeoclimate shaping population structure

4.1: Introduction

Pliocene-Pleistocene (the last ca. 5.3 Ma) climatic variation is often invoked to explain many patterns of faunal diversification in Africa (e.g. Crowe & Crowe 1982; Diamond & Hamilton 1980; Mayr & O'Hara 1986; deMenocal 2004; Fjeldså & Bowie 2008; Voelker *et al.* 2010). Despite the increasing number of phylogeographic studies of African bird species, a closer look at the literature reveals a strong bias to investigating species' that occupy montane highlands (e.g. Roy 1997; Beresford *et al.* 2003; Bowie *et al.* 2004b, 2006). In contrast, bird species that occupy more open habitats in Africa, which have been a primary focus of mammalian phylogeographic studies (e.g. Flagstad *et al.* 2001; Moodley & Bruford 2007; Lorenzen *et al.* 2010), have been sorely neglected by avian phylogeographers. Thus, it remains to be established if the largely concomitant patterns observed for montane taxa (Fjeldså & Bowie 2008) are also recovered in lowland taxa.

One such habitat configuration that has received almost no attention despite a suite of endemic bird species (Burgess *et al.* 1998), is the mosaic of lowland forests distributed along the coastal plain of east and southern Africa. A better understanding of the pattern of genetic diversity among coastal forest taxa is important, as many populations of these taxa are now fragmented as a consequence of the extensive loss of lowland coastal forest (Fjeldså & Lovett 1997; Burgess *et al*. 1998; Rodgers 2000; Lawes *et al.* 2007). Further, recent coalescent based analysis of the widely distributed montane white-starred robin (*Pogonocichlsa stellta*) suggested that coastal forests in East Africa may be providing a vital habitat corridor, thereby maintaining gene flow between isolated montane populations (Bowie *et al.* 2006).

In this chapter, I focus on investigating population structure in the coastal endemic Sombre greenbul (*Andropadus importunus*, Pycnonotidae). This species' widespread occupancy of coastal forest in east and southern Africa, its polytypic plumage, and because of it's basal phylogenetic position within the large African greenbul radiation (Fjeldså *et al.* 2007; Johansson *et al.* 2007) makes the Sombre greenbul an ideal taxon with which to investigate the influence of global climatic

perturbations on the genetic structure and diversity of a coastal forest species. By comprehensive sampling throughout the species' distribution range, I specifically investigate phylogeographic structure within and among populations of Sombre greenbul in order to: (1) test the pattern and extent of genetic variation within widespread coastal forest bird species, and (2) highlight areas where genetic uniqueness occurs in order to inform local government and conservation agencies.

Coastal forest of eastern and southern Africa

The contemporary biological diversity present along the 4000 km stretch of eastern and southern African coast has led this ecological zone to be considered an important area of endemism in Africa (Fjeldså & Lovett 1997; Burgess *et al.* 1998; Burgess & Clarke 2000). At least five endemic bird areas (East African coastal forests, Eastern Zimbabwe mountains, Pemba island, South-east African coast and South African forests) have been recognized within the coastal belt extending from SE Somalia to the Eastern Cape Province in South Africa (Stattersfield *et al.* 1998; Fishpool & Evans 2001). This landscape is diverse and occupies a range of different geomorphological features, from flat and low-lying coastal forest and shrub next to the sea (e.g Arabuka-Sokoke Forest in SE Kenya and Ngoye Forest in South Africa), to river valleys dissecting the plain (e.g. the Tugela River in South Africa and Shire River in Malawi), as well as uplifted areas which rise up to 800 m above sea level (e.g. Mts Gorongosa and Namuli in Mozambique, the Rondo Plateau in Tanzania).

The fragmented forest space found within this heterogeneous landscape is often described as a mosaic (see White 1983 for details) owing to the diverse array of associated vegetation types. Eastern African coastal forests are located within the Swahili regional centre of endemism and Swahili-Maputaland regional transition zone in eastern Africa, between 1° North and 25° South, and 34-41° East (Burgess *et al.* 1998). Along the southeastern coastline, the forests are typically referred to as the Indian Ocean coastal belt (Lawes 2004) and are included in the KwaZulu-Cape

Coastal Forest Mosaic (Moll and White 1978; Geldenhuys 1992) and Maputaland coastal forest mosaic (Moll & White 1978).

Complex biological interactions driven by different paleogeological events (e.g. post Gondwanaland fragmentation such as caused by uplifting, rifting and marine introgression; King 1978; Axelord & Raven 1978) and palaeoclimatic oscillations (e.g. high amplitude glacial cycles and subsequent aridification; Hamilton 1981; Fjeldså & Lovett 1997; Clerk & Burgess 2000; Clerk 2000) have been used to explain the development and distribution of forest cover along Africa's eastern and southern coast. . The localized peaks of endemism with a high degree of local turnover in the species between adjacent forest fragments within the landscape mosaic, as well as the high incidence of bird and mammal species exhibiting disjunct distribution patterns characterize the remarkable pattern of endemism with in the region. These features have generally been interpreted to be a reflection of an ancient evolutionary history of the associated endemic lowland fauna (Burgess *et al.* 1998; Fjeldså & Bowie 2008).

Sombre greenbul

The sombre greenbul (*Andropadus importunus*) is an abundant but seldom seen resident endemic inhabiting coastal forest of southern and eastern Africa. Its present distributional range spans the southern and eastern coasts of South Africa, Swaziland, Mozambique, Tanzania, Kenya and Somalia, and the lowland forested areas of Zimbabwe, eastern Zambia, and Malawi (Figure 4.1). In general this species' range is littoral forests with extensions inland to varying degrees along river valleys. They are shy and inconspicuous, found in dense cover, usually in the canopy. The species is thought to be monogamous, territorial (Keith 1992) and a solitary nester. Sombre greenbuls occur in pairs or, when non-breeding, in small family groups. They feed on insects, small snails, fruits, berries and succulent leaves and buds. The species is a persistent vocalist, singing all day throughout the year. The Sombre greenbul (Keith 1992) is polytypic with four described subspecies (Hockey *et al.* 2005) (Table4.1).

4.1: Materials and Methods

Population sampling

A total of 108 samples (Table 4.2 and Appendix 4.1) spanning the known distribution range of the Sombre greenbul and representing the four recognized subspecies (Hockey *et al.* 2005) were included in this study (Figure 4.1). In order to obtain thorough sampling throughout the present distribution I had to depend heavily on toe-pads from museum skins, as fresh tissue samples could not be easily collected from certain areas due to the presence of landmines in Mozambique. A basal member of the African greenbul radiation, *Calyptocichla serina* and a member of the "Asian" greenbul radiation *Pycnonotus tricolor (*Moyle & Marks 2006; Johansson *et al.* 2007; Zuccon & Ericson 2010) were selected as outgroup taxa.

Laboratory procedures

Tissue (blood or frozen tissue; n=50) or toe-pads (n=58) from museum skins were used to extract total genomic DNA using either a Puregene DNA isolation kit (Gentra Systems, Minneapolis, Minnesota) or a DNeasy Tissue Kit (Qiagen, Beckman Instruments Inc®) with an overnight proteinase K digestion at 56°C. In the case of museum skins, protein digestion time was expanded to 24h and double the volume of tissue lysis buffer and Protinase K was added (DNeasy Tissue Kit) together with 10mM DTT (dithiothreitol) solution. Extraction and amplification of museum skin samples was carried out in a dedicated biosafety level two laboratory following stringent precautions to avoid contamination.

Three mitochondrial and two nuclear gene regions were amplified and sequenced for all 108 individuals. The mitochondrial regions were ND2, ND3 and a fragment of ATP6. The nuclear regions sequenced were Beta-fibrinogen intron 5 (Fib5) and Glyceraldehyde Phosphate Dehydrogenase intron 11 (GAPD). Several internal primers were developed to amplify and sequence genomic DNA from museum skins (Table 4.3). PCR-amplification conditions follow the protocols in Bowie *et al.* (2004) for ND2 and ND3, Fuchs *et al.* (2004) for Fib5 and Fjeldså *et al.*

(2003) for GAPD. However, longer annealing times (45-60s) with an increased number of amplification cycles (40-45) were employed for templates from museum skins despite the smaller fragment sizes (250-350 bp.). Amplitaq Gold (Roche) DNA polymerase which has a proof reading ability was used to amplify PCR products to increase the accuracy of base calls and to minimize possible taq errors. PCR products were electrophoresed on 1.5% agarose gels, stained with ethidium bromide and visualized under UV light and purified using standard protocols using either GELase™ (EPICENTRE® Biotechnologies) or ExoSap-IT (USB Corporation, OH, USA). The purified products were cycle-sequenced using Big Dye terminator chemistry (Applied Biosystems, Inc [ABI]), precipitated with 3M ammonium acetate or 100% isopropanol, rinsed in ethanol, dried and resuspended in formamide-EDTA solution, and run on either an AB 3100 or AB 3730 automated DNA sequencer. Sequences were obtained from both strands of DNA for each individual, and some individuals were sequenced several times if any base ambiguity was encountered.

Length polymorphism was detected in some individuals for Fib5 and GAPD and these PCR products were cloned into the PCR®2.1-TOPO® vector (Invitrogen) and transformed into TOP10 chemically competent *Escherichia coli* cells following the manufacturer's recommendations (TOPO TA Cloning Kit, Invitrogen). Eight positive colonies were selected from each plate and used as a template for PCR-amplification to check the size of the insert. Colonies with inserts of the anticipated size were then purified using the standard Miniprep protocol (Sambrook *et al.* 1989) and sequenced in both directions using the M13 forward and reverse primers or internal sequencing primers (Table 4.3) developed in this study.

Sequence alignment

All mitochondrial DNA (mtDNA) sequences were checked using the program Sequencher 4.7 (Gene Codes Corp) to determine if there were any insertions or deletions, as well as to establish that no stop codons were present in the coding sequence. Nuclear intron sequences were

also edited and assembled using Sequencher 4.7 (Gene Codes Corp). For the two intron data sets allelic phase was determined using PHASE 2.1 (Stephens *et al.* 2001; Stephens & Donnelly 2003). I preformed two independent phase estimates using the recombination model as implemented in PHASE 2.1, with different starting seeds to check for consistency across runs. Individuals with excessive missing data or poor allelic phase estimates were excluded from our per-locus analyses. This resulted in the final alignments containing 101 individuals and 106 individuals for Fib5 and GAPD, respectively.

Phylogenetic analyses

Phylogenetic trees were estimated using the model-based methods of maximum likelihood (ML) and Bayesian inference (BI). All three mtDNA gene regions I sequenced showed high congruence in base composition and substitution dynamics, hence I concatenated the three mitochondrial gene regions for further analysis as mtDNA is inherited as a singe linked locus. However, the two nuclear introns I sequenced differ considerably in their base composition and substitution dynamics hence, intron data were analyzed separately. Two different maximum likelihood (ML) analysis methods were performed on the combined mtDNA dataset as implemented in GARLI v.95 (Zwickl 2006) and RAxML (Stamatakis *et al.* 2008). In the RAxML analysis the mtDNA data set was partitioned into first, second and third codon positions for each gene. Clade support for the ML analyses were estimated using 100 nonparametric bootstrap pseudoreplicates (Felsenstein 1985).

The computer program MrBayes 3.1 (Huelsenbeck & Ronquist 2001) was used to conduct a Bayesian approach to phylogenetic inference. The GTR + I model was selected as the best-fit model for the combined mtDNA data set with the assistance of MrModeltest (Nylander 2004) and the Akaike information criterion (Akaike 1973). Two replicates with four Metropolis-coupled Markov chains Monte-Carlo simulations (one cold and three heated) were run for five million iterations, with trees sampled every 100 iterations. The first 300,000 iterations (3000 trees) were discarded ('burn-

in' period), and the posterior probabilities were estimated from the remaining sampled iterations. The progress of runs was checked periodically with the program AWTY (Wilgenbusch *et al.* 2004), a web based tool to graphically assess convergence of the MCMC algorithm by plotting posterior probabilities of a given set of 20 splits during the course of a run (Figure 4.2).

Population genetic analysis

Statistical parsimony networks were constructed for each intron dataset and for the combined mtDNA dataset using TCS 1.21 (Clement *et al.* 2000). Gaps were treated as a fifth character state in the intron data sets. Analysis of molecular variance (AMOVA) using Tamura and Nei's pairwise distances (1000 permutations) (Excoffier *et al.* 1992) was used to the estimate the extent to which genetic variability was partitioned within and among major lineages. Two data partitions were analyzed independently for the combined mtDNA dataset and two intron data sets; first by partitioning the individuals by subspecies designation, and second by partitioning the individuals according to the three main mtDNA clades recovered in the ML and BI analyses. AMOVA analyses were conducted in Arlequin 3.1 (Schneider *et al.* 2000). Pairwise mismatch distributions for the three haplotype clusters recovered in the ML and BI analyses were estimated for the combined mtDNA and nuclear intron datasets. Two tests of selective neutrality, Tajima's D (Tajima 1989) and Fu's Fs test (Fu 1997) were calculated for the combined mtDNA and both nuclear intron datasets to test whether populations were in mutation-drift equilibrium under an infinite sites model as an indication of recent demographic change, which could have resulted from either population expansion or contraction (Rogers & Harpending 1992; Harpending *et al.* 1993: Fu 1997). Both tests were implemented using Arlequin 3.1 (Schneider *et al.* 2000).

Estimation of divergence times

I used Beast v1.4.6 (Drummond *et al.* 2002, 2006; Drummond & Rambaut 2007), to estimate divergence dates within *A. importunus* using mtDNA sequence data. I assigned the best-fitting model, as estimated by MrModeltest2 to each of the partitions. I used two alternate mutation rates,

the classical rate of $1.05*10^{-8}$ substitutions/site/year (s/s/y) for mtDNA (2.1% per million years, Weir & Schluter 2008), as well as the faster rate of $3.05*10^{-8}$ (6.1% per million years, Arbogast *et al.* 2006, see also Warren *et al.* 2003). This extrapolation carries with it a significant margin of error and thus I emphasize the importance of thinking of the time estimates only as a rough attempt to place diversification events within *A. importunus* in a historical context. I assumed a Yule Speciation Process for the tree prior and an strict molecular clock model for the clock rate (Drummond *et al.* 2006; Ho 2007). I used default prior distributions for all other parameters and ran MCMC chains for 20 million generations. The analysis was repeated twice. I used the program Tracer (Rambaut & Drummond 2007) to assess convergence diagnostics.

Gene flow versus ancestral polymorphism

Since most of the computer programs use to test historical demography (e.g. IMa; Hey & Nielsen 2004) assume that there has been no recombination or gene conversion within the genealogical history of a locus, I used the Difference of Sums of Squares (DSS) test (McGuire & Wright 2000) as implemented in the program TOPALI v 2.5 (McGuire & Wright 2000) to test for evidence of recombination in both intron datasets. Based on the result of the DSS test, I carried out a Hudson-Kreitman-Aguade (HKA) test (Hudson *et al.* 1987) to test for neutrality across the two nuclear loci with 100,000 simulations using the program HKA (Jody Hey, Rutgers University). Population divergence times, historical gene flow, and effective population sizes were estimated using a model of Isolation with Migration as implemented in the software IMa (Hey & Nielsen 2004). In this case, the basic 6-parameter model was chosen. This includes q1, q2, and qA, referring to the effective size of population 1, population 2 and the ancestral population, respectively; m1, the migration rate of genes from population 1 to population 2 in the coalescent (or, as time moves forward, from population 2 to population 1) and m2, migration rate in the opposite direction; and *t*, the time since divergence of the populations. As the program estimates only the parameters between two populations in a single run, the analyses were conducted three times for three

different data partitions; between clade 1 & clade 2, between clade 2 & clade 3 and finally between clade 1 & clade 3. For each data partition, a series of test runs were carried out to check how the parameter estimates behave in coalescence analysis and finally when satisfying parameter estimates were seen with the convergence of MCMC chains, the detailed analyses were performed. I defined inheritance scales to reflect the difference in modes of inheritance among the loci used: 0.25 for the mtDNA locus and 1.0 for the two autosomal loci. I used an HKY model of nucleotide substitution for all three data partitions. Parameters and genealogies were sampled every 100 steps for both the 10 and 20 million step runs. I used a generation time of 1.7 years, which reflects the average for several passerine species (Sæther *et al.* 2005). I used the same two alternate mutation rates, described previously for the dating for the mtDNA partitions. For the autosomal loci I selected a rate of $3.61*10^{-9}$ s/s/y (Axelsson *et al.* 2006). This translated into per locus rates of: Fib5 $2.04*10^{-6}$ s/y, and GAPD $1.02*10^{-6}$ s/y. The geometric mean of the combined mtDNA and ncDNA loci was $4.4*10^{-6}$ s/y (mtDNA 2.1%/ Ma) or $6.41*10^{-6}$ s/y (6.1%/ Ma), and for the two nuclear loci $2.23*10^{-6}$ s/y.

4.3: Results

Mitochondrial DNA Phylogeography

The mitochondrial DNA matrix contained 94 ingroup haplotypes out of 108 individuals sampled for 2026 bp (1041 ND2, 351 ND3 and 634 ATP6) of which 328 (16%) were parsimony informative (Table 4.4). Both ML and BI analyses of the combined mtDNA data set produced a strongly supported topology of three clades with strong geographic structuring (Figure 4.3). The haplotype clusters have a biogeographic rather than a taxonomic component. The only exception to this being that the two individuals sampled from Hluhluwe, South Africa cluster with clades two and three rather than to one clade alone In fact, Hluhlulwe forms the clade boundary between clade 2 and 3. This three clade relationship is largely incongruent with the currently recognized four subspecies taxonomy of the Sombre greenbul (Fig. 4.3).

Mitochondrial sequence divergence values estimated using uncorrected p-distances among the three clades were 7.6+- 0.2% between clades 1 and clade 2, 6.9+- 0.2% between clades 1 and 3, and 4.6+-0.01% between clades 2 and 3 suggesting that these three clades have been separated for a long time. In contrast, estimated sequence divergence values within clades range from 0.4+-0.3% to 0.2+-0.1%. AMOVA for the three mtDNA clades indicated significant structuring of the genetic variability among clades (95%) rather than within clades (5%) with a highly significant Fst value of 0.95 (d.f.=2, $P < 0.0001$). Partitioning the individuals into the subspecies groups, yielded an Fst value of 0.70 (d.f. = 3, $P < 0.001$) and high within subspecies (94%) and low among subspecies (6%) genetic variability, suggesting that grouping the individuals by geography has a better fit to the data than grouping them by subspecies, although the low amount of variability attributed to among-subspecies variation does suggest that these taxa have some validity.

A statistical parsimony network for the combined mtDNA resulted in three unlinked networks that correspond to the three major clades recovered in the ML and BI analyses (Figure 4.6). The network for clade 1 consisted exclusively of unshared or unique haplotypes. Networks for clades 2

and 3 indicated a signature of recent expansion by having characteristic star-like patterns in which several more recently derived haplotypes are connected to a centrally located common ancestral haplotype by one or two mutational steps. Two haplotypes from Zanzibar Island and one from extreme SE Tanzania (Nambiga Forest) form a distinct sub-network with 13 mutational step differences to the main ancestral haplotype in the network for clade 3. Clades 1 and 2 had significantly negative Tajima's D and Fu's Fs statisticss (Table 4.4) and mismatch profiles which follow a Poisson distribution (Figure 4.4), suggesting that these populations may have undergone recent population growth following range expansion. The mismatch profile for clade 3 has a multi-modal distribution, as a consequence of the divergent haplotypes from SE Tanzania and Zanzibar.

Nuclear gene phylogeography and diversity

The final nuclear intron alignments after pruning both 3' and 5' ends to remove missing data consisted of 323 bp and 555 bp for GAPD and FIB5 respectively. Merging identical sequences in the GAPD and FIB5 alignments (after estimating haplotypic phase) yielded 53 and 49 unique alleles respectively. Both nuclear loci revealed substantial genetic variation across the three recovered mtDNA lineages. The number of haplotypes, segregating sites, gene (H) and nucleotide (π) diversities for both introns are provided in Table 4.6. However ML analyses conducted on GAPD and FIB5 (not shown here) were incongruent with each other as well as with the combined mtDNA ML tree. Further the ML trees for introns neither resolve the expected relationships nor obtain high support (poorly resolved) for the branches.

Both nuclear loci indicate similar patterns of genetic variability when analyzed using either subspecies grouping or the three main mtDNA clades. AMOVA for the three main clades revealed Fst values of 0.49 (d.f.=2, $P < 0.001$) and 0.04 (d.f. = 3, $P < 0.001$) for Fib5 and GAPD respectively. Fib5 indicated roughly equal partitioning of genetic variability between and among clades (49%) and within clades (51%), whereas in GAPD much of the genetic variability (95%) lay within clades. When analyzed as the subspecies groups, a Fst value of 0.40 (d.f. = 3, $P < 0.001$)

and higher (60%) within group and lower (40%) among groups genetic variability were recovered for Fib5. In contrast a Fst value of 0.04 (d.f. = 3, $P < 0.001$) was recovered for GAPD, with most of the genetic variability (96%) occurring within groups rather than among groups (4%).

The statistical parsimony network constructed for either GAPD or FIB5 phased haplotyic data (Figure 4.6) did not sort into reciprocally monophyletic clades generally recovered from the mtDNA data. The networks for the nuclear introns show multiple shared haplotypes involving individuals from the three major mtDNA clades. GAPD shows a signature for recent expansion by having a star-like pattern. The central ancestral haplotype in the GAPD network was shared in a 2:1:1 ratio among mtDNA clades 2, 3 and 1 respectively. FIB5 had a network that is more complex, but still largely unstructured, with several quite divergent alleles (Figure 4.6). All three clades have significantly negative Tajima's D and Fu's F statistics (Table 4.6) and mismatch profiles which follow a Poisson distribution (Figure 4.7), suggesting that these populations may have undergone recent population growth following range expansion.

Dating analyses

Divergence date estimates obtained using a molecular clock hypothesis for the two clock rates (6.1% and 2.1%) implemented in BEAST suggest that the lineages north of Zanzibar Island (clade 1) and south of Kahe forest, Tanzania (clades 2+3) diverged between 1.6563 Mya (95% HPD: 1.3494-2.1269) and 4.4776 Mya (95% HPD: 2.8738-8.2942) depending on which clock rate was used (Table 4.6). The 95% HPD interval does not overlap between these two alternate rates of divergence. Under both divergence rate estimates the TMRCA of the clade 2 was slightly older than for clade 1 and clade 2 (Table 4.6). The divergence between the lineages south of Kahe forest in Tanzania (clade 3) and west of Hluhluwe (clade 2) is estimated to have occurred about 3.2 Mya (95% HPD: 1.726-5.118myrs) or 0.9 (95% HPD 0.781-1.349).

Evidence for a model of isolation-with-migration

The DSS algorithms implemented in Topali v 2.5 (McGuire & Wright 2000) did not detect any evidence of recombination within the two nuclear loci Hence, I used the complete sequence alignments of each intron separately for the IM analyses. Appropriate mixing and satisfactory effective sampled sizes were achieved using 12 chains and a geometric heating scheme (g1=0.15 and g2=0.70). Isolation with migration derived peak posterior distribution estimates and 90% highest posterior densities (HPD) of demographic parameters (scaled by neutral mutation rate) of three data partitions (mtDNA, Fib5 and GAPD). For all three comparisons, the sharp peak of the mode of the marginal posterior distribution suggested that the population migration rates between the two populations concerned were low although the confidence interval was large (Table 4.7). For the comparison between the clade 1 and the clade 3, all models that assume no gene flow or unidirectional gene flow could be rejected. A model of symmetrical gene flow could not be rejected. Similarly, although the respective modes of the marginal distribution of theta (effective population size scaled by mutation rate) differed among the two extant populations (clade1-19.947, clade3-35.753) as well as relative to the ancestral population size (24.256), likelihood-ratio tests suggested that equal effective population sizes could not be rejected. However, any model that assumed equal population size and migration rates could be rejected. For the other two comparisons (between clades 1&2 and clades 2 & 3) peak posterior distribution estimates of population diversity ($\emptyset = 4 N\mu$) of the two population concerned since divergence, and of their ancestor were all different, with ancestral theta estimated to be larger than both extant populations concerned (Table 4.7). All models that assume no gene flow and symmetrical gene flow, and all models that assume equal effective population sizes in the two current lineages were strongly rejected (all $P<0.001$) for the coalescence analysis done on both clades 1 & 2 and clades 2 & 3.

4.4: Discussion

This study has provided a detailed analysis of genetic variation within a widespread southern and east African coastal forest endemic bird, the Sombre greenbul *(Anrdopadus importunus).* In fact, this was the first detailed phylogeographic study covering most of the southern and east African coastal forest (except Ethiopian coast) regions. The Sombre greenbul being a basal node of the larger African greenbul radiation (Johansson *et al.* 2007, Zuccon & Ericson 2010) also provides a new perspective on the evolutionary history of the African greenbuls. All three mtDNA clades recovered in this study are reciprocally monophyletic and this is largely incongruent with the currently recognized four subspecies taxonomy. Since, all three haplotype clusters have a biogeographic rather than a taxonomic component, this work justifies the need for a detailed reappraisal of the current taxonomy of the Sombre greenbul.

Genetic variation

The analysis revealed that *A. importunus* is composed of three well-supported lineages that strongly correlated with their geographical locations. The analysis of molecular variance indicated that the current taxonomy does not adequately reflect the extent of genetic variation within *A. importunus* suggesting that grouping the individuals by geography offers a better fit to the data than grouping them by subspecies, although the low amount of variability attributed to among subspecies variation does suggest that these taxa have some validity. This is particularly true for the nominal subspecies (clade 2) and the subspecies *A. i. insularis* (clade1). I found evidence of recent population expansion in the mtDNA dataset for all three clades as revealed by significant Fu's Fs (Table 4.4) and coalescent analyses (Table 4.7) and the shape of the mismatch profiles (Figures 4.5). Population expansion also is evident in the intron data sets particularly with the GAPD data set (Table 4.6 and Figure 4.7). Given the significance values from several tests of demographic change across the data sets and the number of segregating sites per locus (Tables 4.4 and 4.6), it is likely that the demographic expansions, when present, were moderate and not

sudden. The shared nuclear alleles (Table 4.6) could be a consequence of lineage sorting not yet having been completed due to slow rate of evolution of the nuclear markers (see also Zink & Barrowclough 2008).

Biogeography

We estimated that the first split, the split between northern *insularis* (clade 1) and the two southern clades (clades 2+3) within *A. importunus* occurred during the Pliocene around 4.48 Mya (95% CI = 2.87-8.29) or during the early Pleistocene 1.66 Mya (95% CI = 1.35-2.13), depending which clock rate was use, 2% and 6.1% respectively. The southern clades (clades 2 and 3) split about 3.19 Mya (95% CI = 1.732-5.12) or 0.90 Mya (95% CI = 0.78-1.35). This indicates an ancient split and together with the basal position of *Andropadus importunus* in the African greenbul phylogeny (Johansson *et al.* 2007; Fjeldså *et al.* 2007; Zuccon & Ericson 2010) suggests that this species has been widely distributed in coastal forest for a long time. The, confidence intervals estimated for these two splits for the two clock rates used did not overlap. However, in a recent dating analysis of the major clades within the Pycnonotidae by Fuchs *et al.* (in press) using multiple calibration points has revealed that the youngest age estimate of African greenbuls is around 3.1 Mya. Therefore the estimated divergence time based on the 2% clock rate seems to be somewhat unrealistic.

According to my coalescent analyses the southern clades (clades 2 and 3) are expanding and that is probably why the two specimens collected from Hluhluwe belong to these two clades. The historical signal is now being obscured as these clades through expansion are coming into contact with each other at clade boundaries as evident from the isolation with migration parameter estimates (Table 4.7). According to the TMRCA estimated for individual clades, clade 3 has the deepest within clade divergences. This could be explained by the retention of larger ancestral

population size (Table 4.7) and ancestral diversity of clade 3 as estimated from the coalescence analyses.

Over the past 10 million years the climatic trend has been for ever increasing aridity, with the Milankovitch climatic fluctuations superimposed onto this trend (cf.Bennet, 1990; 1995, 2004). Periodically, however, there have been wetter climatic regimes (e.g. between 9 and 6.4 million years, and between 4.6 and 2.43 million years ago) (Lovett, 1993) in east Africa that resulted in expansion of forest cover. The more severe climatic fluctuations during the Holocene (Ice Ages) have further reduced forest cover inland from the eastern African coast, although close to the coast, relatively moist and warm climatic conditions are believed to have persisted as the Indian Ocean remained at a similar temperature as today (Prell *et al.* 1980). Thus the Ice Ages may not have had as much of an effect on the coastal forests as the longer term and slower gradual desiccation of the area. The East African coastal zone has been under a constant climatic influence from the Indian Ocean (Lovett and Wasser 1993) and seems to be a centre of diversification, or ‘‘species pump’’, as populations of different phylogenetic ages could persist in suitable places, and could eventually disperse and enrich the regional biota (Fjeldså & Lovett 1997; Jetz *et al.* 2004).

Over the past two million years, southern Africa has experienced some 20 climatic cycles, each lasting 100,000 years and mirroring periods of expansion and contraction of the glacial ice sheets at higher latitudes (Deacon, 1983; Tyson, 1986; Deacon & Lancaster, 1988). Hyperthermal periods have been characterized in general by warmer, wetter conditions and the expansion of the forest biome, while hypothermal periods experienced cooler drier climates and a reduction in the extent of forest (Deacon, 1983; Tyson, 1986; Deacon & Lancaster, 1988). However, westerly winds are unlikely to have had any major effect on patterns of rainfall in the more easterly regions and KwaZulu-Natal Province of South Africa. Here, generally cold, dry conditions prevailed during the winter, with strong winds and cold air drainage off the Drakensberg mountains exacerbating the drying effect, particularly in the southern part of the province where the high mountains (altitude

1750 m) lie closer to the coast. Thus conditions in KwaZulu-Natal during the LGM were considerably colder and drier than at present, leading to a contraction in forest extent in the province, during the Holocene altithermal conditions were warmer and wetter and conducive to forest growth and expansion.

Two phylogeographic breaks in coastal forests

Clade one recovered in this study forms a geographically contiguous range from the northern most boundary of the species' distribution in S Ethiopia to extreme NE Tanzania and consists exclusively of individuals from subspecies *A. i. insularis.* Clade 2 include samples from the area extending from the species southeastern boundary (Cape Town) up to the west of Hluhluwe in Kwazulu-Natal, South Africa and all the individuals included were from the nominal subspecies *A. i. importunus*. Clade 3 consists of the samples from areas extending from east of Hluhluwe up to Central Tanzania including Zanzibar Island, Mozambique, Zimbabwe and Malawi. Therefore the clade 3 consists of individuals representing all four currently recognized subspecies. Two phylogeographic barriers for faunal evolution in the southern and east African coastal forest could be recognized. The northern barrier lies somewhere between Kahe forest and Zanzibar Island, probably along the 115 km stretch of lowland forest between the Usambara and Nguu Mts. My sampling does not cover each and every forest patch along this stretch of coastal forest, it is difficult to conclude the exact boundaries of this break. Recent literature on east African montane bird phylogeography (e.g. Eastern Double-collared Sunbirds, Bowie *et al*. 2004a; Olive thrushes, Bowie *et al.* 2005 and Voelker *et al.* 2007; Green barbet R.C.K. Bowie, M. Thomasset, M.J. Lawes & J. Fjeldså unpublished data; Streaky canary R.C.K. Bowie, E. Mostert, J.M. Bates & J. Fjeldså , unpublished data; spot throat R.C.K. Bowie & J. Fjeldså unpublished data, see also Fjeldså and Bowie 2008) has suggested a genetic break between the northern Eastern Arc (Taita Hills, Pare and Usambara Mts.) and the central Eastern Arc (Nguu, Nguru, Ukaguru and Rubeho Mts) (Fjeldså & Bowie 2008). However no comparable literature is available for bird species exclusively

inhabiting coastal forest. The second break seems to lie in Kwazulu-Natal, South Africa as evident by the shared haplotypes collected from Hluhluwe. By close examination of locality details of the haplotypes collected in South Africa it is clear that this genetic break could lie along the Maloti, Drakensberg and Witteberg Mountain ranges and Tugela valley as all the samples collected from localities east of this, cluster into mtDNA clade 3 recovered in this study except at Hluhluwe which is probably a current contact zone for clades 2 and 3. The Maloti, Drakensberg and Witteberg Mountain ranges cover the 700 km stretch of mountains from its southern extreme near the Eastern Cape Province town of Elliot in South Africa and straddle the eastern Lesotho-South Africa border northwards to Golden Gate Highlands National Park in the Free State Province of South Africa. Several studies have highlighted the importance of the region for faunal turnover (e.g. for amphibians (Poynton & Boycott 1996); dragonflies (Davis *et al.*1999); mammals (Lawes 2004); millipedes (Hamer & Slotow 2009).

Taxonomic implications

All three mtDNA clades recovered in this study are reciprocally monophyletic and this is largely incongruent with the currently recognized four subspecies taxonomy. Since, all three haplotype clusters have a biogeographic rather than a taxonomic component, this justifies a detailed reappraisal of the current taxonomy of the Sombre greenbul. This is also supported by the high genetic distance (uncorrected p=7%) observed between clade one and two and may warrant elevation of subspecies *A. i. importunus* and *A. i. insularis* to species level should morphological and vocal characters be consistent with the genetic data.

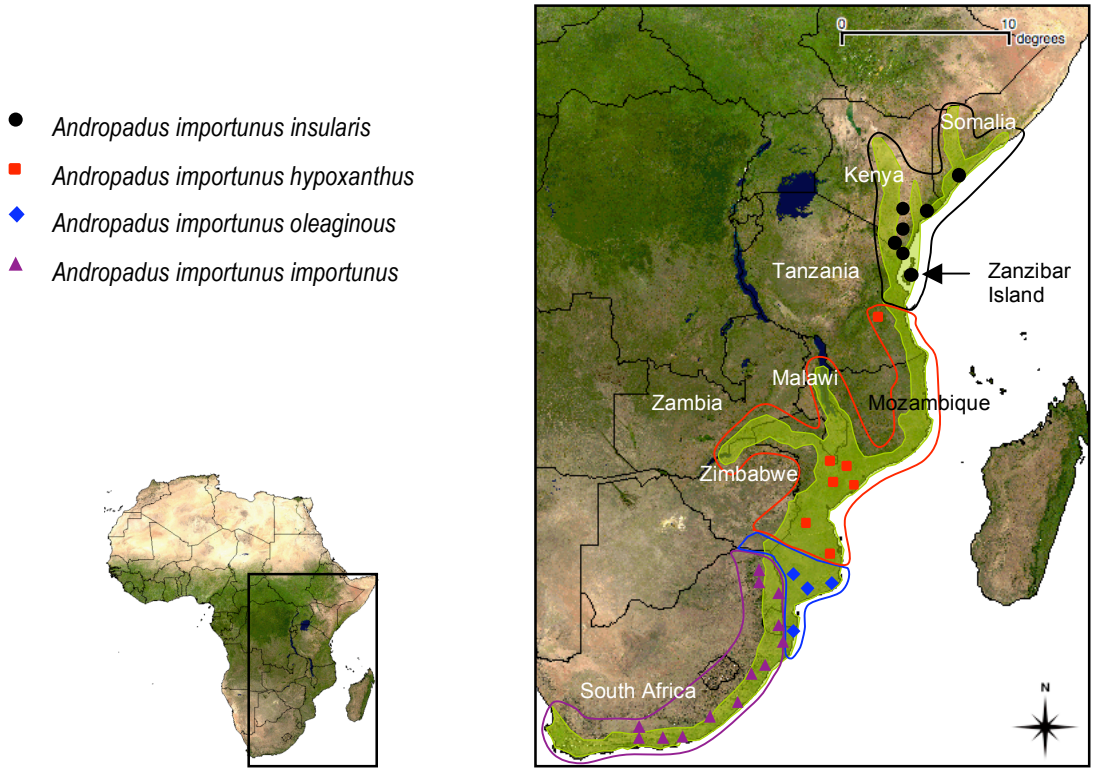


Figure 4.1 - Geographic distribution of Sombre greenbul in southern and east African coastal forest and sample localities. (Modified from Keith 1992). Shaded area in light green indicates the present distribution range and the colored symbols indicate sample localities for each subspecies. Areas demarcated by colored lines indicate the distribution of the subspecies.

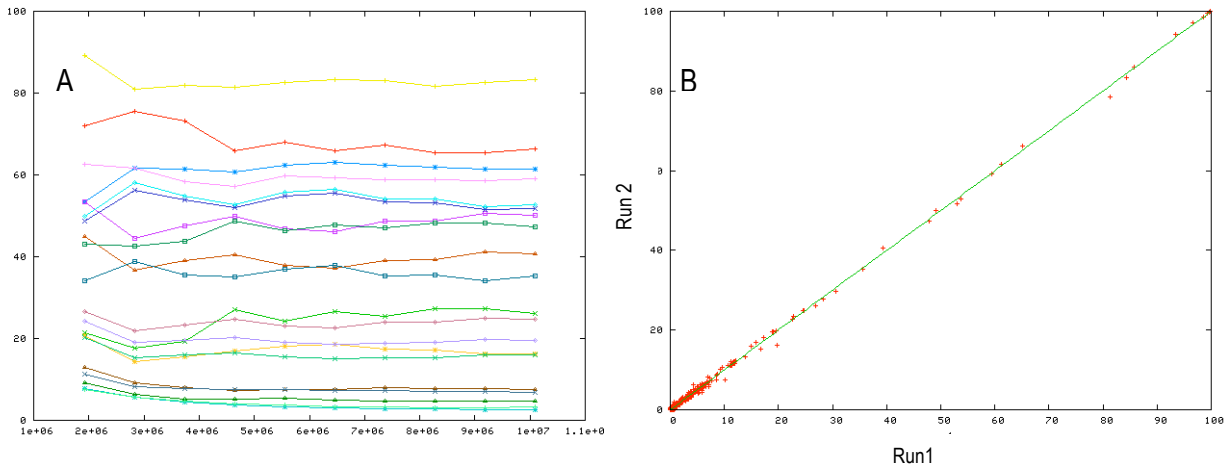


Figure 4.2 - Visual exploration of MCMC convergence with the web based tool AWTY (Wilgenbusch *et al.* 2004). Each line in 'A' represents the combined posterior probability of splits 1-20 during the course of the Bayesian phylogenetic analysis of mtDNA DNA sequences performed with MrBayes3.1.1 (Ronquist & Huelsenbeck 2003). In this case, convergence occurred slightly beyond 5 x 10^6 generations. 'B' shows the convergence between run 1 and 2.

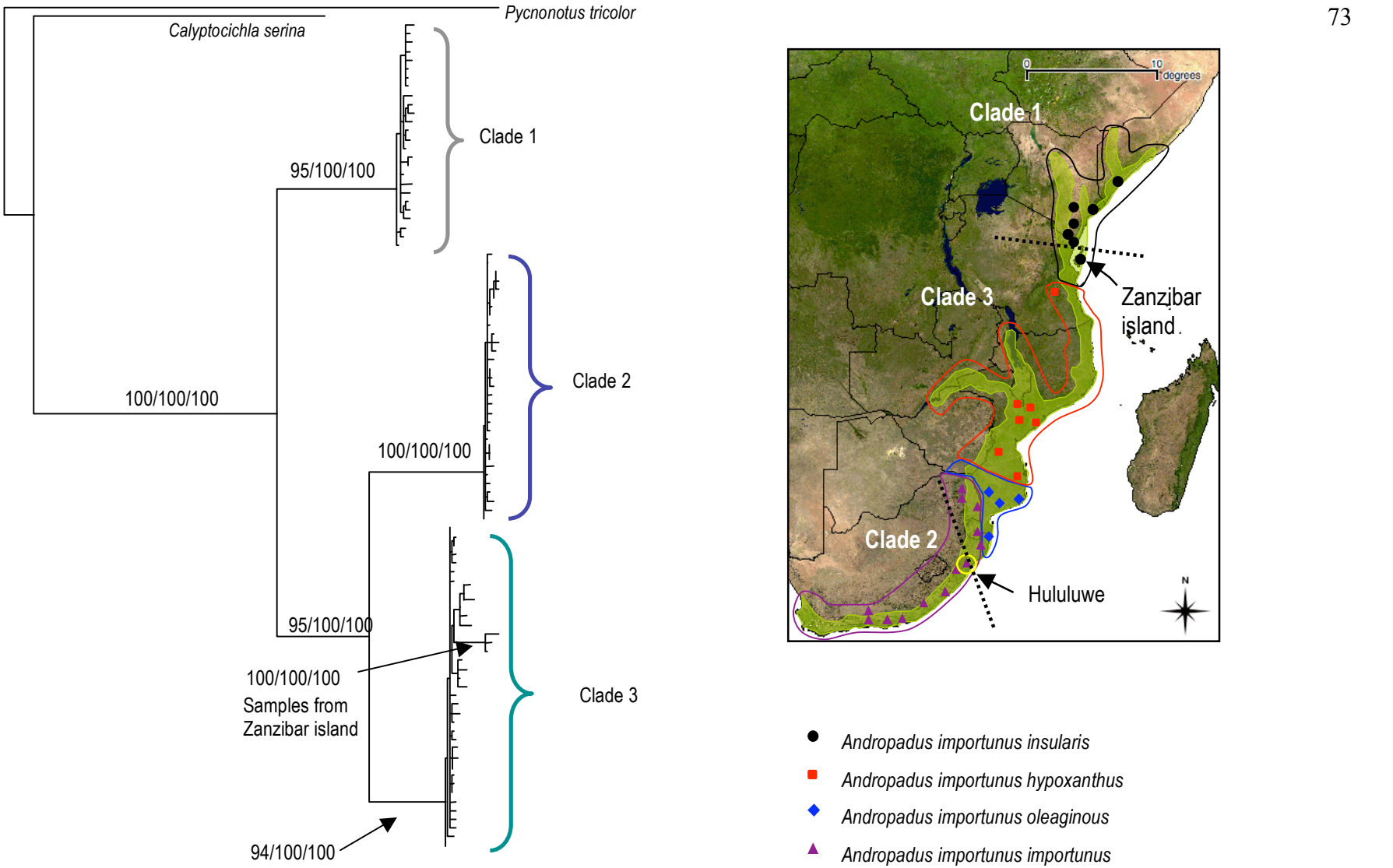


Figure 4.3 - Maximum likelihood tree obtained with the combined mtDNA data. Numbers at nodes indicate ML bootstrap obtained with GARLI (left), RAxML (middle), and Bayesian percent posterior probabilities (right). In all three analysis trees were rooted with *Pycnonotus tricolor* and *Calyptocichla serina*. Dotted lines in the map indicate the clade boundaries recovered in the study. Colored lines indicate subspecies distribution ranges and the colored symbols indicate sample localities for each subspecies. Hululuwe is part of the contact zone between clade 2 and 3 where we sampled haplotypes belong to both clades.

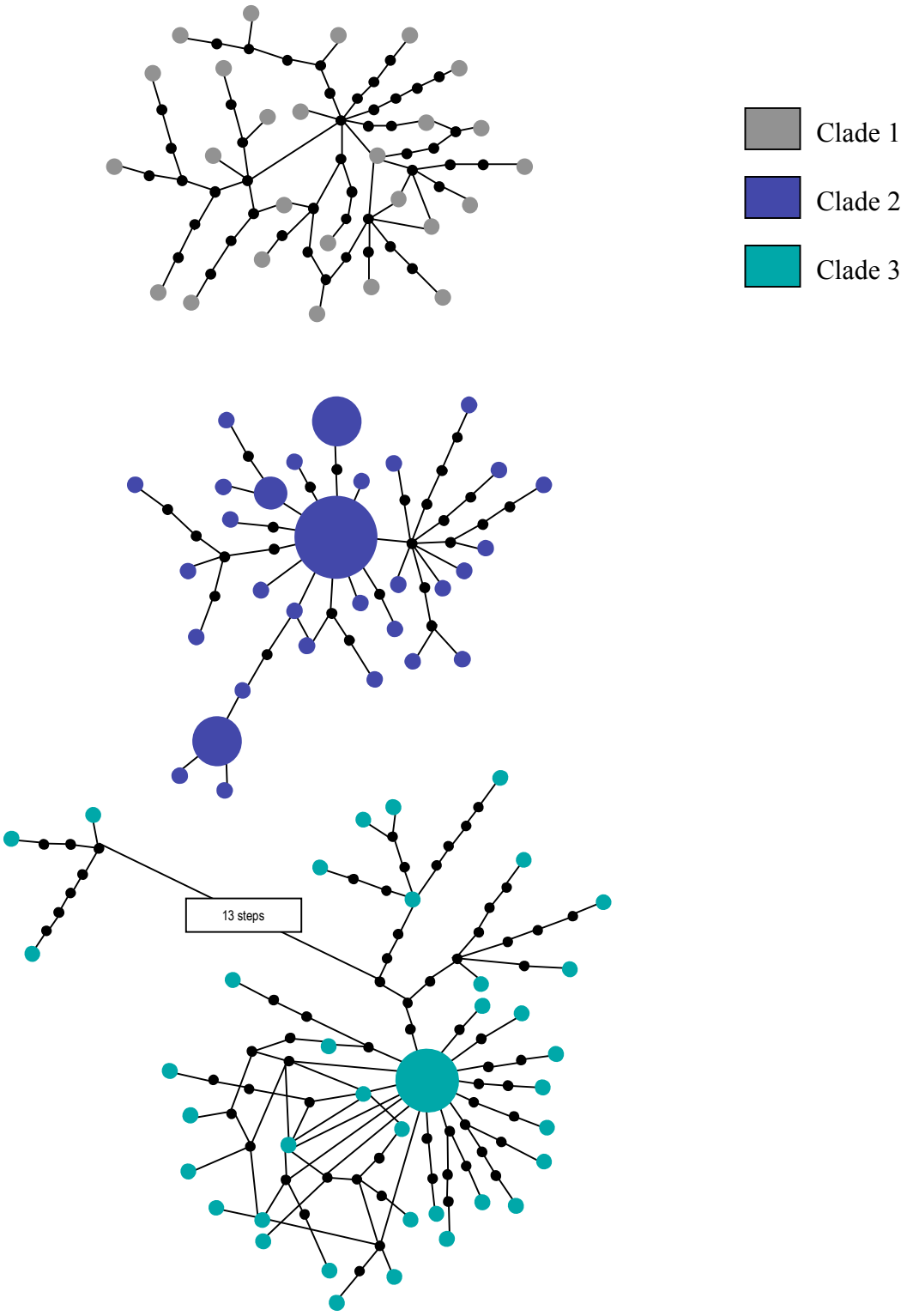


Figure 4.4 - Unrooted statistical parsimony haplotype networks for the three main mtDNA clades recovered in this study. Circle diameter is proportional to the number of individuals possessing this haplotype. Extinct or unsampled haplotypes are indicated by black dots.

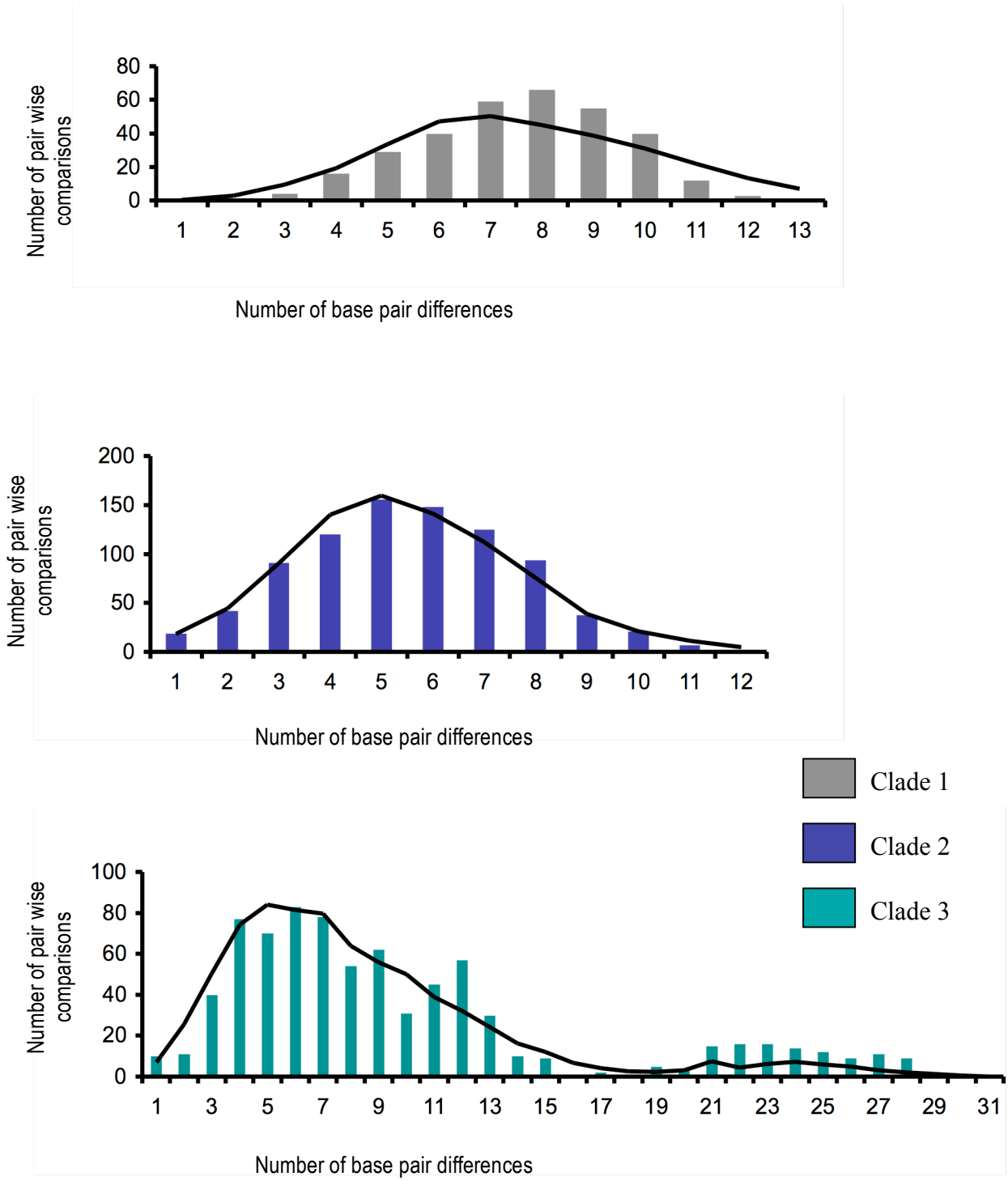


Figure 4.5 - Mismatch distributions for three main mtDNA clades recovered in this study for *Andropardus importunus.* Colored bars in the mismatch graph represent observed data and lines represent the simulated data used to test the goodness of fit to a model of sudden expansion.

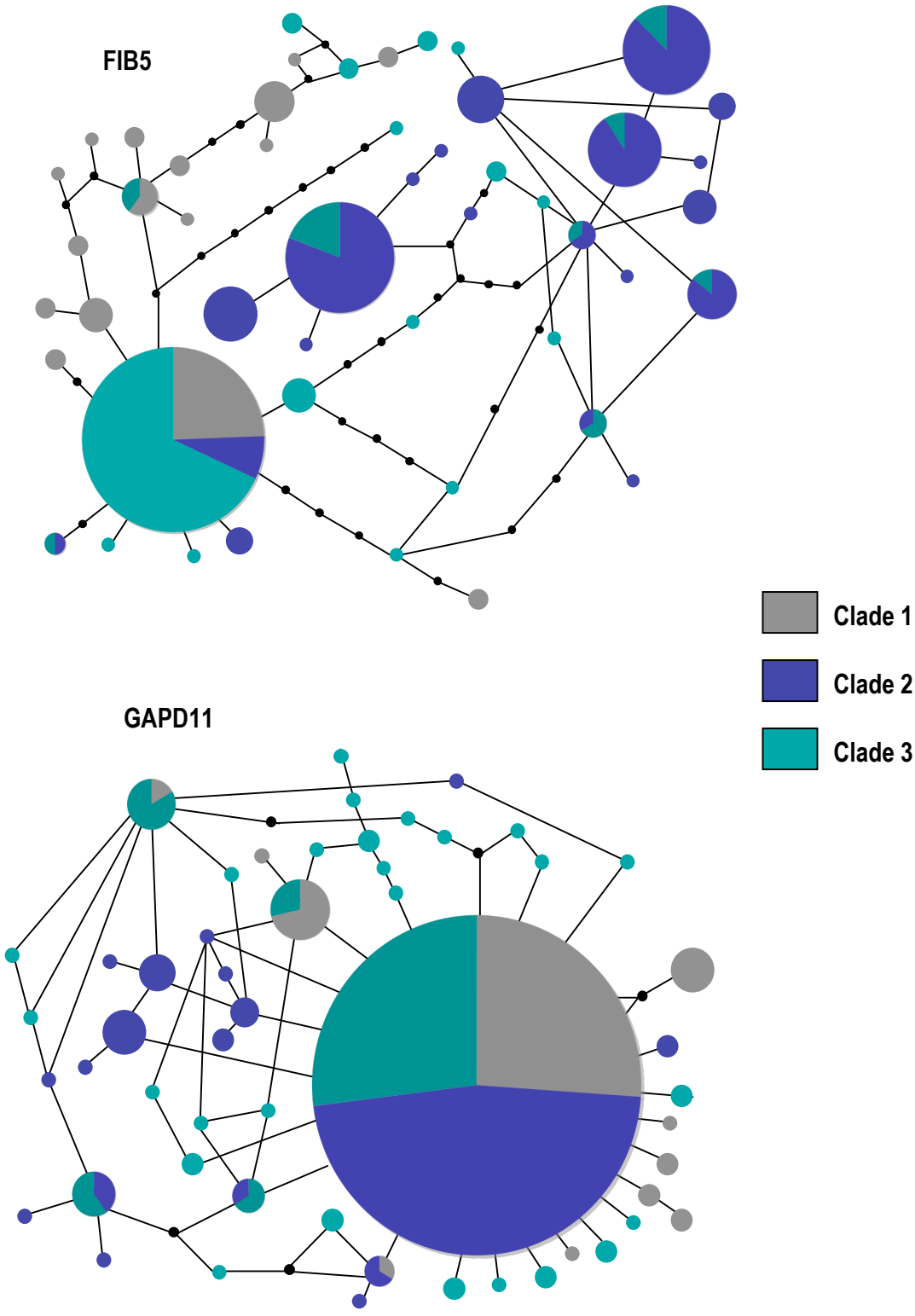


Figure 4.6 Unrooted haplotype statistical parsimony network of the FIB5 and GAPDH11 introns. Circle diameters are proportional to the number of individuals possessing this haplotype. Extinct or unsampled haplotypes are indicated by black dots.

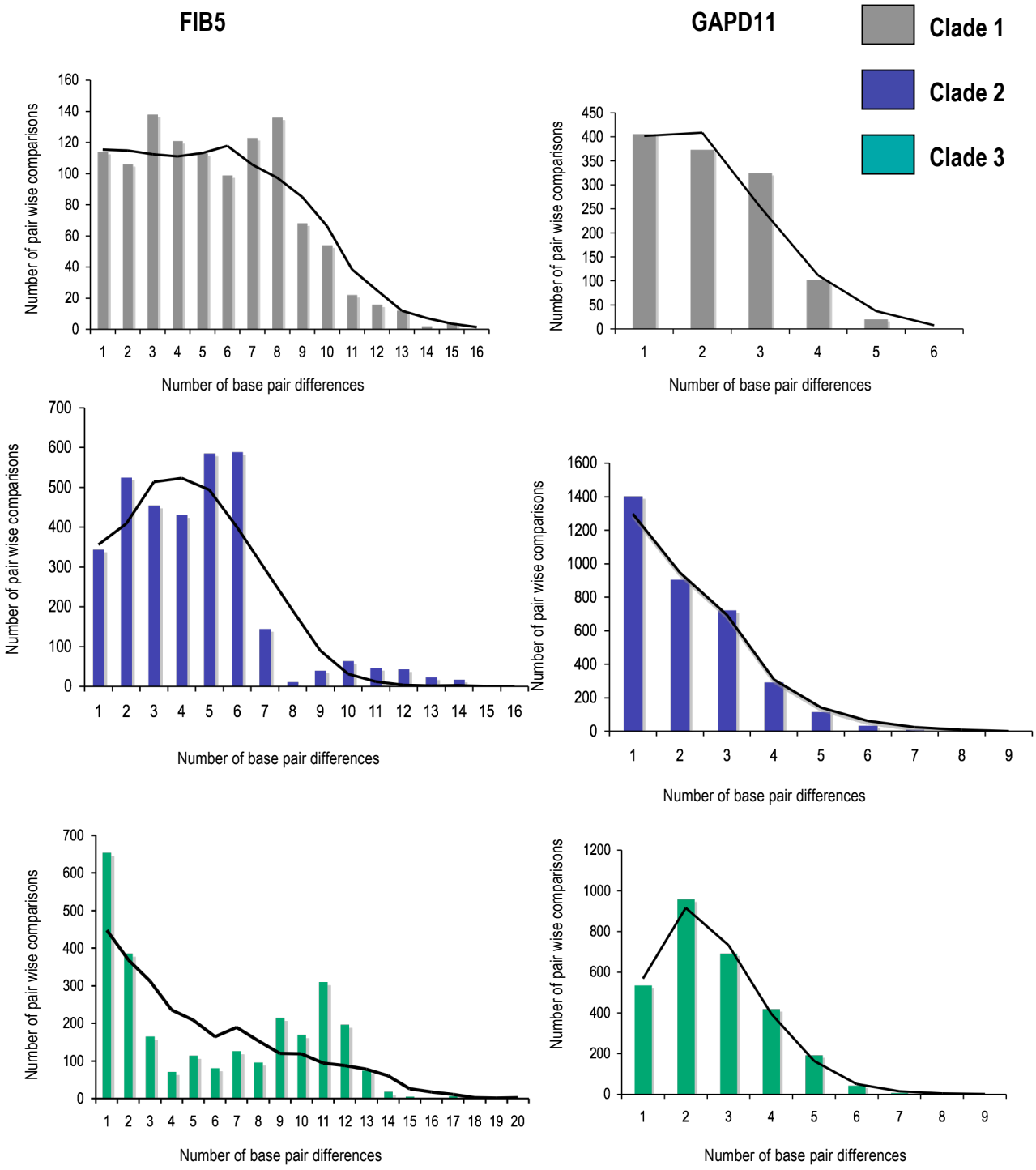


Figure 4.7 - Mismatch distributions for the introns FIB5 and GAPD11 when the data is partioned according to the three main mtDNA clades recovered in this study. Colored bars in the mismatch graph represent observed data and lines represent the simulated data used to test the goodness of fit to a model of sudden expansion.

Table 4.1 - Morphological variation and geographical distribution of the Sombre greenbul (*Andropadus importunus)* species complex (Keith1992**)**

Subspecies	*Upper part coloration*	*Lower part coloration*	Geographic distribution
Andropadus importunus insularis (Hartlaub)	Top and sides of head including wing coverts brownish-olive. Eye, white to pale yellow. Dark olive brown flight feathers without green margins. Darker tail feathers sometimes with narrow whitish green outer margins.	Throat breast and flank lighter brownish-olive, paler on chin with a pale yellow belly. Rich brownish-yellow under tail coverts and thighs, bright ochreous-yellow axillaries, under wing coverts and inner web of flight feathers.	S Ethiopia, Somalia S of *ca.* 4^{0}N chiefly along the Juba and Webbe Shebelle rivers, Kenya and Tanzania, except extreme SE Tanzania.
Andropadus importunus importunus (Vieillot)	Much darker than other races. Upper parts including wing are dark olive green.	Under parts uniform olive grey with slightly paler belly. Under wing is pale yellow without ochreous tinge.	South Africa (Cape Province, KwaZulu-Natal, S Zululand, highlands of Limpopo and Mupumalanga Provinces) and Swaziland
Andropadus importunus oleaginous (Peters)	Slightly paler upper parts than *A. i. insularis*	Grayish yellow throat, upper breast and flank. Bright rich yellow lower breast, belly and under tail coverts.	South Africa (N Zululand, lowlands of Mupumalanga), E Swaziland, Mozambique S of Save river, Zimbabwe along Limpopo river.
Andropadus importunus hypoxanthus (Sharpe)	Greener upper parts than *A. i. oleaginous*.	Brighter yellow underparts.	Mozambique N of Save river, Zimbabwe except along Limpopo river, Zambia, Malawi and SE Tanzania.

Table 4.2 - Sample origins and the number of individuals included in the study

Subspecies	*Number of samples used*	*Extraction numbers**
Andropadus importunus insularis	28	LGS020, LGS634 to LGS640, LGS642, LGS646, LGS771 – LGS772, LGS774, LGS776 to LGS784, LGS789, LGS791 to LGS793, LGS796, LGS848
Andropadus importunus importunus	58	LGS022 to LGS029, LGS050 to LGS055, LGS079 to LGS081, LGS086 to LGS088, LGS612 to LGS614, LGS775, LGS785 – LGS786, LGS788, LGS790, LGS801, LGS806 – LGS807, LGS814, LGS823 - LGS824, LGS829 to LGS831, LGS840 to LGS843, LGS845, LGS850 – LGS851, LGS854 to LGS863, LGS865 to LGS868
Andropadus importunus oleaginous	11	LGS799, LGS804, LGS816, LGS821 - LGS822, LGS827 – LGS828, LGS837 to LGS839, LGS846
Andropadus importunus hypoxanthus	11	LGS641, LGS773, LGS800, LGS802, LGS808, LGS812 - LGS813, LGS819, LGS833, LGS835 - LGS836

* Detailed information (museum catalog number, source, country of origin, locality etc.) on the samples can be found in the Appendix 4.1).

Table 4.3 Novel primers specifically designed for *Andropardus importunus*. F and R stand for forward and reverse directions, respectively.

Primer name	*Gene region*	*Primer location within the gene*	*Primer sequence 5'-3'*
1. ND2AimR1	NADH subunit 2	328-346 bp	GGA YCC AGA AGT GGA ATG G
2. ND2AimF2	NADH subunit 2	283-302 bp	TGC CTR ATC CTA ACC TCA GC
3. ND2AimR2	NADH subunit 2	584-602 bp	GTG AGT TTG GGG TCR TAG G
4. ND2AimF3	NADH subunit 2	546-566 bp	CTC CTC YAT CTC MCA CYT AGG
5. ND2AimR3	NADH subunit 2	816-836 bp	GCT GCA GGG ATY ATG YTT TG
6. ND2AimF4	NADH subunit 2	749-767 bp	TCA CTR GCA GGC CTT CC
7. ND3AimF	NADH subunit 3	151-170 bp	CGA CTC CCM TTC TCA ATC C
8. ND3AimR	NADH subunit 3	201-220 bp	TAG GAG GGC GAT CTC TAG G
9. ATP6AimF	ATP6	310-330 bp	TTC CCC CTA TGA CTA GC
10. ATP6AimR	ATP6	340-355 bp	GCG GAT GGY TGG TTG C
11. FIB5AimF	FIB5	248-265 bp	GCA GRG TAA GAA ATG AGC CAG G
12. FIB5AimR	FIB5	270-291 bp	ATA GGT GGG CAG ACC AGG
13. GAPDHAimF	GAPDH11	148-169 bp	AGA CCT CCA GCC TCA GTG TGG
14. GAPDHAimR	GAPDH11	198-221 bp	GAG GGR AAG GAG GCA GYT ATA GT

Table 4.4 - Number of haplotypes, number of polymorphic sites, haplotype diversity (H), nucleotide diversity (π), Tajima's D and Fu's Fs statistics obtained when the combined mitochondrial data set is partitioned as subspecies or according to the three major clades recovered in the study.

	Main mtDNA clades			*Subspecies*			
	Clade 1	*Clade 2*	*Clade 3*	*Andropadus importunus importunus*	*Andropadus importunus insularis*	*Andropadus importunus hypoxanthus*	*Andropadus importunus oleaginus*
Number of individuals	26	42	40	58	28	11	11
Number of haplotypes	26	31	37	47	28	10	11
Number of polymorphic sites	51	52	96	162	186	33	36
H	1.0000+/- 0.011	0.9779+/- 0.012	0.9923+/- 0.009	0.9903+/- 0.006	1.0000+/- 0.010	0.9722+/- 0.064	1.0000+/- 0.030
π	0.0032+/- 0.002	0.0022+/- 0.001	0.0043+/- 0.002	0.0203+/- 0.010	0.0127+/- 0.006	0.0037+/- 0.002	0.0037+/- 0.002
Tajima's D	-1.9399 (P=0.012)	-2.2029 (P=0.002)	-2.2253 (P=0.004)	0.6871 (P=0.815)	-1.7860 (P=0.018)	-1.9245 (P=0.003)	-1.5590 (P=0.047)
Fu's Fs	-21.3495 (P=0.000)	-25.3816 (P=0.000)	-24.7512 (P=0.000)	-6.4151 (P=0.061)	-11.1159 (P=0.001)	-1.6806 (P=0.141)	-7.0938 (P=0.001)

Table 4.5 - Number of haplotypes, haplotype diversity (H), nucleotide diversity (π), Tajima's D and Fu' Fs statistics obtained for the introns FIB5 and GAPD11 when the datasets were partitioned according to the three major clades recovered in the study.

	FIB5			*GAPDH11*		
	Clade 1	*Clade 2*	*Clade 3*	*Clade 1*	*Clade 2*	*Clade 3*
Number of samples*	48	82	74	50	84	76
Number of haplotypes	17	19	24	11	18	30
Number of shared haplotypes	1 (between 1 & 2) 3 (between 1 & 3)	1(between 2 & 1) 9(between 2 & 3)	2(between 3 & 1) 9(between 3 & 2)	2 (between 1 & 2) 3 (between 1 & 3)	2(between 2 & 1) 3(between 2 & 3)	3(between 3 & 1) 3(between 3 & 2)
Number of polymorphic sites	27	18	30	10	12	20
H	0.8989 +/- 0.029	0.8964 +/- 0.016	0.7575 +/- 0.053	0.6686 +/- 0.070	0.5978 +/- 0.063	0.8649 +/- 0.036
π	0.0077 +/- 0.004	0.0061+/- 0.003	0.0087 +/- 0.005	0.0035 +/- 0.003	0.0035 +/- 0.002	0.0051 +/- 0.003
Tajima's D	-1.02485 (P=0.172)	-0.06771 (P=0.525)	-0.39024 (P=0.525)	-1.38832 (P=0.066)	-1.44152 (P=0.050)	-1.80358 (P=0.015)
Fu's Fs	-3.42849 (P=0.109)	-4.76651 (P=0.065)	-6.47341 (P=0.034)	-5.91018 (P=0.004)	-15.20157 (P=0.000)	-27.88973 (P=0.000)

*total number of haplotypes resulted after estimating allelic phase with the software PHASE v 2.1.1 together with the cloning of some PCR products.

Table 4.6 - Estimated time-to-most-recent-common-ancestry (TMRCA) of *A. importunus* and the three mtDNA lineages under two alternate clock rates. Upper and lower 95% confidence levels for estimates are presented in parenthesis.

Lineage	*TMRCA estimated at 2% clock rate*	*TMRCA estimated at 6.1% clock rate*
A . importunus sensu lato	4.4776 (2.8738-8.2942)	1.6563 (1.3494-2.1269)
Clade 2 + 3	3.194 (1.7261-5.1179)	0.9003 (0.7807-1.3490)
Clade 1	0.4436 (0.2052-0.7342)	0.1460 (0.0923-0.2061)
Clade 2	0.4645 (0.2229-0.7744)	0.1546 (0.0981-0.2151)
Clade 3	0.6777 (0.3222-1.1480)	0.2227 (0.1369-0.3148)

Table 4.7 - Demographic estimates under the Isolation with Migration model using the program IMa (Hey & Nielsen 2004). Values represent the mean of the distribution and the 90% HPD confident interval.

	clade 1 & clade 2	*cle 2 & clade 3*	*clade 1 & clade 3*
q1	20.125 (14.771-25.386)	12.435 (9.022-15.743)	19.947 (13.492-26.128)
q2	14.527 (10.765-18.276)	35.662 (26.445-44.491)	35.753 (25.602-45.741)
qA	44.057 (0.050-86.675)	51.882 (0.062-107.960)	24.256 (0.0658-37.8438)
m1	0.058 (0.002-0.072)	0.067 (0.002-0.137)	0.073 (0.002-0.157)
m2	0.027 (0.002- 0.117)	0.280 (0.117-0.437)	0.060 (0.002-0.132)

CHAPTER 5

Unraveling the origin of a “leapfrog” distribution pattern in the Yellow-streaked greenbul (*Phyllastrephus flavostriatus*) species complex in southern and eastern Africa

5.1: Introduction

Evolutionary biologists have long recognized the significance of studying the mechanisms underlying geographic variation in organisms (e.g. Mayr 1963, 1970; Gould & Johnston 1972; Zink & Remsen 1986) because the data generated from such studies are crucial for testing hypotheses of phylogenetic relationships and provide valuable insight into the processes of speciation (Hayes 2001). Among the diverse geographic variation of plumage patterns in African songbird (Passeriformes) species, one of the most striking is 'leapfrog' distribution pattern owing to its scarcity among African birds (Fjeldså & Lovett 1997). Remsen (1984) was the first to highlight this evolutionary phenomenon in birds, where geographically disjunct populations are separated by an intervening but morphologically distinct population/s of what is presumably the same species. Leapfrogging is known to be most frequent among humid Andean montane bird species (Remsen 1984; Johnson 2002) and has been reported in several other avian species complexes (e.g. Schodde & Calaby 1972; Ford 1983; Haeys 2001; Norman *et al.* 2002; Sánchez-González *et al.* 2007;Cadena *et al.* in press). Such patterns also exist in a variety of other organisms including insects (Hovanitz 1940), plants (Matsumura *et al*. 2006), and amphibians (Noonan & Gaucher 2006).

Despite nearly a century of study, explicit tests of mechanistic explanations for leapfrog patterns are scarce (Cadena *et al.* in press). Hypotheses proposed to explain the origin of leapfrog distributions can be divided into two groups (Remsen 1984; Norman *et al.* 2002). In one, phenotypic similarity of geographically distant terminal populations is attributed to their phylogenetic affinities. The other group of hypotheses attributes leapfrogging patterns to divergent patterns of phenotypic change among geographically proximate populations (Remsen 1984; Chapman 1939). The phylogenetic affinity hypotheses typically invoke demographic explanations (Cracraft 1983; Johnson 2002), such as long-distance dispersal, loss of connections between currently disjunct populations, or extinction of intervening populations, to explain leapfrogging patterns. In contrast,

hypotheses which focus on divergent patterns of phenotypic change, posit that either convergent evolution of distant populations or accelerated divergence of central populations compared to those at the periphery of species' range is the route cause of this pattern (Remsen 1984; Fjeldså & Krabbe 1990; García-Moreno & Fjeldså 2000).

Fjeldså and Lovett (1997) state that leapfrog patterns in morpholological variation among African songbird species are rare. Three species that exhibit this morphological pattern, together with the Yellow-streaked greenbul examined here, are the Mountain Greenbul (*Andropadus tephrolaemus*), and the Bar-throated Apalis (*Apalis thoracica)* and Chestnut-throated Apalis (*Apalis porphyrolaema)*. Fjeldså and Lovett (1997) suggest that the potential vicariance induced by the Eastern Arc montane circle could be the unifying underlying evolutionary process generating these patterns. In general, models of speciation in African songbird species have stressed either the role of isolation or ecological gradients (Smith *et al.* 2004). For example, the effects of glacial refugia on isolated mountains has commonly been invoked to explain the formation and present-day distribution of closely related, but largely allopatric, songbird taxa (Moreau 1966; Haffer 1969; Crowe and Crowe 1982; Mayr & O'Hara 1986, Fjeldså & Lovett 1997; Roy 1997; Garcia-Moreno & Fjeldså 2000; Beresford *et al.* 2005; Bowie *et al.* 2004a, 2006; Voelker *et al.* 2010). In contrast, ecological gradient hypotheses have been highlighted as a way in which phenotypic diversity can be generated among habitat mosaics even in species, that are widespread and abundant (Smith *et al.* 2004).

The Yellow-streaked greenbul

The Yellow-streaked greenbul (*Phyllastrephus flavostriatus,* Pycnonotidae) is a common polytypic songbird with a highly disjunct distribution pattern in southern and east African forests (Figure 5.1). In the western half of its distribution range, it largely inhabits montane rainforest of the Albertine Rift (Uganda, eastern DRC, Rwanda and Burundi) as a series of allopatric populations. In East Africa the species occurs in both the montane and lowland coastal forests of Tanzania, the

eastern part of Zambia, and the northern and southern Malawi highlands (Keith 1992). In southern Africa the species' distribution is patchy occurring in the montane forests of the highlands of eastern Zimbabwe and adjacent Mozambique, as well as in the coastal forests of eastern Mozambique and South Africa (Hockey *et al.* 2005).

Distinguishing features of the species are the yellow colour streaks found from the lower breast to the under-tail coverts, medium body size, long bill and dark-brown irides eyes. The species is represented by eight recognized subspecies (Table 5.1) based on plumage variation and geographic distribution (Keith 1992). Of the eight recognized subspecies seven have a grey head and green upper parts, and one *P. f. alfredi* has a brown head with brownish upper parts. The geographical distribution of grey-headed and brown-headed forms of the Yellow-streaked greenbuls form 'leapfrog' pattern in morphology where the grey-headed subspecies in coastal (South Africa to Tanzania) and montane (Albertine Rift and Eastern Arc Mountains) parts of the range are separated from each other by the brown-headed *P. f. alfredi* found in southwestern Tanzania, Zambia and northern Malawi. Further, some forms of grey-headed Yellow-streaked greenbuls do seem to have different but subtle colour morphs within their respective distributional ranges (Keith 1992). Owing to variation in head color, Hall and Moreau (1970) have regarded the brown-headed *alfredi* as an incipient species. However, due to its resemblance in habit, vocalization, and egg morphology to *Phyllastrephus flavostriatus vincenti* (Keith 1992) ambiguity in taxonomic status remains. Most of the recent bird checklists recognize the brown-headed form as full species *Phyllastrephus alfredi* (e.g. Sibley & Monroe 1996; Dickinson 2003). Recent phylogentic analysis (Johansson *et al.* 2007) of african greenbuls using a limited number of samples suggest that *alfredi* may be a distinct species.

Here, I use a mitochondrial DNA-based phylogenetic approach to understand the pattern of genetic variation within the Yellow-streaked greenbul species complex in order to elucidate the evolutionary history of this unusual and complex geographical differentiation by sampling

throughout the known distribution range of the species complex. Further, I examined the potential consequences that major Quaternary geological and climatic events in eastern and southern Africa had on the evolution of genetic and morphological diversity in the *Phyllastrephus flavostriatus species* complex. Thereby, I test the applicability of phylogenetic hypothesis for the evolution of a leapfrog distribution pattern of morphological diversity in Yellow-streaked greenbul.

5.2: Materials and Methods

Population sampling

A total of 248 individuals representing brown-headed and grey-headed populations of the Yellow-streaked greenbul complex were included in this study (Table 5.2; Appendix 5.1). This represents samples from all eight recognized subspecies (Table 5.1) and includes samples from nearly every known allopatric population (Figure. 5.1). Three or more individuals were included for all populations sampled except for Itombwe forest, DRC and Mahale Mountain National Park, Tanzania where we could obtain only two and one sample respectively. Four related greenbul species (*Andropadus importunus*, *Phyllastrephus cerviniventris, Phyllastrephus placidus and Phyllastrephus terrestris)* were selected based on a recently published African greenbul phylogeny (Johansson *et al.* 2007) as the out-group taxa.

Laboratory procedure

Both fresh material (either blood or frozen tissue) and toe-pads from museum skins were used to extract total genomic DNA. Either the Puregene DNA isolation kit (Gentra Systems, Minneapolis, Minnesota) or a DNeasy Tissue Kit (Qiagen, Hilden) were used for the extraction of genomic DNA following the manufacturer's standard protocols with an overnight proteinase K digestion at 56°C. In the case of museum skins, protein digestion time was expanded to 24h and double the volume of tissue lysis buffer and Protinase K enzyme (DNeasy Tissue Kit) together with 10mM DTT (dieritheritiole) solution was added to facilitate tissue digestion. Extraction and amplification of museum skin samples was carried out in a dedicated laboratory following stringent

precautions to avoid contamination. Two mitochondrial protein coding genes, the second and third subunits of Nicotinamide Adenine Dinucleotide Dehydrogenase (ND2 and ND3, respectively) were PCR-amplified and sequenced. Primers used to amplify ND3 also amplified the adjacent flanking tRNA. PCR-amplification conditions follow the protocol in Fuchs *et al.* (2004) for ND2 and Bowie *et al.* 2003 for ND3. However, longer annealing times (45-60s) with an increased number of amplification cycles (40-45) were employed for DNA extracted from museum skins despite the smaller fragment sizes (250-350bp). Several internal primers (Table 5.3) were developed to amplify and sequence genomic DNA from museum skins. PCR products were electrophoresed on 1.5% agarose gels, stained with ethidium bromide and visualized under UV light. Amplicons of the appropriate length were cut out of the gel and purified using standard gel purification kits (E.g. QIAquick gel extraction kit). The purified products were cycle-sequenced using Big Dye terminator chemistry (Applied Biosystems, Inc [ABI]), precipitated with 3M ammonium acetate or 100% isopropanol, rinsed in ethanol, dried and resuspended in formamide-EDTA solution, and run on either AB3100 or AB3700 automated DNA sequencer. Sequences were obtained from both strands of DNA for each individual, and some individuals were sequenced several times if any base ambiguity was encountered. All sequences were checked using the program Sequencher 4.7 (Gene Codes Corp) to test for the presence of any insertions or deletions, as well as to check that no stop-codons were present.

Genetic diversity and phylogenetic reconstruction

Standard molecular diversity indices for the combined two-gene mitochondrial DNA dataset [haplotype diversity (h), nucleotide diversity (π), nucleotide frequencies, transition/transversion ratio, number of polymorphic sites] were calculated using Arlequin 3.5 (Excoffier *et al.* 2005). Pairwise genetic distances were calculated across all taxa using uncorrected-p distances in PAUP* (Swofford 2002). Phylogenetic trees were estimated using the model-based methods of maximum likelihood (ML) and Bayesian inference (BI). Maximum likelihood analyses used the RAxML-VI-

HPC algorithm (Stamatakis 2006; Stamatakis *et al.* 2008); with computations performed on the computer cluster of the Cyber-Infrastructure for Phylogenetic Research project (CIPRES, www.phylo.org) at the San Diego Super computing Center. In the RAxML analysis the mtDNA data was partitioned in to first, second and third codon positions. Bayesian analyses were conducted in MrBayes version 3.1 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003) using one cold and seven incrementally heated Markov Monte Carlo chains (MCMC) in two simultaneous runs. Chains were run for 1 million cycles, with trees sampled every 100th generation, with a random tree as the starting point and a temperature parameter value of 0.2. The first 20% trees of each run were discarded as the burn-in; converging log-likelihoods, potential scale reduction factors for each parameter, and inspection of tabulated model parameters suggested that stationarity had been reached thereafter. The remaining trees were used to compute posterior probabilities of nodes. Models or Bayesian analyses were established using MrModeltest 2.1 (Nylander, 2004) employing the Akaike information criterion (AIC). The general time-reversible (GTR) model plus a gamma shape parameter (Γ)) and proportion of invariable sites (I) received the best AIC score in MrModeltest, this was also the model used in the ML analyses implemented in RaxML. Clade support was assessed with posterior probabilities (PP) computed with MrBayes and 100 nonparametric bootstrap pseudoreplicates (Felsenstein 1985) as implemented in RAxML.

Population structure, networks and historical demography

A network of haplotypes was constructed with a statistical parsimony (Clement *et al.* 2000). In addition, a neighbor-net (NN) splits graph (Bryant & Moulton 2002), implemented in SplitsTree v4.1 (Huson & Bryant 2006), was constructed to infer the distribution of incompatible subgroups among haplotypes. Hierarchical analysis of molecular variance (AMOVA), as executed in Arlequin 3.5, was used to determine how contemporary genetic variation is partitioned over the range of the Yellow-streaked greenbul. Two competing groupings were analysed: the first based on colour form (Albertine Rift grey-headed, Malawi/Tanzania brown-headed, and coastal forest grey–headed) and

the second based on subspecies distribution (Table 5.1), with the exception of the single *P .f. kungwensis* sample and the three *P. f. graueri* individuals, which were excluded due to small sample size. In the analysis of colour forms, the grey-headed *P. f. kungwensis* individual was included in the grey-headed coastal forest clade, as suggested by the phylogenetic analyses (see Results), despite the geographical proximity to the Albertine Rift grey-headed forms (Figure 5.1).

I also used Arlequin (Schneider *et al.* 2000) to conduct the Tajima's D and Fu's Fs tests of selective neutrality (1000 replicates), as well as to estimate mismatch distributions for the major clades recovered in the maximum likelihood analysis. Demographic changes were also examined by calculating the raggedness index of the observed mismatch distribution for each of the mtDNA terminal lineages according to the population expansion model implemented in Arlequin (Schneider *et al.* 2000). Due to limited number of individuals included I excluded subspecies *graueri* and *kungwensis* from the analysis. Small raggedness values represent a population that has experienced sudden expansion whereas higher values of the raggedness index suggest stationary or bottlenecked populations (Harpending *et al.*1993; Harpending 1994). Further, the demographic history of whether Yellow-streaked greenbul mtDNA main lineages underwent a range expansion was investigated by the spatial expansion model in Arlequin (Schneider & Excoffier 1999; Excoffier 2004; Excoffier *et al.* 2005). We employed parametric bootstrapping (1000 replicates) as implemented in Arlequin to test the goodness-of-fit of the observed mismatch distribution to that expected under the spatial expansion model using the sum of squared deviations (SSD) statistic for regional and pooled samples. A significant SSD value is taken as evidence of departure from the estimated demographic model of spatial expansion in subdivided populations (Schneider & Excoffier 1999; Excoffier 2004) using Arlequin (Excoffie *et al.* 2005].

Estimation of divergence times

I used Beast v1.4.6 (Drummond *et al.* 2002, 2006; Drummond & Rambaut 2007), to estimate divergence dates within *P. flavostriatus* using mtDNA sequence data. I assigned the best-fitting model, as estimated by MrModeltest2 to each of the partitions. I used two alternate mutation rates, the classical rate of $1.05*10^{-8}$ substitutions/site/year (s/s/y) for mtDNA (2.1% per million years, Weir & Schluter 2008), as well as the faster rate of $3.05*10^{-8}$ (6.1% per million years, Arbogast *et al.* 2006, see also Warren *et al.* 2003), which translate into per locus (mtDNA: ND2 & ND3 combined) rates of $1.82*10^{-5}$ and $5.29*10^{-5}$ substitutions per year, respectively. This extrapolation carries with it a significant margin of error and thus I emphasize the importance of thinking of the time estimates only as a rough attempt to place diversification events of *P. flavostriatus* in a historical context. I assumed a Yule Speciation Process for the tree prior and an uncorrelated lognormal distribution for the molecular clock model (Drummond *et al.* 2006; Ho 2007). I used default prior distributions for all other parameters and ran MCMC chains for 50 million generations. The analysis was repeated twice. I used the program Tracer (Rambaut & Drummond 2007) to assess convergence diagnostics.

5.3: Results

Phylogenetic relationships

The mitochondrial DNA matrix contained 112 ingroup haplotypes from 248 individuals sampled and 975 bp (ND2=593 bp and ND3=382 bp) of which 213 (22%) sites were parsimony informative (Table 5.4). Both ML and BI analyses of the combined mtDNA data set recovered a topology of several with strong geographic structuring (Figure 5.2). Examination of the deepest nodes in the ML and BI phylogenetic trees identifies three broad clades that correspond to the Albertine Rift grey-headed, Malawi–Tanzanian Brown-headed, and coastal grey-headed forms, in accordance with the leapfrog pattern of plumage variation. However the nodel support for two of

the three (brown-headed and coastal grey-headed forms) basal nodes was poor in both ML and BI analysis. The Albertine Rift clade was well-resolved and received very strong support in both the ML and BI analysis. Within each major clade (basal nodes) several well-supported terminal clusters corresponding to the currently recognized subspecific taxonomy were also recovered (Figure 5.2).

The Albertine Rift clade included all the samples from the subspecies *olivaceogriseus* (including *itombwensis)* and *graueri.* In this clade both terminal and basal nodes were well-resolved. Interestingly, the molecular analyses *graueri* rendered *olivaceogriseus* paraphyletic with one clade containing birds from the DRC (Itombwe & Kabobo), and the second birds from east of the rift valley in Burundi and Uganda. Within the brown-headed *alfredi* clade two well-supported terminal clusters representing Malawian and Tanzanian populations were recovered. These two terminal brown-headed clusters were geographically separated by the Rukwa Rift, which is a part of the East African Rift System.

The widely distributed coastal grey-headed cluster also shows remarkable levels of genetic structure. The only sample of *kungwensis* included, forms the basal lineage of this cluster with high nodel support. The population from the Udzunguwa Mountains in the Eastern Arc Highlands (*uzungwensis*) forms the next distinct terminal cluster with strong nodal support. All the *tenuirostris* individuals included from different coastal localities and the northern Eastem Arc form a single terminal cluster; however the support for this node was poor. My analysis failed to recover subspecies *flavostriatus* and *vincenti* as distinct terminal clusters. Some individuals of *flavostriatus* from Ngoye Forest (South Africa) show a sister relationship to some *vincenti* individuals from southern Malawi and northern Mozambique. However the nominate subspecies from the Zimbabwe highlands was recovered as a distinct terminal clade.

Population structure and demography

Sequence divergence values estimated using uncorrected-p distances among the three basal clades were: 4.9 +- 0.2% between the brown-headed and coastal clades, 4.5 +- 0.2% between brown-headed and Albertine Rift, and 4.1 +- 0.01 % between the coastal and Albertine Rift clades, respectively. AMOVA for the three main basal clades (brown-headed and two grey-headed forms) indicated significant structuring of genetic variability among clades (Fst = 0.72, df = 2, $P < 0.001$) rather than within clades. Partitioning individuals into the subspecies groups, yielded a higher among subspecies (Fst = 0.86, d.f. = 5, $P < 0.001$) and lower within subspecies (14%) distribution of genetic variability suggesting that grouping the individuals by subspecies had a better fit to the data than lumping them purely by three basal mtDNA clades corresponding to two color forms.

Phylogenetic networks of the concatenated mtDNA using uncorrected-p genetic distances, and the NeighborNet algorithm recovered the same relationships among mtDNA clades as in the RAxML and Bayesian analysis (Figure 5.3). The three geographic clades correspond to two color forms and within each distinctive subspecies/terminal clusters were recovered with similar levels of bootstrap support as in the RAxML analysis. The statistical parsimony network resulted in several unlinked networks corresponding to the subspecies or geographic locations of samples (Figure 5.4) and due to the 95% probability constraint of joining haplotypes this method was unable to recover the three basal clades as in the ML and BI analysis of the combined mtDNA data. All three clades have significantly negative Tajima's D and Fu's Fs statistics (Table 5.4) and mismatch profiles follow a multi-modal distribution (results not shown here). However the mismatch profile for at least four(A,B,C, and D in Figure 5.5) of the main mtDNA lineages shows signature of expansion based on smaller raggedness index value (results not presented) and significantly negative Fu's Fs values (Table 5.4) where as the others shows multimodal distribution patterns. According to the non-significant values of the sum of squared deviations observed a fit for spatial range expansion was

found for five (A,B,C,D and H) terminal clades.

Divergence times

Divergence date estimates obtained using a molecular clock hypothesis for the two (6.1% and 2.1%) clock rates implemented in BEAST suggest that the first deep split within the Yellow-streaked greenbul complex occurred in the Plio-Pleistocene (Table 5.5). The lineages from the Albertine Rift and Mt. Mahale versus all the lineages from the rest of the distribution range diverged between 1.47 (95% HPD: 1.21-1.77) and 4.47 (95% HPD: 2.49-6.90) Ma for the 6.1% and 2% clock rates respectively. The 95% HPD interval does not overlap between these two alternate rates of divergence for the first deep split. However for some of other comparisons made there was a certain degree of overlap between the 95% HPD intervals for the two clock rates used (Table 5.5) particularly within the brown-headed *P. f. alfredi* clade. The brown-headed *P. f. alfredi* diverged from the Eastern Arc and east African coastal forest grey-headed forms between 1.07 (95% HPD: 0.85-1.31) and 3.26 (95% HPD: 1.82-5.07) Ma suggesting possible involvement of Plio-Pliestocine climatic changes for driving this divergence.

5.4: Discussion

Genetic variation and taxonomic implications

The mtDNA evidence suggests that the Yellow-streaked greenbul complex is comprised of at least nine highly differentiated geographic populations (Figure 5.2). The recovered three deep lineages are consistent with there being two grey–headed and brown–headed groupings of the species complex indicating an association between the evolution of head colour and genetic structure based on mtDNA. Further, the reconstructed mtDNA phylogeny strongly supports the validity of current subspecific level phylogeny for most of the populations studied. The results of AMOVA for the subspecies data partition also strongly support the notion of grouping individuals by subspecies. This geographic separation of disjunct or parapatric populations, was also corroborated by the observed pairwise genetic distance (5-4%) and the statistical parsimony

networks (Figure 5.4). There is evidence of recent population expansion in this data set for *uzungwensis, vincenti, olivaceogriseus*(Burundi) and *olivaceogriseus*(DRC*)* terminal lineages recovered as seen with the unimodel distribution of mismatch profile (Figure 5.5) and significant Fu's Fs values (Table 5.4). The hypothesis of historical range expansion could not be rejected for same above four terminal clusters and *alfredi* from Malawi and Zambia.

In the reconstructed phylogeny the Albertine Rift lineages received the greatest support in both the ML and Bayesian analysis. All three lineages recovered here are reciprocally monophyletic. The presense of the two highly divergent mtDNA lineages of the subspecies *P. f. olivaceogriseus* is inconsistent with the current subspecies level phylogeny in this region, but otherwise the pattern of mtDNA lineage diversity is congruent with the geographic separation among disjunct or parapatric populations (Figure 5.2). The well-supported lineage of *P. f. olivaceogresius* from South Kivu and Itombwe Forest DRC may indicate that the subspecies *P. f. itombwensis* described by Prigogine (1975) is a valid taxon. However, thorough reassessment of morphology and if possible voice/behaviour, is required to confirm this. All three Albertine Rift lineages show patterns of population expansion, but my study did not find any evidence of mitochondrial introgression.

The brown-headed Malawi-Zambia-Tanzania clade was recovered as a distinct lineage in both the ML and Bayesian analyses, but the support for the node was poor. This poses some uncertainty with respect to the species level taxonomic status of *Phyllastrephus alfredi* (Sharpe's greenbul; Shelly 1903). The genetic differentiation across the Rukwa-Malawi Rift was evident with the recovery of two well-supported lineages of brown-headed forms without any evidence of mitochondrial introgression (no shared haplotypes).

The most complex phylogenetic structure was observed within the Eastern Arc, East African coastal forest and southern grey-headed forms, thus reaffirming the complex biogeography of the region. The subspecies *P. f. kungwensis* from Mt. Mahale was recovered *as* the basal lineage of

this clade and a sister taxon to the Albertine Rift lineages. The subspecies *P. f. uzungwensis* was recovered as strongly supported genetically distinct lineage. Both *kungwensis* and *uzungwensis* were reciprocally monophyletic and the demographic analysis of the *uzungwensis* revealed population expansion. However, the subspecies, *P. f. vincenti* and the nominal subspecies were not recovered as monophyletic (Figures 5.2) lineages. Within the nominal subspecies a discordance was observed between montane and coastal populations. The nominal subspecies from the montane Zimbabwe highlands was recovered as a strongly supported monophyletic lineage. Haplotypes in this lineage show deep structuring (long mutational steps in SPN network; Fig.5.4) and a multimodel mismatch profile (Figure 5.5) perhaps indicative of retained ancestral polymorphism from long-term persistence. All the individuals from South Africa (Ngoye Forest) were paraphyletic with respect to some individuals of *P. f. vincenti* from Mt. Mulanje in Malawi (Figures 5.2, 5.3 and 5.4). Fu's Fs and the mismatch profile indicate that the Mt. Mulanje population is expanding. This sister relationship of individuals from Ngoye Forest (the nominal subspecies) to some individuals of *P. f. vincenti* on Mt. Mulanje either indicate retained ancestral polymorphism or gene flow via patchily distributed intervening coastal forest. However there were no shared haplotypes between the two subspecies concerned, and this makes hybridization unlikely. In a previous study on the White-starred Robin *Pogonochchila stellata* Bowie *et al.* (2006) suggested that coastal forest may provide a vital habitat configuration for gene flow between montane populations of this species. The Eastern Arc and east African coastal forest populations *(P. f. tenuirostris)* of Yellow-streaked greenbul were recovered as a single clade in both ML and Bayesian analysis with poor nodal support. In the SPN network (Figure 5.4) Ngoye Forest nominal subspecies and all the *P. f. vincenti* were nested within the highly structured *P. f. tenuirostris* subnetwork. *P. f. tenurostris* consists of a series of disjunct populations which appear to be expanding and there is evidence in my analysis for introgression between montane and coastal forest populations, thus reaffirming Bowie *et al.* (2006) hypothesis of gene flow among montane

populations via East African coastal forests, at least for some forest birds.

Biogeography

The estimates of divergence times suggest that diversification of the Yellow-streaked greenbul occurred in the Plio-Pleistocene period within the past 4.5 to 1.5 Mya according to the 2% and 6.1% clock estimates, respectively. It is well known that the Plio-Plestocine period was characterized by high amplitude glacial cycles and particularly during the last 2.5 Ma there have been about 20 such glacial cycles of major forest expansion and retreat in Africa (deMenocal 2004). This further implies that the diversification is primarily due to climatic perturbation rather than the major geological events associated with the EARS. This is primary because the initial divergences of the three deep basal clades post-date the final Miocene uplift of the EARS at c. 7 Ma (Griffith, 1993) for both clock rates used. All estimates done with 6.1% clock rate had occurred exclusively during the Pleistocene and the youngest split had occurred around 60 Ka. This suggests that the underlying biogeography pattern may be explained by the popular Pleistocene forest refuge hypothesis (Diamond & Hamilton, 1980; Crowe & Crowe, 1982; Mayr & O'Hara, 1986). This is in agreement with the three putative refuges described for central Africa (western flank of the Albertine Rift), Eastern Arc sky islands and the southern refuge (Burgess *et al*. 1998; Fjeldså & Lovett 1997;Lawes *et al.* 2001). However, for most African forest bird taxa it has proven that the Pleistocene forest refuge hypothesis is not the best explanation (Bowie *et al.* 2004a,b, 2006; Fjeldså & Bowie 2008; Njabo *et al.* 2008). In contrast the divergence time estimates with classical clock rate (2%) suggest that Pliocene climate change was a major driver of speciation in the Yellow-streaked greenbul for the deep lineages. However, even with the 2% clock rate most of the within clade radiations has taken place during the Pleistocene. This further supports the postulated Pleistocene forest refugia as a primary hypothesis to explain lineage diversification within the Yellow-streaked greenbul complex.

Estimating divergence dates with molecular data remains controversial for songbirds, where few fossil data exist (Voelker *et al.* 2010). There is compelling evidence to suggest that the 2.0%

rate is not as general as previously advocated due to several factors (e.g. metabolic rate, generation time) that may affect substitution rates (e.g. Warren *et al.* 2003, Arbogast *et al.* 2006, Pereira & Baker 2006). Thus the 6.1% rate may be more reliable than the traditional rate, at least at the time scale concerning incipient species of divergence among members of recently formed species complexes such as the Yellow-streaked greenbul complex studied here.

Looking at the broader geographical scale it seems that the Pleistocene forest refuge alone may not well explain the pattern of observed genetic differentiation and lineage evolution. Perhaps geographically restricted mtDNA lineages might have their own complicated evolutionary histories, which probably resulted primarily from the Pleistocene cyclic climatic scenario, facilitated by isolation across already existing barriers and subsequent recolonization due to habitat reconnection. Further, the topology of the mtDNA phylogeny and strongly supported geographically disjunct populations suggest the possibility of a stepping-stone fashion of diversification from Albertine Rift. The direction of colonization during the window of climatic stability is difficult to speculate as two of the deep splits did not received sufficient nodal support in both ML and Bayesian inference analyses. It is possible that the complex originated or the common ancestor was distributed in the Albertine Rift which is the region with the highest vertebrate species richness in Africa (Brooks *et al.* 2001; Plumptre *et al.* 2007), and many avian species are either endemic to, or reach a distributional limit here (Sinclair & Ryan 2003). For the stepping-stone model two possible routes for the colonization of adjacent areas from the Albertine Rift refuge can be identified; first a north-east route via the Kenyan highland to the eastern Arc, and a second route south-east to the Eastern Arc and Malawi Rift via Mt. Mahale. As stated earlier the exact direction is difficult to predict. Yet by being absent from the Kenyan highlands and existence of the grey-headed population on Mt. Mahale, it does seem to suggest that the south–east dispersal route may be more likely. This is also in agreement with my divergence time estimates (Table 5.5). Three reciprocally monophyletic lineages recovered in the Albertine Rift provide further evidence of long-

term persistence in this region. There is a clear geographic separation of the lineages in the left and right rift valley regions (east-west break). All individuals included left of the Rift valley in the highlands NW of L. Tanganyika (Itombwe massif) and Mt. Kabobo in DRC formed the basal lineage of the entire complex. Estimated divergence times for the three lineages coincide with the upper Pleistocene period where the isolation can be attributed to the high amplitude glacial cycles. These isolated lineages appear to show recurrent patterns of population expansion.

In the Eastern Arc and East African coastal forest there is a clear evidence for lowland and montane interaction (Figures 5.2 and 5.3). The pattern observed in east Africa is somewhat different to the more common patterns where adjacent mountains contains sister taxa (Burgess *et al.*1998; Fjeldså, 1994; Fjeldså & Lovett, 1997; Roy *et al.* 2001; Voelker et al. 2010). The occurrence of the same subspecies in adjacent mountains (historical sky islands) without much genetic structuring suggest that recurrent gene flow and multiple recolonization events among these heterogeneous habitats, as is evident in some other studies (Outlaw et al., 2007; Outlaw, Voelker & Bowie, unpublished - Muscicapidae). Pleistocene forest refuge can explain the genetic structuring observed in southern coastal and montane caldes. In Zimbabwe highlands there is evidence for a long term stable population and the comparatively younger South African coastal forest would have been colonized from this Zimbabwe highland refuge. Yet there is no evidence to reject the possibility of south African coastal forest from ancient east African coastal forest refugia (Burgess *et al.*1998). This pattern has been observed for other taxa as well (see Laws 2007 for details). Parapatric taxa (*flavostriatus/vincenti*) occurring on Mt. Mulanje, Malawi can be attributed to the northward migration of coastal nominal subspecies along the southeastern coast of Africa. Similarly the reciprocal monophyly of the isolated population in Udzungwa Mt. and the escarpment and adjacent areas can be explained with Pleistocene forest refuge hypothesis. However, rapid fixing of characters within a shorter duration in response to local conditions unique to the region might have accelerated lineage sorting and hence the reciprocal monophyly. *P. f. alfredi* in

Zambia, north and central Malawi and Mbizi Forest in Tanzania seem to represent a derived state, probably from the adjacent Mt. Mahale or from the Uzungwa Mt. populations. Further, the possibility of deriving the brown-headed morphotype from the southern Mt. Mulanje population cannot be convincingly rejected owing to the similarities in egg morphology and vocalization in both *P. f. alfredi* and *P. f. vincenti*. Intriguingly, the genetic differentiation observed across the Rukwa-Malawi Rift is in accordance with that observed for *P. cerviniventris* (chapter 3).

Evolution of a leapfrog pattern of colour morphology

The recovered three deep mtDNA lineages correspond to the two grey–headed and the brown–headed populations of the species complex indicating the association between the evolution of colour morphology and phylogenetic affinity. Brown-headed plumage seems to be a derived state from the grey-headed plumage and radiations would have occurred only within the past 290,000 years, corresponding to the last of two or three major Pleistocene glaciations (deMenocal, 1995). The brown-headed lineages from Zambia, Northern Malawi and Mbizi Forest, Tanzania, were recovered in both phylogenetic reconstructions and networks as distinct clades in an intermediate position, although with varying degrees of support. There were 12 fixed base substitutions in the mtDNA dataset in brown-headed *alfredi* when compared with all the grey-headed lineages. These further prove that the origin of the leap-frog pattern of color morphology can be a consequence of evolutionary change in the intermediate population. It is likely that similarities among disjunct grey-headed populations reflect retained ancestral character states and character states of intervening brown-headed populations are more recently derived(Cadena *et al*. 2010).

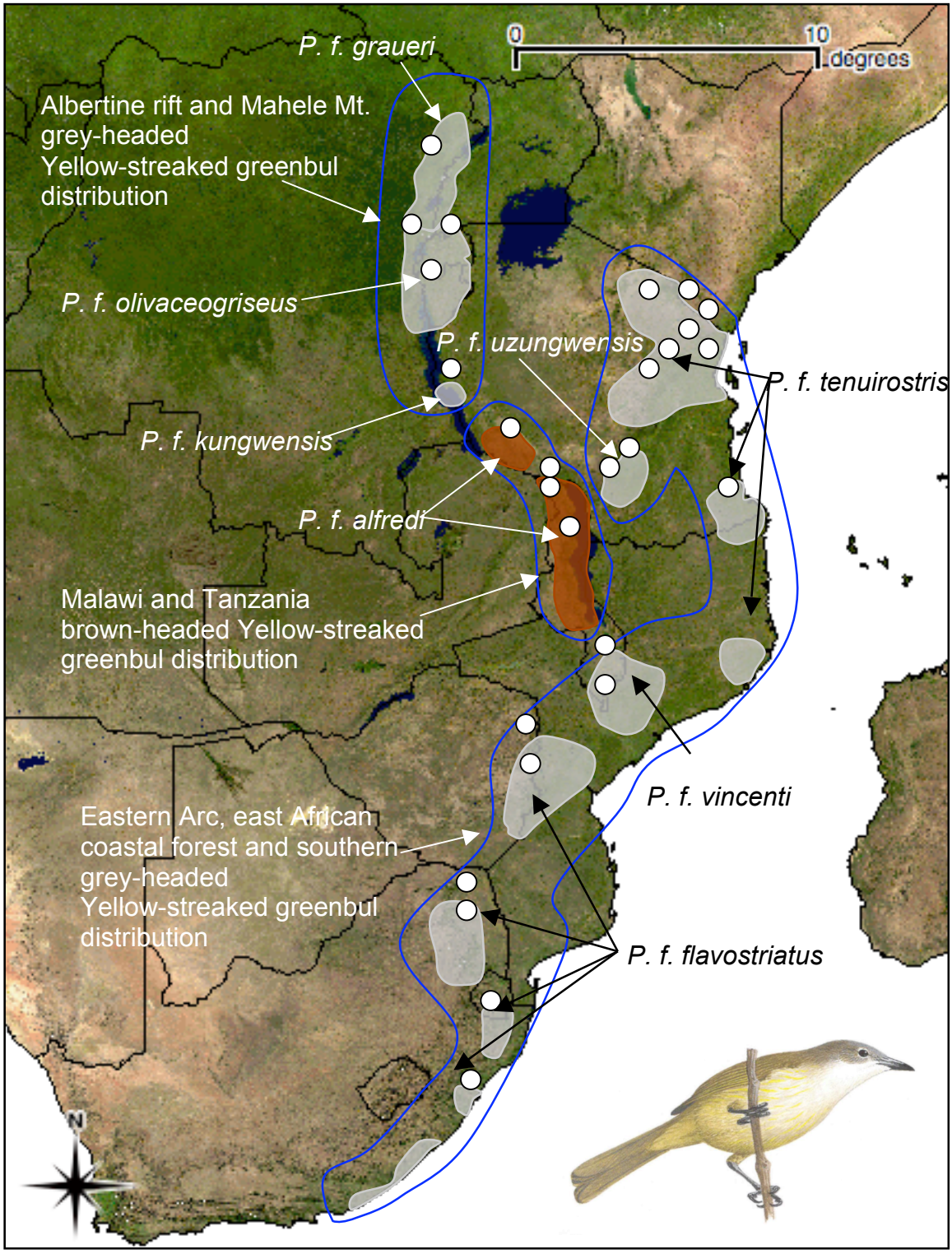


Figure 5.1 - Distribution range of *Phyllastrephus flavostriatus* complex in southern and eastern Africa. Distribution pattern of two colour forms are shown in the area demarcated by blue lines and the current distribution of each putative subspecies is shown in grey and brown highlighted areas in the map. White dots indicate the sample localities. Picture of grey-headed Yellow-streaked greenbul is modified from Keith (1992).

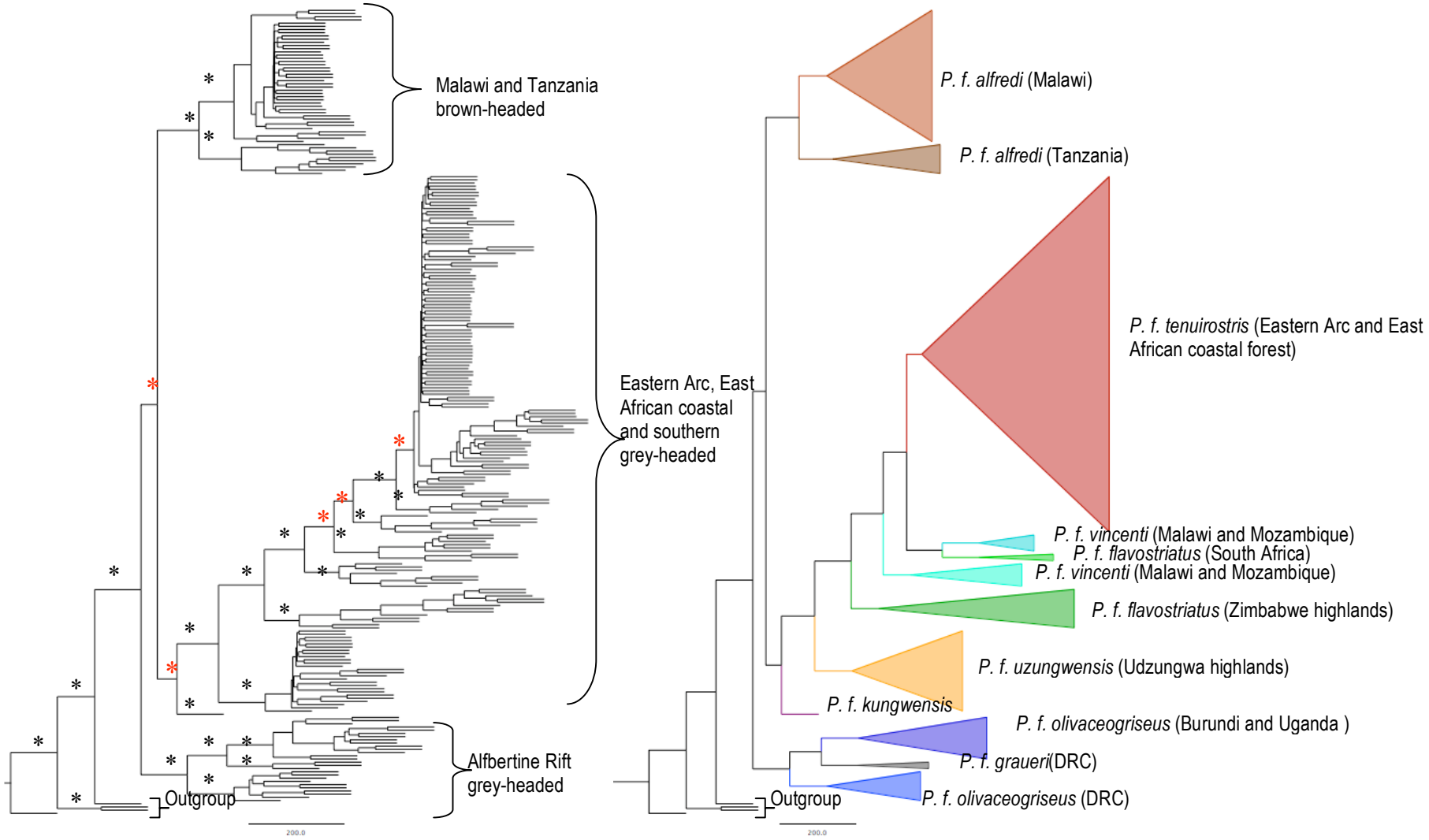


Figure 5.2 - Majority rule consensus tree calculated from 100 bootstrap trees generated using RAXML (Stamatakis 2006; Stamataskis *et al.* 2008). * indicates the nodes which received more than 75% bootstrap support and * indicates nodes which did not receive considerable support. The tree on the right indicates the phylogenetic relationship of the brown headed and two grey-headed Yellow-streaked greenbul (three basal nodes) clades and the tree on the left depicts the subspecies/populations recovered as distinct terminal clusters in the analysis.

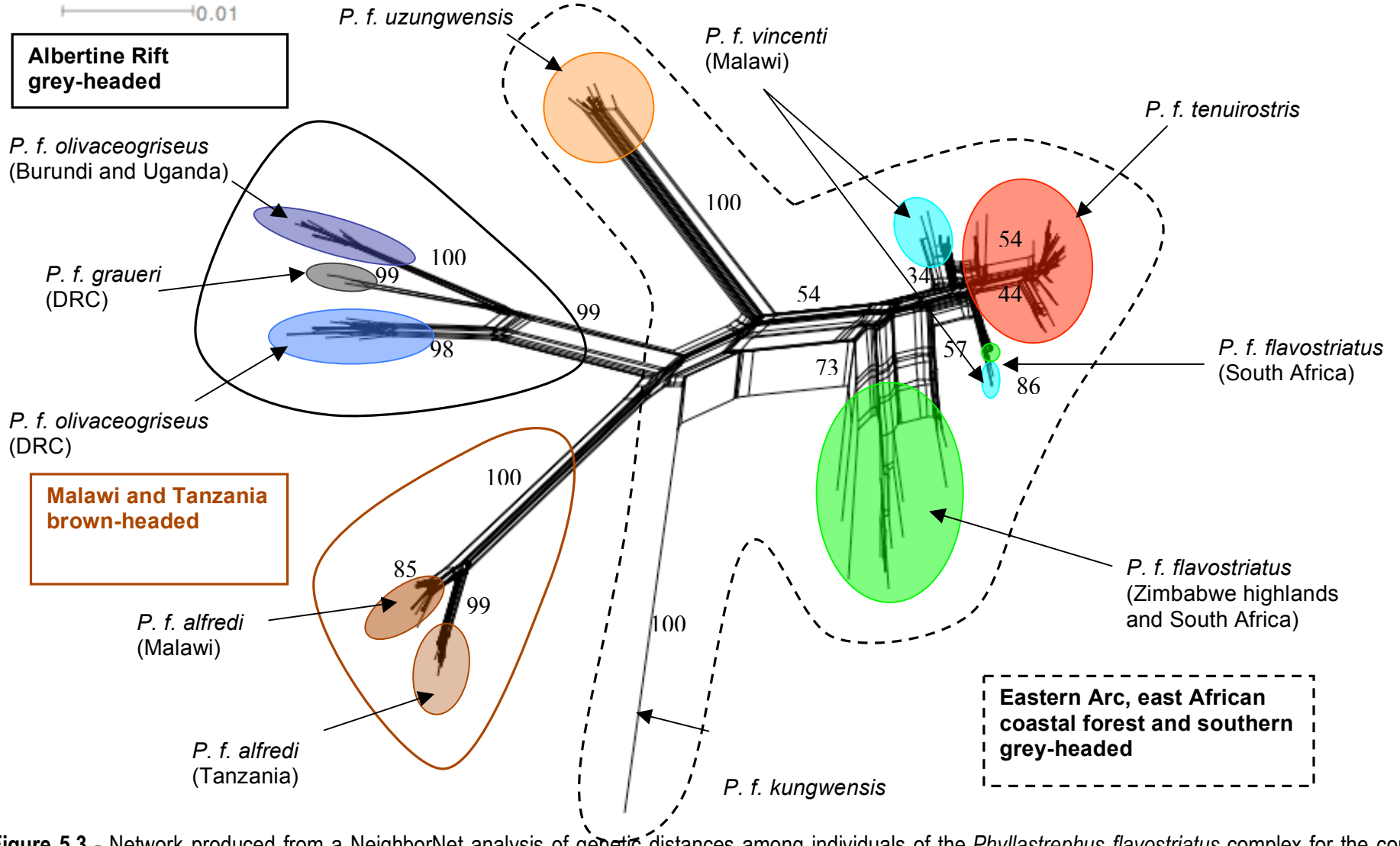


Figure 5.3 - Network produced from a NeighborNet analysis of genetic distances among individuals of the *Phyllastrephus flavostriatus* complex for the combined mtDNA data (n=248). Terminal clusters are coloured to represent different subspecies/populations. Bands of parallel edges represent incompatible splits and the value at each major split indicates the bootstrap support.

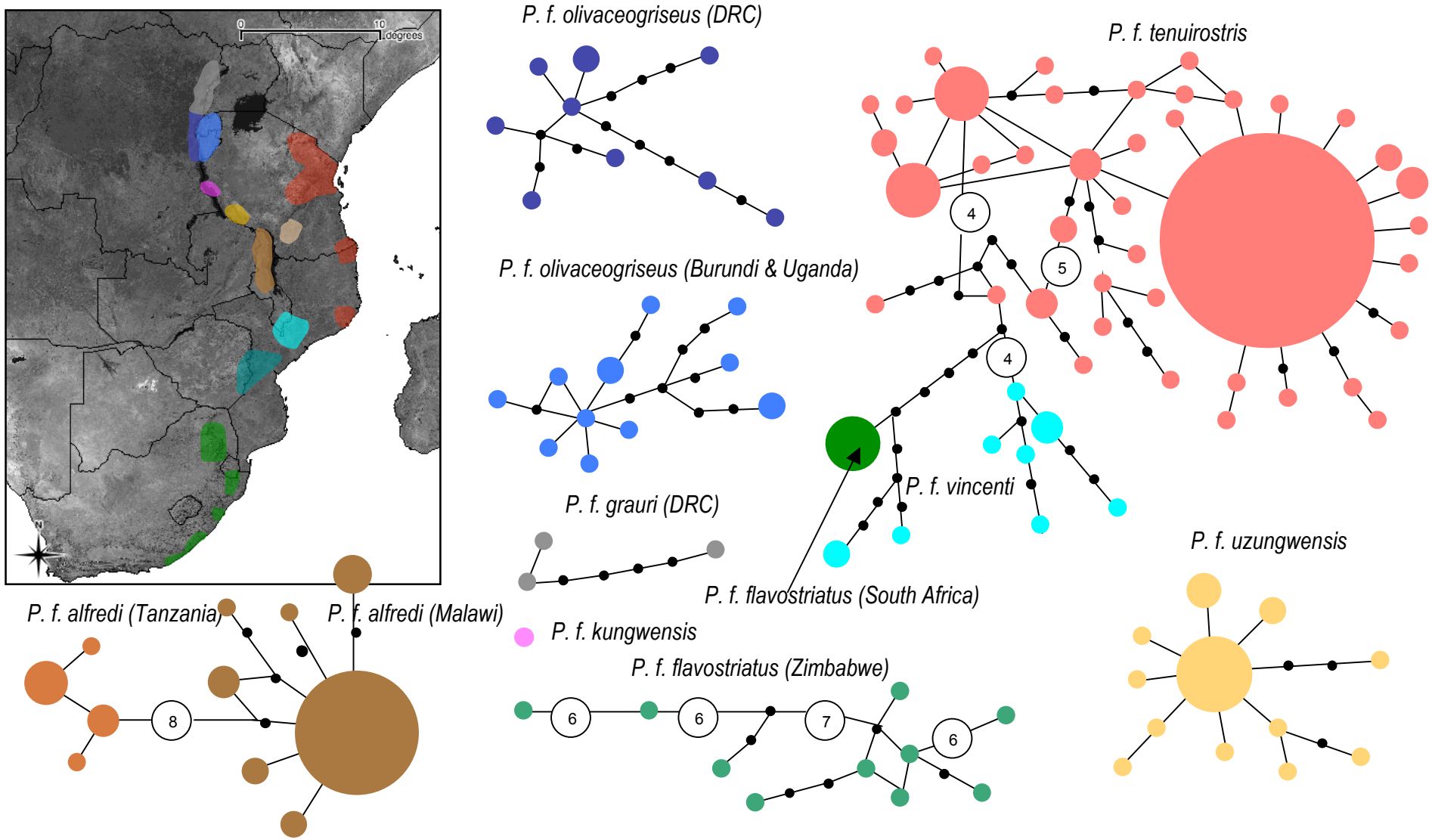


Figure 5.4 - Unrooted statistical parsimony network constructed from the combined mtDNA dataset, which resulted in eight unlinked haplotype clusters. Circle diameter is proportional to the number of individuals possessing this haplotype. Extinct or unsampled haplotypes are indicated by black dots. Numbers within the circled areas indicate the number of mutational steps connecting the adjacent haplotypes.

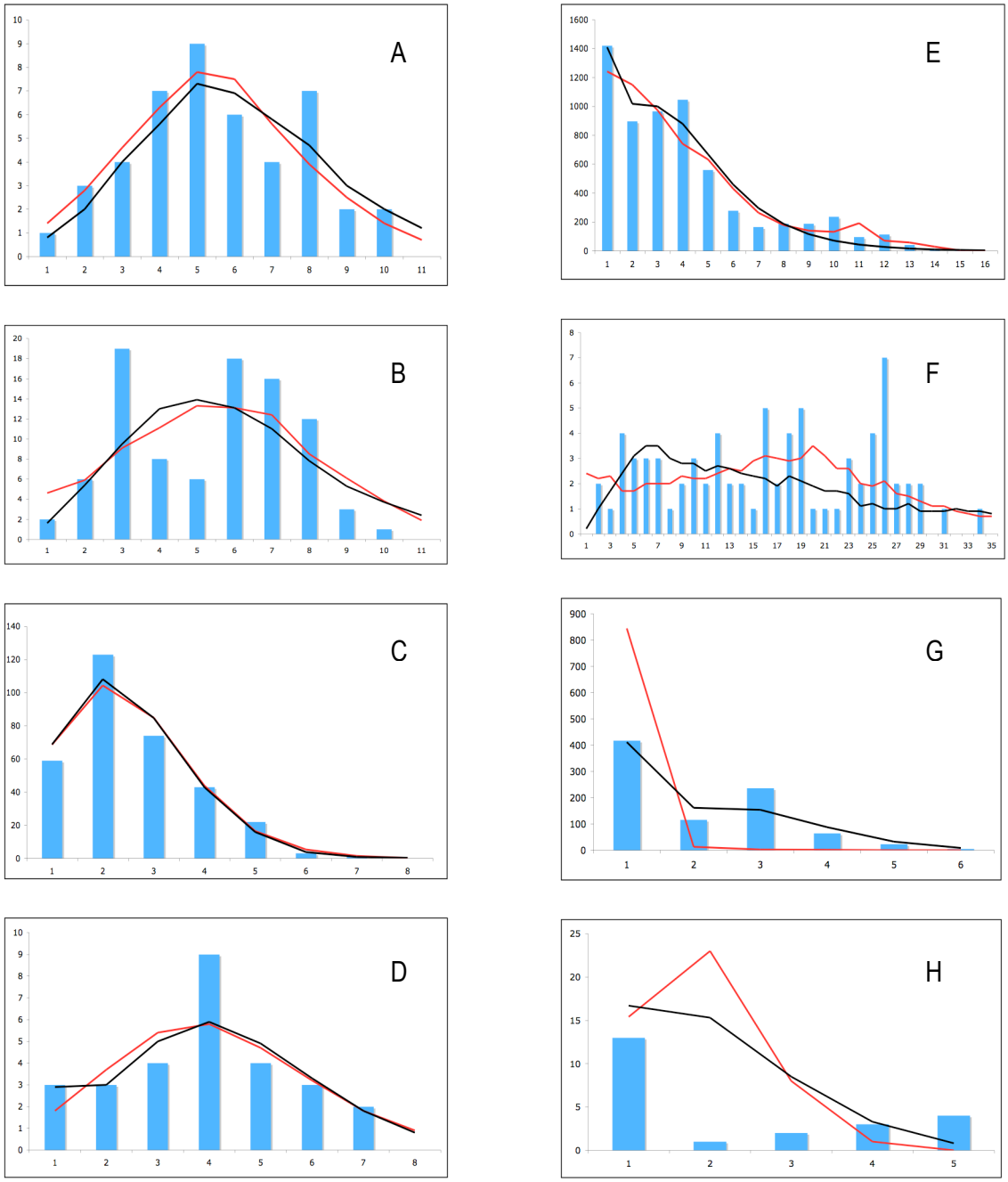


Figure 5.5 - Mismatch distributions for the main mtDNA lineages of the Yellow-streaked greenbul. Coloured bars in the graphs represent the observed data and red lines represent the simulated data used to test the goodness-of-fit to the sudden expansion model. Black lines represent the simulated data used to test the goodness-of-fit to the spatial expansion model. The x-axis represents the number of base pair differences and the y-axis represent the number of pairwise comparisons made. A – lineage *P. f. olivaceogrisesus* from DRC, B – lineage *P. f. olivaceogrisesus* from Burundi and Uganda, C- *P. f. uzungwensis*, D- *P. f. vincenti* excluding the three individuals which showed sister relationship with nominal spp from Ngoye Forest South Africa. E- *P. f. tenuirostris*, F- nominal spp from Zimbabwe, G- lineage *P. f. alfredi* from Tanzania, and H-lineage *P. f. alfredi* from Malawi and Zambia.

Table 5.1 - Morphological variation and geographical distribution of the Yellow-streaked greenbul complex (*Phyllastrephus flavostriatus)* (Keith 1992; Hockey *et al.* 2005)

Subspecies	*Colour morphology*	*Distribution*
*Phyllastrephus flavostriatus flavostriatus**	Grey-headed	South Africa, Zimbabwe, lowlands of Mozambique south of the Zambezi River.
Phyllastrephus flavostriatus vincenti	Grey-headed	Highlands of SE Malawi east of Shire River valley and adjacent N Mozambique.
Phyllastrephus flavostriatus tenuirostris	Grey-headed	Kenya and E Tanzania to the lowlands of Mozambique north of Zambezi River
Phyllastrephus flavostriatus uzungwensis	Grey-headed	Udzungwa Mts. central Tanzania.
Phyllastrephus flavostriatus alfredi	Brown-headed	SW Tanzania, Zambia, N Malawi
Phyllastrephus flavostriatus kungwensis	Grey-headed	Mt. Kungwe and the Mt. Mahale Mt,W Tanzania
*Phyllastrephus flavostriatus olivaceogriseus***	Grey-headed	Ruwenzori Mts., SW Uganda, W Rwanda and N Burundi, and left of rift valley in the highlands NW of L. Tanganyika (Itombwe) and Mt. Kabobo in DRC
Phyllastrephus flavostriatus graueri	Grey-headed	DRC in highlands west of lakes Albert, Edward and Kivu.

* include three morphological variants "*distanes*", "*dendrophilus*" and "*dryobates"* (Clancy 1962,1975), ** include two morphological variants "*itombwensis*" and "*ruwenzorii"* (Prigogine 1975)

Table 5.2 - List of sample extract numbers included in the study. See Appendix 1 for further details.

Subspecies	*Number of samples*	*DNA number in the analysis*
Phyllastrephus flavostriatus flavostriatus	21	LGS001, LGS004 to LGS005, LGS077, LGS646 to LGS647, LGS649 to LGS652, LGS666, LGS671, LGS679 to LGS681, LGS742, LGS744, LGS746, LGS749 to LGS750
Phyllastrephus flavostriatus vincenti	11	LGS309, LGS370, LGS412 to LGS419, LGS907
Phyllastrephus flavostriatus uzungwensis	25	LGS149 to LGS150, LGS166 to LGS167, LGS169, LGS177, LGS252, LGS271 to LGS272, LGS279, LGS282 to LGS283, LGS291 to LGS293, LGS708 to LGS716, CVV604
Phyllastrephus flavostriatus tenuirostris	85	LGS151 to LGS 152, LGS154, LGS159 to LGS165, LGS170, LGS175 to LGS176, LGS181 to LGS185, LGS191, LGS193, LGS253, LGS257 to LGS258, LGS260 to LGS261, LGS263, LGS266 to LGS269, LGS284 to LGS287, LGS290, LGS299 to LGS300, LGS302 to LGS306, LGS311 to312, LGS314 to LGS318, LGS320 to LGS324, LGS367, LGS691 to LGS707, CVV566, CVV577, CVV584, CVV586 to CVV587, CVV593, CV606 to CVV610, CVV619, CVV621
Phyllastrephus flavostriatus alfredi	52	LGS172 to LGS173, LGS222 to LGS229, LGS326 to LGS330, LGS332, LGS334 to LGS347, LGS349 to LGS355, LGS357 to LGS358, LGS674 to LGS678, LGS908, LGS910 to LGS911, LGS914 to LGS916, LGS921 to LGS922
Phyllastrephus flavostriatus kungwensis	01	LGS319
*Phyllastrephus flavostriatus olivaceogriseus**	24	LGS359 to LGS365, LGS682 to LGS684, LGS686 to LGS690, LGS735 to LGS736, LGS764 to LGS770
Phyllastrephus flavostriatus graueri	03	LGS366, LGS737, LGS738

*include subspecies *P. f. itombwensis*

Table 5.3 - Novel PCR primers developed for this study to amplify the mitochondrial ND2 and ND3 gene regions from museum skin samples.

Primer name	*Gene region*	*Primer location with in the gene*	*Primer sequence -5'-3'*
1. ND2PFintR1	*NADH subunit 2*	*332-352 bp*	CCT CTG GGA ATC AGA AGT GG
2. ND2PFintF2	*NADH subunit 2*	*316-334 bp*	CTA GGC CTA GCC CCA TTC
3. ND2PFintR2	*NADH subunit 2*	*781-805 bp*	CTT GGA TGA TRA GTC ATT TGA GGA G
4. ND2PFintF3	*NADH subunit 2*	*736-757 bp*	CTA ACC CTR CTC TCC CTA GC
5. ND2PFint4F	*NADH subunit 2*	*637-654 bp*	GCY GCC GTA TTC CTA ACC
6. ND3PFintF	*NADH subunit 3*	*112-130 bp*	ATC CCC ATA YGA ATG TGG C
7. ND3PRintF	*NADH subunit 3*	*144-160 bp*	GGG RAG TCG GGC RGA TC

Table 5.4 - Number of haplotypes, number of polymorphic sites, haplotype diversity (H), nucleotide diversity (π), Tajima's D and Fu's Fs statistics obtained when the combined mitochondrial data set is partitioned into the three major mtDNA clades (colour forms), as well as when partitioned according to subspecies. Samples from subspecies *P. f. kungwensis* and *P. f. graueri* were not included when data were portioned as subspecies. In the three mtDNA clade analyses *P. f. kungwensis* was included in the coastal grey-headed clade.

	Three primary mtDNA clades ("colour forms")			*Subspecies*					
	Eastern Arc, east African coastal and southern grey-headed	*Malawi and Tanzania brown-headed*	*Albertine rift gey-headed*	*P. f. flavostiatus*	*P. f. vincenti*	*P. f. tenuirostris*	*P. f. uzungwensis*	*P. f. alfredi*	*P. f. olivaceogriseus*
Number of individuals	169	52	27	19	11	113	26	52	24
Number of haplotypes	77	11	24	14	8	42	13	11	21
Number of polymorphic sites	156	22	69	65	20	88	15	22	57
H	0.8945 +/- 0.0213	0.6757 +/- 0.0694	0.9915 +/- 0.0125	0.9123 +/- 0.0605	0.9273 +/- 0.0665	0.7756 +/- 0.0423	0.8185 +/- 0.0733	0.6757 +/ - 0.0694	0.9891 +/ - 0.0152
π	0.0192 +/- 0.0950	0.0043 +/- 0.0024	0.0211 +/- 0.0107	0.0203 +/- 0.0105	0.0079 +/- 0.0045	0.0041 +/- 0.0022	0.0016 +/- 0.0011	0.0043 +/- 0.0024	0.0195 +/ - 0.0099
Tajima's D	-0.9959 p= 0.1580	-0.4025 p= 0.3530	-0.5780 p= 0.7960	0.2787 p=0.6770	0.5843 p=0.7590	-2.4538 p=0.000	-0.0882 p=0.0050	-0.40258 p= 0.35300	0.9677 p= 0.8920
Fu's Fs	-20.7701 p= 0.009	0.5454 p= 0.6570	-5.6873 +/- 0.0310	0.1134 p=0.4930	-0.3613 p=0.3900	-25.8418 p=0.0000	-9.6311 p=0.0000	0.5454 p= 0.6570	-4.3866 p=0.0440

Table 5.5 - Estimated time-to-most-recent-common-ancestry (TMRCA) of *P. flavostriatus* and the main lineages representing the subspecies under two alternate clock rates. Upper and lower 95% confidence levels are presented in parenthesis.

Lineage	*TMRCA estimated at 6.1% clock rate*	*TMRCA estimated at 2% clock rate*
Albertine Rift + *P. f. kungwensis* vs. *P. f. alfredi* + *P. f. uzungwensis* + *P. f. flavostriatus* + *P. f. tenuirostris* + *P. f. vincenti* (Albertine Rift and Mahale Mt. vs rest of the distribution range)	1.47 (1.21-1.77)	4.47 (2.49-6.90)
P. f. kungwensis vs. Albertine Rift (Mahle Mt. vs Albertine Rift)	1.45 (1.18-1.77)	4.43 (2.44-6.87)
All Albertine Rift	0.73 (0.55-0.91)	2.21 (1.18-3.46)
P. f. olivogreseus (DRC)	0.21 (0.13-0.29)	0.63 (0.29-1.05)
P. f. olivogrreseus (Burundi+Uganda)	0.33 (0.19-0.47)	1.01 (0.46-1.71)
P. f. graueri (DRC)	0.28 (0.16-0.42)	0.87 (0.36-1.46)
P. f. alfredi vs + *P. f. flavostriatus* + *P. f. tenuirostris* + *P. f. vincenti* + *P. f. uzungwensis*	1.07 (0.85-1.31)	3.26 (1.82-5.07)
P. f. uzungwensis vs *P. f. flavostriatus* + *P. f. tenuirostris* + *P. f. vincenti*	0.84 (0.64-1.04)	2.56 (1.43-4.07)
P. f. uzungwensis	0.11 (0.06-0.16)	0.32 (0.13-0.55)
All *P. f. flavostriatus* + *P. f. tenuirostris* + *P. f. vincenti*	0.55 (0.42-0.70)	1.65 (0.90-2.59)
P. f. flavostriatus	0.55 (0.42-0.70)	1.65 (0.90-2.59)
P. f. tenuirostris	0.29 (0.19-0.39)	0.87 (0.43-1.43)
P. f. vincenti	0.28 (0.17-0.38)	0.81 (0.40-1.36)
P. f. alfredi (Tanzania+Malawi)	0.30 (0.19-0.41)	0.91 (0.43-1.47)
P. f. alfredi (Malawi)	0.18 (0.12-0.25)	0.55 (0.25-0.89)
P. f. alfredi (Tanzania)	0.06 (0.02-0.10)	0.20 (0.06-0.36)

CHAPTER 6

A reassessment of species boundaries within the Yellow-streaked greenbul species complex (*Phyllastrephus flavostriatus*): a mulitlocus perspective and application of the species tree concept to reconstruct phylogeny

6.1: Introduction

A phylogenetic tree represents the evolutionary history of a set of taxa, and the use of "tree-thinking" has greatly facilitated the discussion and testing of multiple hypotheses (Huson & Bryant 2006). However, it is well-known that phylogenetic trees alone have limited power to describe more complex evolutionary scenarios, particularly among closely related species or at the intraspecific level. Traditional molecular phylogenetic analyses broadly consist of two steps: (1) obtaining and aligning molecular sequences and (2) inferring gene trees for those sequences. These gene trees have generally been considered to be synonymous with species trees. However, it is increasingly recognized that the evolutionary histories of individual genes may differ substantially from the underlying species tree (e.g., Nei 1986; Neigel & Avise 1986; Doyle 1992; Ruvolo 1994; Maddison 1997; Nichols 2001; Nordborg 2001; Degnan & Rosenberg 2006; Kubatko & Degnan 2007), with the potential for incongruence between gene trees and species trees resulting from several different process, including: horizontal gene transfer, gene duplication, and incomplete lineage sorting (deep coalescence) (Maddison 1997; Maddison & Knowles 2006).

The possibility that the majority of independent gene trees may be incongruent with the true underlying species tree under several circumstances (Degnan & Rosenberg 2006; Kubatko & Degnan 2007) is a cautionary warning against relying on a single locus for evolutionary analysis and species delimitation (Leache 2009). When phylogenetic trees representing the species history are of primary interest, it is therefore useful to gather multilocus data and not a single gene tree. Further, although several recent studies have claimed that the commonly used procedure of concatenating multi-gene data prior to phylogenetic analysis performs well (Chen & Li 2001; Rokas *et al*. 2003), others have highlighted situations in which such procedures fail (Carstens & Knowles 2007; Kolaczkowski & Thornton 2004; Kubatko & Degnan 2007; Mossel & Vigoda 2005). The concatenation approach is analogous to the total evidence philosophy in systematic biology advocated by Kluge (1989) and may offer the advantage of revealing a predominant (or underlying)

species phylogeny despite the presence of conflict in the data (e.g. Rokas *et al.* 2003).

With the advent of large-scale comparative genomic data sets and enhanced computational power, statistical methods such as maximum likelihood and Bayesian inference have provided sophisticated approaches to incorporation of heterogeneous models of sequence evolution often termed mixed-model analyses in phylogenetic analysis of multigene concatenated (i.e total evidence) data sets (Edwards *et al.* 2007). However, the stochastic process of lineage sorting can cause discordance among gene trees inferred from independent loci, which may provide inaccurate estimations of the species tree in separate and combined phylogenetic analyses (Pamilo & Nei 1988; Maddison 1997; Edwards *et al.* 2007; Rannala & Yang 2008). New multilocus phylogenetic methods are emerging to reconstruct species trees from independent loci (Edwards 2009), and some approaches are integrating population genetics and phylogenetic methods to explicitly accommodate the process of lineage sorting (Liu & Pearl 2007; Liu *et al.* 2008; Kubatko *et al.* 2009).

Phylogenetic networks are a generalization of phylogenetic trees that allow for the representation of conflicting signals or alternative evolutionary histories in a single diagram. Networks should be used if the underlying history is not treelike. For example, recombination, hybridization, and gene transfer can all lead to histories that are not adequately represented by a single tree (Grunewald *et al.* 2007). Moreover, even when the history is treelike, parallel evolution, model heterogeneity, and sampling error can make it hard to find a unique tree. In such cases, networks provide a tool for representing ambiguity or for visualizing a collection of feasible trees (Bryant & Moulton 2004; Huson & Bryant 2006). Thus, phylogenetic networks should be employed when reticulate events such as hybridization, horizontal gene transfer, recombination, or gene duplication and loss are believed to be involved, and even in the absence of such events, phylogenetic networks have a useful role to play to help resolve intraspecific relationships because gene evolution does not necessarily follow a strictly bifurcating model (Posada & Crandall 2001;

Huson & Bryant 2006). Network representations of data circumvent this problem and can help identify hybrid individuals or sequences that have undergone recombination (Bryant & Moulton 2004; Joly & Bruneau 2006). This characteristic of genetic networks enables them to provide portrayals of species relationships that are distinct from standard phylogenetic trees.

Analyses of multiple nuclear loci (nuDNA) have the potential to provide additional insights into historical processes and evolution of taxa, but incorporating nuDNA sequences in phylogeographic analyses is not as straightforward as traditionally done for mitochondrial DNA (mtDNA) (Spinks *et al.* 2010). Challenges include intragenic recombination, marker availability, and low variability of nuclear sequences (Hare 2001; Zhang & Hewitt 2003). Recombination rates have to be relatively high to bias phylogeographic analyses (Schierup & Hein 2000), and network analyses can be used to help elucidate intraspecific polytomies (Posada & Crandall 2001; Huson & Bryant 2006). More generally, the prospects for nuclear DNA phylogeography of organisms with relatively low nuclear sequence variation seem poor (Spinks *et al.* 2010), even as an ever-increasing body of work demonstrates that single marker analyses are often plagued by introgressive hybridization and/or incomplete lineage sorting (e.g. Bernatchez *et al.* 1995; Melo-Ferreira *et al.* 2005; Robertson *et al.* 2006; Linnen & Farrell 2007; Peters *et al.* 2007; Good *et al.* 2008; Spinks & Shaffer 2009). However, I am fortunate that abundant markers are presently available for the avian genome (e.g. Backstrom *et al.* 2008, Kimball *et al.* 2009) and next-generation sequencing technology enables the relatively straightforward development of additional loci. Further, the avian genome has been shown to harbour high-levels of polymorphism (e.g. Primmer *et al.* 2002), consequently with sufficient funding, studies of avian phylogeny and phylogeography may be well poised to take advantage of several newly proposed species-tree algorithms.

In this chapter I make use of several independent loci to reassess the evolutionary history of the Yellow-streaked greenbul (*Phyllastrephus flavostriatus*) species complex in depth. First, I make

used of gene trees and the traditional concatenated approach to phylogeny reconstruction (using mixed-models) to elucidate to what extent the phylogeny from the eight nuclear intron data set is concordant with the mitochondrial DNA phylogeny recovered in chapter five. Secondly, I make use of the hierarchical Bayesian algorithm implemented in the program BEST (Bayesian estimation of species trees; Liu *et al.* 2008) to estimate a species trees for the *P. flavostriatus* complex that explicitly accommodates incomplete lineage sorting and coalescence of multiple alleles within a species'. Finally I construct individual statistical parsimony networks (Templeton *et al.* 1992) for each locus as well as a mulitilocus split network (Huson & Bryant 2006) for the combined dataset.

6.2: Materials and Methods

Population sampling

Seventy individuals representing the major clades of the Yellow-streaked greenbul recovered in chapter five were included in this analysis. The number of individuals sampled varied according to subspecies and geographic location (Table 6.1, Figure 6.1, Appendix 5.1). Two individuals of *Phyllastrephus fisheri* were included as an out-group.

Multilocus nuclear data

I sequenced eight independent nuclear loci (Table 6.2). Two of these loci were sex-linked (BRAM5 and CHDZ) and the others six represent autosomal nuclear introns. Most of these loci have been widely used in avian phylogenetic studies (e.g., Fuchs *et al.* 2004; Moyle & Marks 2006; Johansson *et al.* 2007; Kimball *et al.* 2009) and have been demonstrated to be useful in delimiting species boundaries. Standard laboratory procedures (visualization, PCR-amplification, purification of PCR products and cycle sequencing) follow the protocol detailed in chapters three, four and five. All loci were sequenced in both forward and reverse directions using an ABI 3730 capillary sequencher. Some individuals were either sequenced several times or cloned into PCR®2.1-TOPO® vector when length polymorphism/base ambiguities were detected. Cloning followed the protocol described in chapters three and four. Contiguous DNA sequences were edited, assembled

into contigs, and aligned using Sequencher v4.7.

Gene trees versus total evidence

The eight nuclear loci were concatenated to conduct mixed-model maximum likelihood (ML) analysis using RAxML-VI-HPC v7.0.4 (Stamatakis 2006) and mixed-model Bayesian analysis using MrBayes v3.1.2 (Ronquist and Huelsenbeck 2003). The Akaike information criterion (AIC) in MrModeltest v2.2 (Nylander 2004) was used to determine the best-fit nucleotide substitution model (Table 6.2) for each locus, resulting in eight data partitions. Nonparametric bootstrapping (100 pseudoreplicates) was used to generate support values for ML analyses of the concatenated data. All ML analyses utilized the general-time-reversible (GTR)+I+Γ model. Partitioned Bayesian analyses used the substitution models selected using the AIC criterion in MrModeltest and were run for 50 million generations (sampling every 25 000 steps). Convergence was assessed using cumulative posterior probability plots constructed with the program AWTY (Nylander *et al.* 2008). I also used the program Tracer (Rambaut & Drummond 2007) to assess convergence diagnostics.

Bayesian estimation of species trees

I made use of the hierarchical Bayesian model implemented in BEST v2.2 (Liu *et al.* 2008) to estimate species trees for the yellow-streaked greenbul complex that take into account incomplete lineage sorting and accommodates multiple alleles within species. Two Markov Chain Monte Carlo (MCMC) runs were initiated at different starting seeds and allowed to proceed for 200 million generations (sampling every 100 000 steps). Convergence was assessed using burn-in plots of likelihood values and parameter estimates. The gene mutation prior was set at (0.5, 1.5) for all analyses. The prior distribution for the effective population size parameter theta (Θ) was modeled using an inverse gamma distribution, which is a 2-parameter probability distribution with a mean = $\beta/(\alpha - 1)$ (when $\alpha > 1$). Posterior probability values for species relationships were obtained by summarizing the posterior distribution of species trees (post burn-in) with a 50% majority-rule consensus tree.

Multilocus network and statistical parsimony networks

In order to produce a multilocus genetic network linking the 70 specimens of yellow-streaked greenbul included in this study, I first calculated the genetic distance among alleles at each nuclear locus. Resolving the phase of heterozygous genotypes was accomplished using PHASE v2.1.1 (Stephens & Donnelly 2003). I tested for intragenic recombination using the difference of sum of squares test in TOPALi v2.5 (McGuire & Wright 2000). Genetic distance matrices among alleles at each locus were calculated using uncorrected-p distances and the HKY85 (Hasegawa *et al.* 1985) model using PAUP v4.0b10 (Swofford 2001). The genetic distance matrices for separate loci were combined into a single distance matrix of individuals using the program POFAD v1.03 (Joly & Bruneau 2006). A genetic network among individuals was constructed using the NeighborNet algorithm (Bryant & Moulton 2004) in SplitsTree v4.6 (Huson & Bryant 2006). I also constructed statistical parsimony networks using phased haplotypic data (Clement *et al.* 2000) for each individual intron to elucidate how individual gene networks differ from the combined network as implement in TCSv 1.21 (Clement *et al.* 2000).

Population assignment of individuals

I used Structure v2.2.3 software (Pritchard *et al.* 2000) and allelic data from the eight nuclear locus data set generated from Phase 2.1.1 (Stephens *et al.* 2001; Stephens & Donnelly 2003) to assess geographic groupings of the individuals used in this study. I used the admixture model, with population information prior (LOCPRIOR), five runs with a burn-in of 100,000 and a run length of 1,000,000 for a number of clusters from K = 1 to 10, and allowed for correlation of allele frequencies among clusters. I performed five independent runs per K to ensure that the results were consistent. I investigated the most likely number of clusters in two different ways, by considering the log 'probability of data' lnP(D) for the different numbers of K, and by using the statistic ΔK (Evanno *et al.* 2005), which considers the rate of change in lnP(D) among successive K values.

6.3: Results

Multilocus Nuclear Data

All eight nuclear loci were sequenced for all 70 specimens sampled in the *P. flavostriatus* complex (Table 6.1, Appendix 5). The eight nuclear loci range in size from 311 to 830 bp (mean = 550 bp), contain 20–51 variable sites (mean = 31.75), and include 8–36 parsimony informative characters (mean = 21.4; Table 6.2). The eight nuclear loci indicated different patterns of genetic variability when analyzed with respect to the three major mitochondrial clades recovered in chapter 5; Albertine Rift grey-headed forms, grey-headed forms from elsewhere, and brown-head forms. FIB5 had the highest gene diversity (0.863 +/- 0.014) whereas ODC had the lowest (0.583 +/- 0.05) (Table 6.2). Uncorrected-p distances among members of the Yellow-streaked greenbul complex were much lower for nuclear DNA (ca. 1-4%; Table 6.2) than corresponding distances from mtDNA (ca. 4-5%; Chapter 5). Per locus AMOVA for the three main clades recovered the largest Fst value of 0.93 (d.f.=2, $P < 0.001$) for GAPD11, and the lowest for Fib7 (Fst = 0.27, d.f. = 2, $P < 0.001$) (Table 6.3). Contrasting results were recovered for the two sex-lined loci (Table 6.3) where BRM15 shows much of the variation to be within population (68%, whereas CHDZ suggested much of variation was among populations (69%).

Gene trees versus total evidence

Individual ML analyses of these eight nuclear loci result in a set of genealogies that support conflicting phylogenetic relationships among *P. flavostriatus* individuals sampled (results not shown but see Figure 6.3 (SPN networks)). Generally most of the basal nodes received poor support. Although no single gene tree is fully resolved; three of the loci (CHDZ, ODC and GAPD) provided strong support for placing brown-headed *alfredi* outside the grey-headed populations representing different subspecies. Only locus CHDZ recovered all three primary mtDNA clades. Relationships among individuals sampled from the coastal clade showed the most conflicting relationships among the eight loci studied.

Concatenation of the eight nuclear loci resulted in a data set containing 4398 bp of aligned nucleotide positions of which 372 (8.46%) were phylogenetic informative. Average pair-wise distance was 0.0057 and the nucleotide frequencies were A=28.67%, C=18.74%, G=20.92%, T=31.08% and other=0.59% (i.e. heterozygous bases) in the concatenated dataset. The partitioned ML phylogenetic analysis of the concatenated data did not recover the three-clade relationship based on geographic distribution as recovered with mtDNA in chapter five. Despite support for most of the nodes being poor, brown-headed *alfredi* was recovered as a distinct clade with high bootstrap support (Figure 6.2). The Albertine Rift grey-headed clade was subdivided into two distinct clades, whereas the costal grey-headed clade showed marked mixing of all the subspecies. The partitioned Bayesian phylogenetic analysis resulted in the same topology as the ML analysis, but with slightly better support (Fig. 6.2).

Bayesian Estimation of Species Trees

After running the analyses for 20 million generations with the default prior on theta, posterior distributions for most of the parameters (tree length and mutation rate) were unimodal and the ESS values were close to 100. The BEST tree provided a better topology compared to the concatenated ML tree, however posterior probability support for some subspecies was poor (Figure 6.2b). Strong support (posterior probability - 0.97) was recovered for the basally placed Albertine Rift grey-headed clade containing subspecies *P. f. graueri* and *P. f. olivaceogriseus,* which in turn are placed sister to a clade containing all the other subspecies including brown-headed *alfredi*. Weak posterior probability support received for the subspecies *kungwensis* and *alfredi* hinders interpretation of the phylogeographic pattern and evolutionary history of this complex. Subspecies *uzungwensis* also received high posterior probability support reaffirming the MI and BI results of the concatenated dataset.

Multilocus network and statistical parsimony networks

Individual statistical parsimony networks (Figure 6.3) showed similar patterns to the

topologies recovered in individual gene ML analyses. Three distinct linked haplotype clusters corresponding to the three primary mtDNA clades were recovered for the sex-linked locus CHDZ. Brown-headed *alfredi* was recovered as distinct linked haplotype cluster in locus ODC, however the two grey-headed forms were not clustered separately. In GAPD, brown-headed *alfredi* and the two grey-headed forms were recovered as two unlinked haplotype networks strongly supporting the monophyly of the *alfredi* cluster. In all other loci none of the primary clades were recovered as a distinct haplotype cluster indicating substantial heterogeneity among loci.

Overall the multilocus network (Figure 6.4) supports three phylogeographic groups within the *Phyllastrephus flavostriatus* complex as revealed in the combined mtDNA network (refer to Figure 5.2, Chapter 5). However, Instances of discordance are observed when compared with the mtDNA split network. A single sample of the subspecies *kungwensis* clustered with the Albertine Rift clade in the multilocus network but with the coastal clade in the mtDNA network. Within the coastal clade, subspecies *uzungwensis* was recovered as a distinct haplotype cluster, whereas the other taxa were not monophyletic.

Population assignment of individuals

STRUCTURE analyses performed on the nuclear data set (Figure 6.5) revealed five primary genotype clusters (-ln P(D)= 2025.1, K=5), corresponding to the brown-headed *P. f. alfredi* (blue), *P. f. graueri, P. f. olivaceogriseus* from Burundi and *P. f. kungwensis* from Mt. Mahale (light blue), to *P. f. olivaceogriseus* from DRC (pink), *P. f. flavostriatus* (Ngoye Forest) and *P. f. vincenti* (Mt. Mulanje) (red), *P. f. tenuirostris* from the Eastern Arc and east African coastal forest (yellow) and *P. f. uzungwensis* (light green). One important discordance observed in the structure analysis compared with mtDNA phylogeny reconstructed (Chapter 5) is that *P. f. grauri* and *P. f. kungwensis* cluster with *P. f. olivaceogriseus* from Burundi and Uganda and not with *P. f. olivaceogriseus* from DRC. However one individual from DRC *P. f. olivaceogriseus* clustered with the Burundi and Uganda samples with a posterior probability greater than 0.90. Admixture individuals were

observed with posterior probabilities lesser than 0.8 in the Eastern Arc, East African and southern grey-headed clade, particularly in *P. f. flavostriatus* and *P. f. tenuirostris*. I found the same parapatric relationship recovered between *vincenti* and nominal subspecies from southern Africa with the multi-locus nuclear intron too.

6.4: Discussion

This study provides a new perspective on the phylogenetic relationships of the Yellow-streaked greenbul species complex. Whereas in the previous chapter, I used a mtDNA-based approach to reconstruct the phylogenetic relationships among the disjunct and morphologically distinct populations, the results presented here are based on the phylogenetic analyses of eight independent nuclear loci. The probability of estimating an accurate species tree typically increases with the number of loci added, and the analyses such as the multilocus network and population assignment of individuals using the Structure algorithm (Pritchard *et al*. 2000) presented here takes an additional step toward improving phylogenetic accuracy among closely related lineages (Leache 2009; Fuchs in press) and help me to gain new insight into the underlying biogeography as described in subsequent sections.

mtDNA and multilocus nuDNA phylogenies: Phylogeographic discordance or concordance?

A comparison of the mtDNA and multilocus nuclear data genealogies reveals an intriguing pattern of discordance at the phylogeographic level within the *Phyllstrephus flavostriatus* complex. Although the nuclear and mtDNA data both recover three major phylogeographic groups (brown-headed *P. f. alfredi*, Albertine Rift grey-headed and Eastern Arc, East African coastal forest), the composition of these groups differs with respect to the phylogenetic placement of *P. f. kungwensis, P. f. grauri, P. f. olivaceogresius* from the DRC and *P. f. uzungwensis* specimens (Figure 6.2A and Table 6.1). It is clear with this type of spatial pattern of discordance that lineage sorting is completed for mtDNA but not for the multilocus nuclear intron data in most of the disjunct and

morphologically distinct populations. Further, the admixture individuals detected in the STRUCTURE analysis, and shared haplotypes detected in the SPN networks for individual genes between the Albertine Rift and Eastern Arc and East African coastal forest grey-headed morphotypes provide further evidence for either retained ancestral polymorphism or gene flow between this two heterogeneous segments of the EARS and hence the complex nature of montane–lowland fuanal interactions in the region. This may also suggest that reproductive isolation is not yet completed across the currently established subspecies boundaries within the two regions concerned. On the other hand the concordance observed with the multilocus network (Figure 6.4) and the mtDNA neighbor-net network (Figure 5.4) indicate that there is some genetic basis for current taxonomy and warrant the thorough reappraisal of members of the species complex with more sampling at clade boundaries (specially for the Albertine rift and brown headed forms) together with a robust coalescent analyses to elucidate the complex and unusual genetic and geographic differentiation among the grey- and brown-headed forms of the Yellow-streaked greenbul.

Species trees and phylogeographic implications

Species delimitation is often challenging in studies of recently diverged lineages (e.g. Leache 2009) as in this study. Empirical studies has shown that dense sampling improves the accuracy of phylogenetic resolution in species tree reconstructions using BEST (Belfiore *et al.* 2008; Leache 2009) although an increased number of loci is also typically required to achieve an improvement in phylogenetic accuracy as species sampling increases (Liu *et al.* 2008; McCormack *et al.* 2008). Further Leache 2009 has shown that species tree inference methods are sensitive to phylogeographic sampling and that species relationships can change depending upon which particular specimens are used to represent the "species". Because BEST makes the assumption that species assignments can be accurately made *a priori,* inaccurate species assignments will mislead species tree inference. In this study, I used the subspecies taxonomy to hold true as the *a*

priori assumption in BEST that stemmed from the reconstructed mtDNA phylogeny.

Overall topology of the resulting species tree suggests a stepping-stone model of colonization as hypothesized in chapter five stemming from the Albertine Rift, but the low posterior probability values for the placement of *kungwensis* and *alfredi* makes this hypothesis at present inconclusive even with the BEST method. One important fact that needs to highlight here is that the admixed individuals representing both montane and lowland *P. f. tenuirostris* detected with STRUCTURE show evidence of introgression with Albertine Rift individuals. This poses some uncertainty of the southeast direction of colonization from Albertine Rift to Mt. Mahale and then to form the brown-headed morphotypes and the coastal clade. In fact, it suggests northeast directional colonization from the Albertine Rift to the Eastern Arc and to along the coastal forest down to southern Africa. Absence of any extant populations in the Kenyan highlands may indicate extinction filtering happened during the Pleistocene glacial maxima. This is further supported by the absence of any admixture individuals in the brown-headed clade.

Phyllastrephus alfredi (Sharpe's Greenbul)

Given that the brown-headed lineage is well differentiated genetically in both mtDNA and nuclear intron data as well as morphologically, without any evidence of gene flow between neighboring populations of grey-headed forms (*P. f. kungwensis, P. f. vincenti and P. f. uzungwensis*) it is reasonable to consider populations as a distinct species, *P. alfredi.*

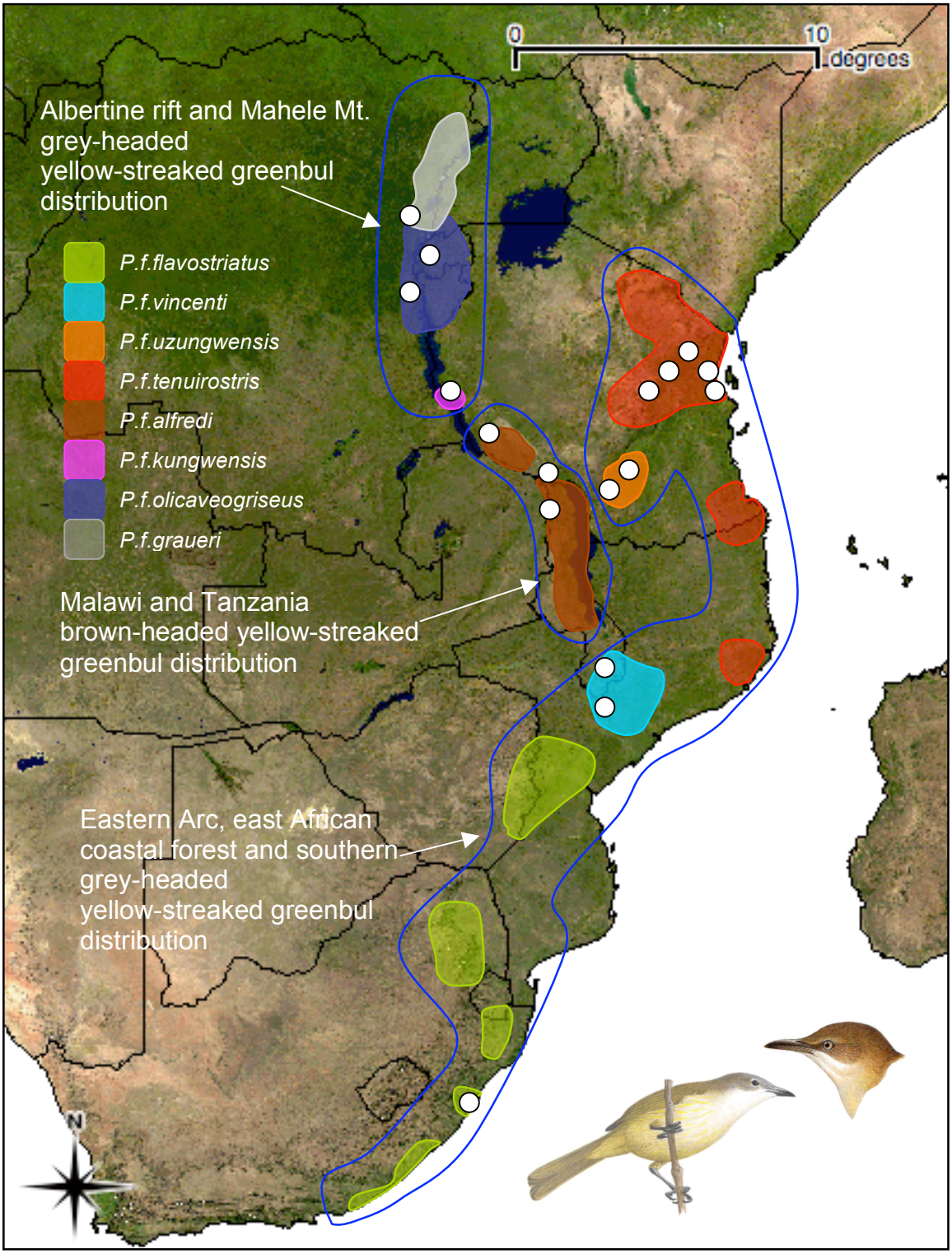


Figure 6.1 - Distribution range of *Phyllastrephus flavostriatus* complex in southern and eastern Africa. Distribution pattern of two color forms are shown in the area demarcated by black lines and the current distribution of each putative subspecies is shown in colored areas in the map. White color dots indicate the sample localities.

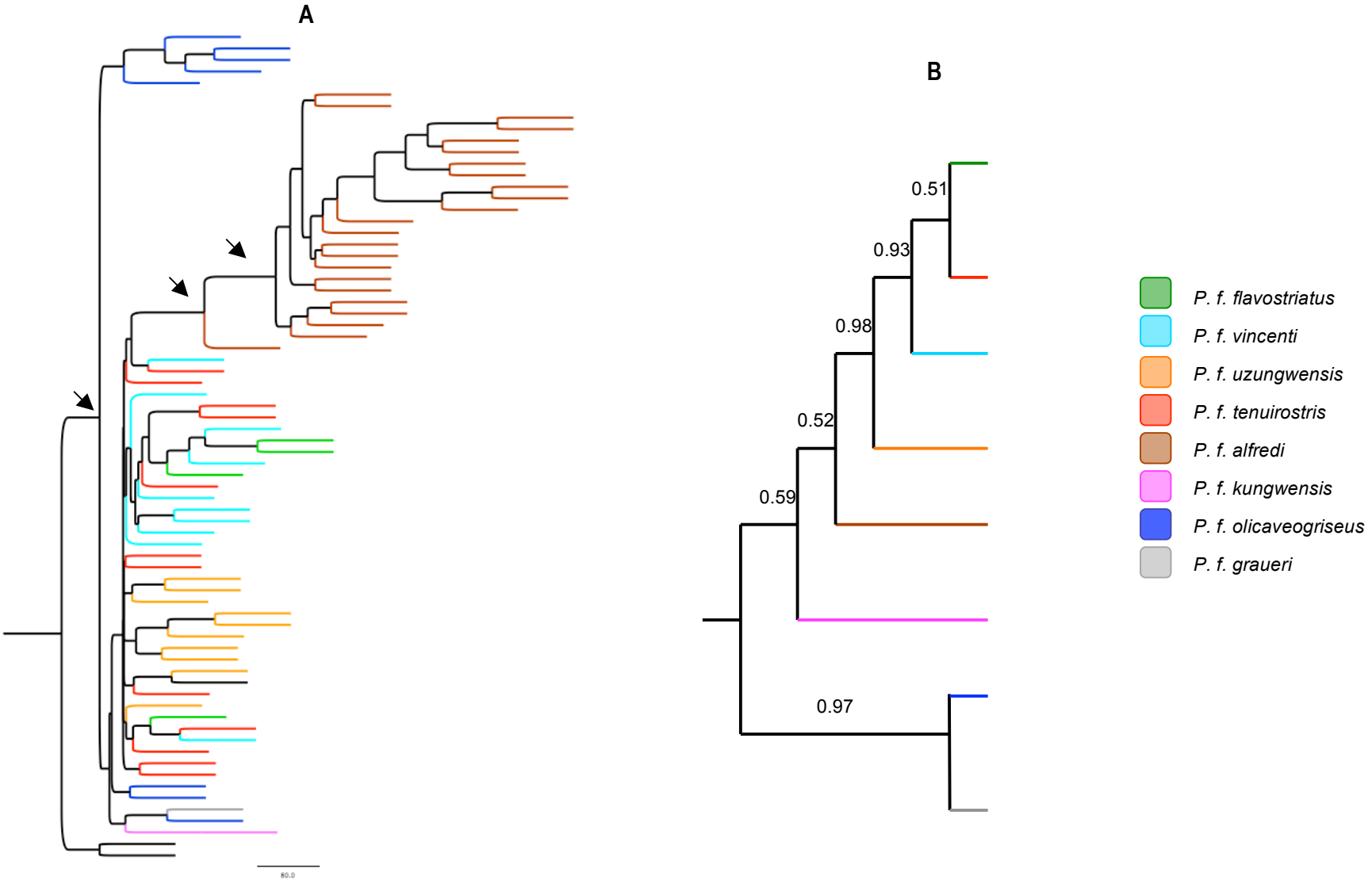


Figure 6.2 – A: Maximum likelihood majority rule consensus tree from the eight nuclear gene concatenated dataset produced from the RAxML. The color of the terminal branches correspond to the subspecies. Arrows indicate the nodes which received >75% bootstrap support and >0.95 Bayesian posterior probability. B- Species tree estimate for the *Phyllastrephus flavostriatus* complex group based on the BEST analysis of the 8 nuclear loci (mean Θ = 0.015). Values at each node indicate the estimated posterior probability.

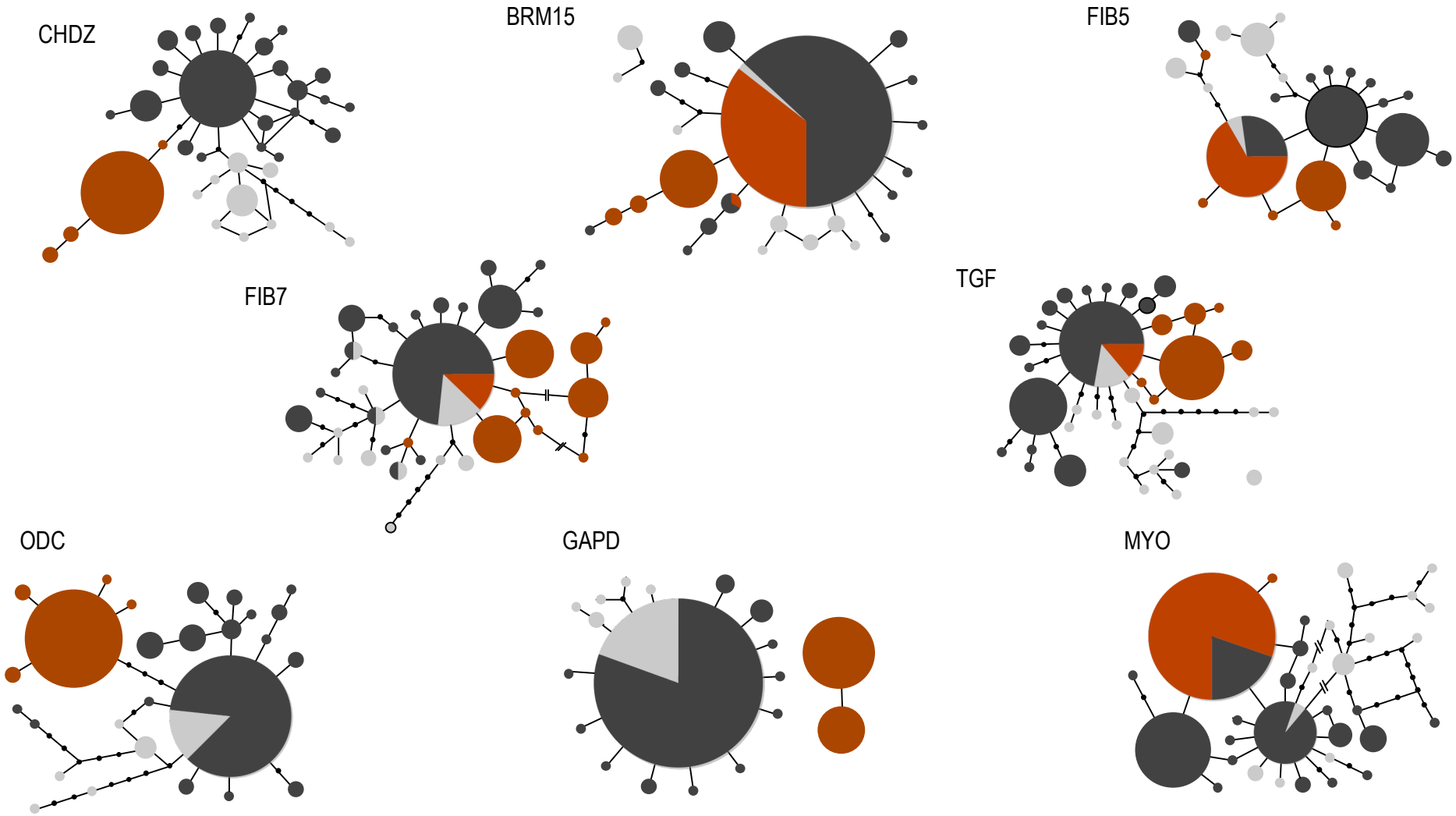


Figure 6.3 – Statistical parsimony networks constructed with phased data using TCS v 3.1(Clement *et al.* 2000) for each nuclear intron. Circled areas in the haplotype networks are proportional to the number of individuals possessing this haplotypes (drawn to scale). Extinct or unsampled haplotypes are indicated by black dots. Black lines indicate the number of connections between haplotypes. Brown color corresponds to the brown–headed subspecies *alfredi*. dark grey corresponds to the individual haplotypes representing the Eastern Arc, east African coastal forest and southern grey-headed form whereas, the light grey represent the individual haplotypes from the Albertine Rift grey-headed clade.

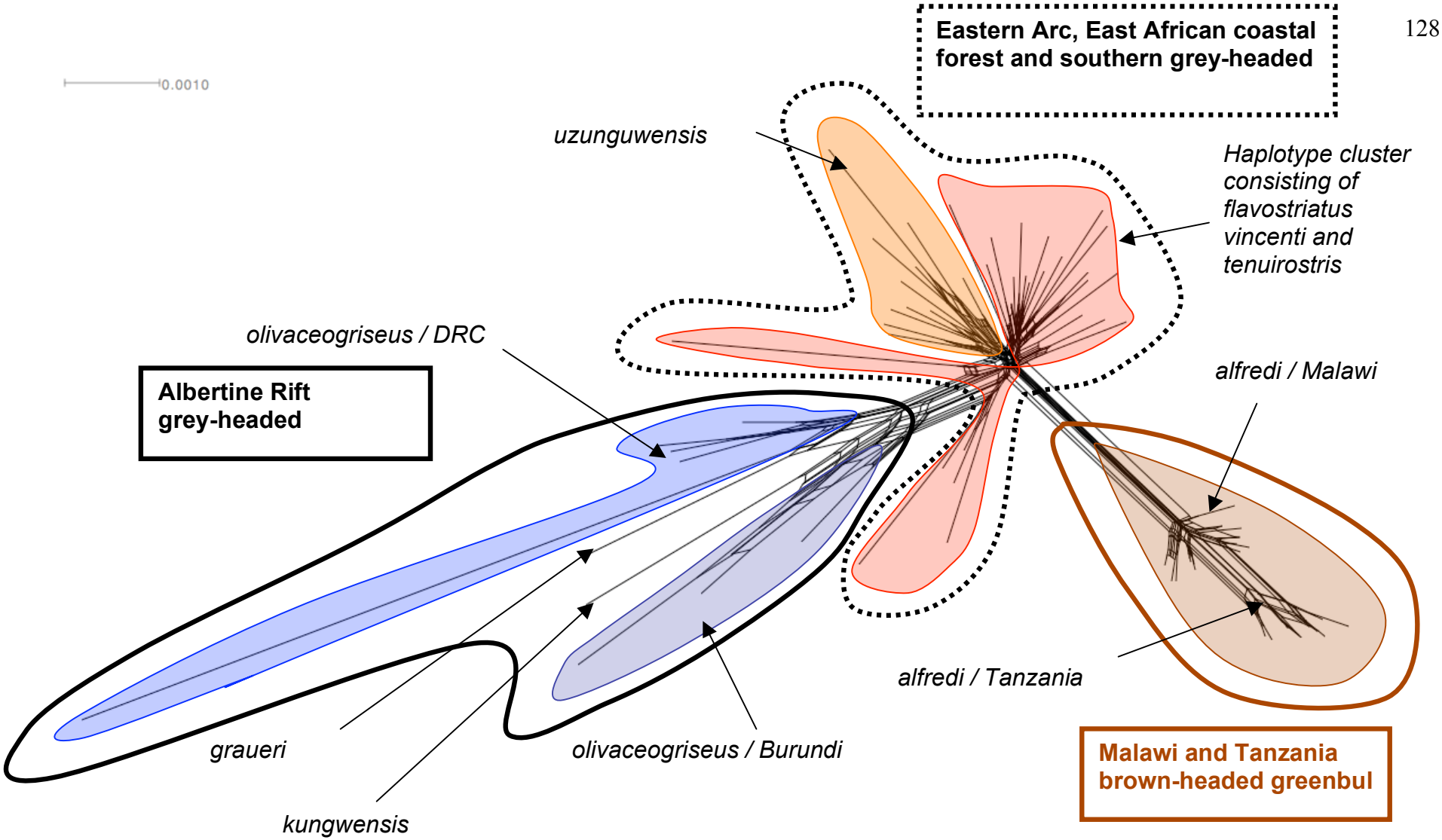


Figure 6.4 - Multilocus network produced from a Neighbor-Net analysis of genetic distances among individuals of the *Phyllastrephus flavostriatus* complex as implemented in program SplitsTrees (Huson & Bryant 2006) for the eight nuclear intron loci (n=70 individuals). Terminal clusters are colored to represent different subspecies/populations. Three main basal color forms are also shown. Bands of parallel edges represent incompatible splits.

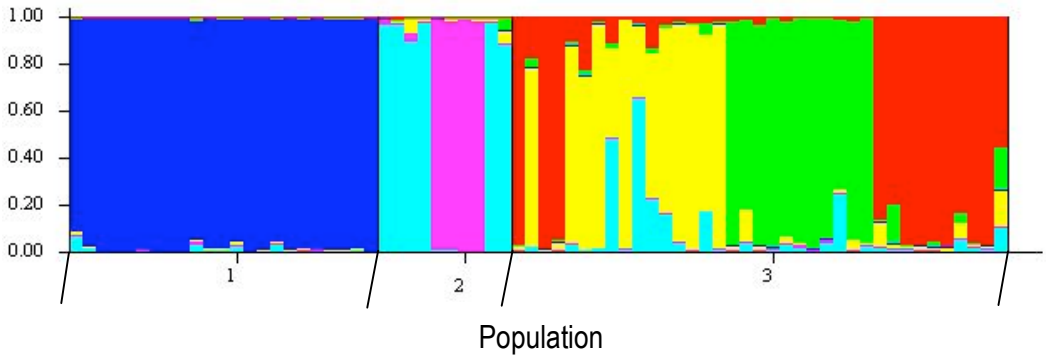


Figure 6.5 - Assignment of individuals to genetic clusters using the STRUCTURE algorithm for K=5 (mean LogLikelihood across five runs, -ln=2025.1,). The blue color corresponds to the brown-headed individuals (*P. f. alfredi*) sampled in Tanzania and Malawi. Light blue corresponds to individuals sampled in the Albertine Rift and on Mt. Mahale (*P. f. grauri, P. f. olivaceogreseus* from Burundi and *P. f. kungwensis* from Mt. Mahale). Pink corresponds to *P. f. olivaceogreseus* from DRC and red represent individuals from subspecies *P. f. flavostriatus* (Ngoye Forest), *P. f. vincenti* (Mt. Mulanje). Yellow corresponds to P. *f. tenuirostris* from the Eastern Arc and East African coastal forest. Light green corresponds to the individuals from *P. f. uzungwensis*. Multicolored bars represent admixture individuals. Population one corresponds to brown-headed yellow-streaked greenbuls, two to Albertine rft and Mahalr grey–headed forms and the three correspond to the Eastern Arc, East African coastal forest and southern grey-headed forms.

Table 6.1 - List of samples included in this study together with their DNA extraction numbers. See Appendix 5.1. for further details of the individuals sampled.

Subspecies	Number of samples used	Extraction numbers*
Phyllastrephus flavostriatus alfredi	23	LGS173, LGS222 to LGS229, LGS326 to LGS330, LGS340 to LGS342, LGS344, LGS346 to LGS347, LGS357 to LGS358, LGS908
Phyllastrephus flavostriatus falvostriatus	04	LGS001, LGS004 to LGS005, LGS077
Phyllastrephus flavostriatus graueri	01	LGS366
Phyllastrephus flavostriatus olivaceogriseus	08	LGS360 to LGS361, LGS365, LGS683 to LGS686, LGS688
Phyllastrephus flavostriatus tenuirostris	12	LGS160, LGS162, LGS182, LGS184, LGS267, LGS303 to LGS 306, LGS314, LGS325, LGS367
Phyllastrephus flavostriatus uzunguwensis	11	LGS177, LGS252, LGS280, LGS283, LGS291, LGS709 to LGS710, LGS713 to LGS716
Phyllastrephus flavostriatus vincenti	10	LGS309, LGS370, LGS412 to LGS419
Phyllastrephus flovistriatus kungwensis	01	LGS319

* Detailed information (museum catalog number, source, country of origin, locality etc.) on the samples can be found in the Appendix 5.

Table 6.2 - Summary statistics of the nuclear loci sequenced in this study.

Locus	Reference for PCR conditions	Number of characters (bp)	Variable sites	Parsimony informative sites	Nucleotide substitution model	No. of haplotypes	H	π	Tajima's D	Uncorrected P distances range	Fu's Fs
Myoglobin	Kimball *et al.* 2009	704	34	22	K80+I	30	0.797 +/- 0.028	0.0031 +/- 0.0003	-1.9564 p<0.05	0.0014 - 0.0158	-21.828
ODC	Allen & Omland 2003	675	33	26	HKY+I	26	0.583 +/- 0.05	0.0023 +/- 0.0003	-2.1772 p<0.01	0.0014 - 0.0133	-21.135
Fib5	Fuchs *et al.* 2004	558	23	16	HKY+I	25	0.863 +/- 0.014	0.0041 +/- 0.0003	-1.5294 p>0.10	0.0017 - 0.0125	-13.061
Fib7	Prychitko & Moore 1997	830	42	25	GTR+I	34	0.856 +/- 0.025	0.0031 +/- 0.0003	-2.0025 P < 0.05	0.0012 - 0.0158	-24.892
TGF	Kimball *et al.* 2009	514	51	36	HKY	32	0.781 +/- 0.035	0.0053 +/- 0.0008	-2.2030 p< 0.01	0.0133 - 0.0394	-20.736
GAPD11	Friesen *et al.* 1997	311	20	8	HKY	18	0.682 +/- 0.035	0.0069 +/- 0.0004	-1.3175 p> 0.01	0.0035 - 0.0257	-7.008
BRM15	Sætre *et al.* 2003	349	22	16	HKY	25	0.685 +/- 0.043	0.0036 +/- 0.0004	-1.9683 p<0.05	0.0028 - 0.0172	-24.382
CHDZ12	Kimball *et al.* 2009	482	29	22	GTR+I	31	0.839 +/- 0.022	0.0060 +/- 0.0004	-1.3666 p>0.01	0.0021 - 0.0269	-17.634

Table 6.3 - Results from AMOVA and population pairwise FST values.

	AMOVA			Population pairwise FSTs		
Locus	Among population variation (%)	Within population variation (%)	FST	Brown & coastal	Brown & Albertine rift	Coastal & Albertine rift
Myoglobin	36.7	63.3	0.3666	0.1583	0.6854	0.4143
ODC	69	31	0.6899	0.7608	0.8251	0.1708
Fib5	35	65	0.3501	0.2485	0.4800	0.4426
Fib7	27.2	72.8	0.2717	0.3150	0.2354	0.0794
TGF	27.38	72.62	0.2737	0.2903	0.3616	0.2432
GAPD11	93.92	6.08	0.9321	0.9619	0.9392	0.0678
BRM15	31.68	68.32	0.3168	0.1516	0.4280	0.4499
CHDZ12	69.2	30.8	0.6920	0.7058	0.8576	0.5331

CHAPTER 7

Synthesis: Climatic Perturbations and Speciation of Southern and East African Greenbuls (Passerifomes, Pycnonotidae)

7.1: Greenbuls, an ideal model group with which to explore biogeographic patterns in Africa

A thorough understanding of phylogenetic relationships among species and clear delineation of monophyletic groups is a fundamental to making modern biogeographical inferences. African greenbuls (family: Pycnonotidae) represent one of the most taxonomically complex groups of passerine birds (Kieth 1992) with more than 12 genera and 57 species (*sensu* Dickinson 2003) distributed across sub-Saharan Africa. Greenbuls are remarkable in their distribution patterns, habitat preferences, behaviour and inter- and intraspecific morphological variation as detailed in chapters one and two. African greenbuls occupy diverse habitat configurations ranging from: afromontane forest, lowland forest, grassland, miombo thickets, broadleaf woodland patches, scrub, riverine forest, swamps and seeps in drainage basins and coastal forests. Hence, greenbuls provide an ideal model system with which to study differences in diversification dynamics among and within the heterogeneous forest biomes of Africa. For these reasons and others detailed in chapter 1, I chose to focus on this diverse group of African birds for my thesis.

The taxonomy of the Pycnontidae, particularly African taxa, has only recently become more settled owing to several molecular phylogenetic revisions (Pasquet *et al.* 2001, Moyle & Marks 2006, Johansson *et al.* 2007; Fjeldså *et al.* 2007; Zuccon & Ericson 2010; Fuchs *et al*. in press). This has also provided greater insight into their distribution patterns, and consequently, the underlying evolutionary mechanisms have just begun to emerge. This wealth of new information will augment our understanding of the role of palaeoclimate, vicariance and dispersal in shaping the remarkable diversity within and among African greenbuls. Although progress has been made, as illustrated in this thesis for three species (chapters 3-6), phylogenetic resolution is still a work in progress for these and several other species complexes owing to unsampled populations, better analysical methods as well as data. Therefore, the history of African greenbuls can be inferred from both parallels in the biogeographic literature and the recent molecular work conducted on this family of birds. This synthesis is attempted in the following sections.

7.2: New perspectives on African greenbul biogeography

The core chapters of my thesis (three to six) focused on three distinct greenbul species: Grey-olive greenbul (*Phyllastrephus cerviniventris*), Sombre greenbul (*Andropadus importunus*) and the Yellow-streaked greenbul (*Phyllastrephus flavostriatus*). In the sections below, I present the new perspectives gained with respect to the biogeography of African greenbuls based on my dissertation research. First, I discuss how palaeodrainage systems particularly the Zambezi and Rufiji River basins have shaped the population genetic structure of bird species and the important of these overlooked habitat configurations as centers of speciation. Second, I detail how an ancient coastal forest endemic diversified in putative eastern and southern coastal forest refugia with special emphasis on phylogeographic breaks and cryptic genetic variation. Third, I discuss how disjunct polytypic populations of the Yellow-streaked greenbul diversified in the Albertine Rift and subsequently dispersed throughout it's current range. The potential stepping-stone fashion of dispersal routes from the Albertine Rift to the Eastern Arc, East African coastal forest, Rukwa and Malawi Rifts, and southern coastal forest are also addressed using a large multilocus nuclear intron dataset together with a species tree analytical approach. Further, I test whether there is any phylogenetic affinity with the evolution of two color forms (brown-headed versus grey-headed) within the Yellow-streaked greenbul complex. I also examine the extent to which lowland-montane interactions in habitat configurations may have shaped diversification in greenbuls and compare my results with parallel literature to gain new insight. Finally, I discuss the phylogeographic breaks unraveled in this study and their relevance to present day avian distribution patterns.

Overlooked refugia: watershed basins as an important habitat configuration for speciation

Grey-olive greenbuls inhabit riverine vegetation (gallery forest, gullies in open woodland, lake littoral, ground water forest, bamboo thickest etc.) in lowland forests in western, eastern and southern Africa. The current distribution range spans the Rukwa and Malawi Rifts separating

Tanzania from Zambia and Malawi. Habitat specificity of this species allowed me to test how palaeodrainage systems particularly the Zambezi and Rufiji River basins shaped the population genetic structure of this taxon. This was the first detailed study carried out on the phylogeography of a bird species distributed in palaeodrainage (river) basins in central and east Africa. Only a few comparable data sets are available for this habitat configuration (e.g. Linder 2001 for plants; Cotteril 2003, 2005 for mammals and fish; Daniels *et al.* 2006 for freshwater crabs; Dijkstra 2007 for dragonflies).

My estimated divergence times indicate a very recent Pleistocene(between 0.218 Ma (95% HPD: 0.1361-0.3062 Ma) and 0.667 Ma BP (95% HPD: 0.3163-1.1052 Ma) depending on which clock rate was used) split between the lineages east and west of the Rukwa and Malawi Rifts for the Grey-olive greenbul suggesting that the region comprises permanent refuges, most probably forest associated with seeps around the large permanent swamps of the Zambian Plateau and the Kilombero Valley, although these habitats are only likely to have been connected during the most humid periods. Further, there is clear indication of the involvement of tectonic activity across the Rukwa-Malawi Rift since the Pliocene, resulting in watershed reconfigurations in the region that facilitated diversification within other taxa, particually mammals, at least based on data published to date (detailed in Chapter 3), thus reaffirming the importance of these woodland drainage basins as a probable centers of recent speciation (Broadley & Cotterill 2004; Dijkstra 2007).

Eastern and southern African coastal forests: ancient forest refuges shape patterns of diversification in old and young species

I examined in detail how two greenbul species with contrasting distribution patterns (disjunct versus continuous and exclusively coastal versus lowland-montane) yet with subtle morphological variation diversified in the southern and east African coastal belt. These two species represent an ancient (*Andropadus importunus*) and a young (*Phyllastrephus flavostriatus*) greenbul

species (Johansson *et al.* 2007) with contrasting evolutionary histories indicating the complex nature of the biogeographic history in this heterogeneous landscape.

The coastal forest endemic Sombre greenbul forms a basal lineage in the African greenbul phylogeny (Johansson *et al.* 2007) and my TMRCA estimates suggest that it evolved around 4.4776 (2.8738-8.2942) MYA or 1.6563 (1.3494-2.1269) MYA based on 2.1% and 6% clock rates respectively. Both these suggest the species is an ancient species compared to the othe two specis I investigated in the study. Its sister relationship to the west African golden greenbul (*Calyptochichla serina*) supports the ancient coastal forest refuge hypothesis in east Africa and the faunal exchange between the Guinea-Congolian forest belt and the Eastern Arc during the period of habitat connectivity in the Miocene (deMenocal 2004, Voelker *et al.* 2010). Furthermore, it is comprised of deeply genetically structured lineages based on the estimated divergence times in this study (Table 4.5). However, the observed genetic structure is largely incongruent with the currently recognized subspecific level taxonomy (Figure 4.3). The recovered three reciprocally monophyletic mtDNA lineages have a geographical component. The northern clade corresponds to the subspecies *insularis* and the southern clade corresponds to nominal subspecies, but third (southeastern) clade is comprised of individuals from all four subspecies. This further indicates that coastal forest has existed across a considerable period despite Plio-Pleistocene climatic perturbations. Using molecular dating methods, I was able to conclude that the diversification of mtDNA lineages were climatically driven and hence the subtle morphological variation observed among the four subspecies may be an adaptation to a mosaic of forest patches in relatively stable climate, perhaps due to moisture off the Indian Ocean (Lawes *et al.* 2004). The ancestral lineage could have been retained in larger numbers in putative eastern and southern refuges (see Burgess *et al.* 1998; Lawes *et al.* 2004 for details). Further, it appears that at present, the southern and eastern populations of the Sombre greenbul are expanding and have come into contact at range boundaries. However, my study did not find any evidence of mtDNA introgression as revealed by

coalescence analyses. It is likely that the same external force, most probably climatically driven, as has been suggested for the diversification of montane avifauna (e.g. Bowie *et al.* 2006, Fjeldså & Bowie 2008, Voelker *et al.* 2010) is responsible for the diversification of coastal lowland *Andropadus importunus*. This is owing to the discovery of the same phylogeographic break (in the same geographical region) recovered for montane birds between the northern and central Eastern Arc also holding true for the lowland coastal taxa.

In contrast shallow divergence and very recent (Holocene) spits were observed among the disjunctly distributed coastal lineages (*tenuirostris, vincenti* and *falvostriatus*) of the Yellow-streaked greenbul complex. This is quiet interesting as some individuals from the nominate subspecies were found to be parapatric with *vincenti* from Mt. Mulanje, Malawi. This suggests that the South African coastal forest has had gene flow with montane forests of northern Mozambique and southern Malawi. Yellow-streaked greenbul diversification took place in response to Pleistocene climatic perturbations and dispersal from the Albertine Rift refuge, possibly in a stepping-stone manner, seems to be the primary cause of diversification and possibly the leap-frog pattern of morphological variation observed today. Discovery of a reciprocally monophyletic lineage of the nominal subspecies from the Zimbabwe highlands and its apparent stable population demographic history indicates that this region may represent a southern coastal forest refuge and the Ngoye Forest population of the nominal subspecies may be recently colonist from this refuge. No comparable avian data are available for the region except the recent work on Tiny greenbul (*Phyllastrephus debilis*; Fuchs *et al.* in review). Therefore my findings support the notion that the East African coastal zone has been under constant climatic influence from the Indian Ocean (Lovett & Wasser 1993) and seems to be a centre of diversification, or ‘‘species pump’’, as populations of different phylogenetic ages appear to have persisted in several putative refugia, and could thus have eventually dispersed to enrich the regional biota (Fjeldså & Lovett 1997; Jetz *et al.* 2004; Fjeldså *et al.* 2007).

Montane-lowland interactions: new insights from species tree analytical approaches

The disjunctly distributed morphologically divergent lineages of the Yellow-streaked greenbul provided an ideal system with which to test different speciation theories (Chapter 1) postulated for diverse African fauna. The heavy bias of avian phylogeographic studies towards montane taxa, has heightened the necessarily of exploring the relevance of putative speciation hypotheses for alternative habitat configurations. Yellow-streaked greenbul lineage diversification has occurred in the upper Pleistocene (Table 5.5) and seems to follow a stepping–stone fashion of colonization stemming from the Albertine Rift as detailed in chapters five and six. The two montane lineages of the Albertine Rift are older compared to the more southern lowland lineages *tenuirostris*. This disagrees with the "montane speciation hypothesis" described for some African bird taxa (Roy 1997) and is thus unlikely to be the evolutionary explanation for disjunct populations of Yellow-streaked greenbul. Dispersal during Upper Pleistocene and subsequent isolation seems to be the most likely mechanism responsible for diversification of this complex. Mitochondrial DNA supports a south-east direction (Chapter 5) of dispersal, whereas nuclear DNA suggests a north-east direction (Chapter 6) of dispersal from the Albertine Rift. However, a species tree approach using the multilocus nuclear DNA data also suggests a south-east stepping-stone dispersal pattern, in accordance with the mitochondrial DNA results.

The subspecies *tenuirostris* is distinct in its habitat preference compared to all other subspecies as this taxon is distributed in both the Eastern Arc Mountains (montane) and East African coastal forests (lowland). This indicates lowland-montane interactions in the region. Shared haplotypes and evidence for introgression strongly suggest that the lowland populations are maintaining gene flow between adjacent mountains. A similar pattern has been observed in *Pogonochichla stellata* populations (Bowie *et al.* 2006). To assess the extent and direction of gene flow detailed population genetic analyses are required with large number of samples and a suitable

panel of polymorphic markers. This provides further evidence for a complex biogeographical history in the region (see also Fjeldså & Bowie 2008).

Leapfrog pattern of plumage evolution in the Yellow-streaked greenbul: phylogenetic affinity versus divergent evolution

Both mitochondrial and nuclear phylogenies recovered three primary lineages corresponding to the two color forms (grey- and brown headed) indicating the association between the leapfrog color morphs and phylogenetic affinity. The species tree constructed with multiple independent nuclear loci support the stepping–stone model of diversification of the lineages in this complex. My analyses suggest that the origin of the leapfrog pattern of color morphology can be a consequence of evolutionary change in the intermediate population and the plumage of the intermediate forms represents a derived state. Accelerated divergence of centrally distributed populations compared to those at the periphery of the species' range seems to be the most likely evolutionary explanation for this leapfrog pattern of plumage evolution.

Phylogeographic barriers

This thesis revealed the existence of at least four phylogeographic breaks in eastern, central and southern Africa. The two breaks recovered in the coastal forest have not previously been described for lowland bird lineages. The proposed phylogeographic break in east Africa is likely to be the same barrier described previously for montane taxa (see Chapter 4) and this may indicate that the same external factor/force is responsible for shaping lineage diversification both in the montane Eastern Arc and east African coastal forests. The southern coastal forest genetic break corresponds to the location of the Maloti, Drakensberg and Witteberg Mountain ranges and Tugela valley. This barrier has previously been described for many taxa with limited dispersal capability (summarized in Chapter 4) and also appears to hold true for the Sombre greenbul, which

at first might appear to have better long distance dispersal ability. Several studies on mammals have shown that the lineages to the east and west of the Rukwa and Malawi Rift basins show genetic differentiation indicating restricted gene flow across the rift or that miombo woodland has restricted the distribution of savannah adapted species on either side of the rift. I recovered a similar pattern of genetic differentiation across the Rukwa and Malawi Rift in the Grey-olive greenbul and brown-headed population of the Yellow-streaked greenbul (Sharpe's greenbul). Yet, these two taxa occupy different habitats (Chapters 3 and 5) indicating the importance of this barrier for diversification of diverse taxa. My study is the first to investigate the Rukwa and Malawi Rift as a potential barrier to dispersal of avian lineages which may facilitate lineage divergence. Finally, the central part of the Albertine Rift has been postulated as a genetic break for many different taxa (Plumptre *et al.* 2007); I found strong evidence for genetic differentiation to across the left and right (east-west) flanks of the rift in *P. f. olivaceogriseus*.(chapters 5 & 6).

7.3: Synthesis

African greenbuls are an exciting system with which to explore different hypothesis that may underlie observed patterns of lineage diversification. In this study, the two key questions I wanted to address (chapter 1) were: 1) what determine the evolutionary relationships among and within three species of southern and east African greenbuls? and 2) what determine the phylogeographic population structure and historical biogeographic processes that have led to the current distribution patterns of the three species/species complexes?

The results of this study have contributed to three higher-level phylogeny papers (two already published and one in review) on greenbuls, with two manuscripts (see Appendix and chapter 2) specifically focused on the African greenbul radiation. Therefore, this thesis has provided a better understanding of the evolutionary relationships among diverse greenbul species.

In the second part of the thesis I tested different hypotheses on genetic structure in three selected greenbul species. My research established that most of the greenbul diversification took place in Plio-Pleistocene and the primary mechanism appears to be climatic cycling, yet dispersal and vicariance have shaped the population genetic structure. The pattern of diversification observed in the three study taxa/species complexes differs substantially and can mostly be explained by the Pleistocene refuge hypothesis. The study did not support the montane speciation hypothesis as articulated by Roy (1997) for some of the montane *Andropadus* taxa. This study found a close association between palaeo-drainage systems and swampy areas (seeps) as an important habitat configuration for diversification of lineages restructured to these patchy habitats. Further, this study revealed the complex nature of East African biogeography, and two possible routes of dispersal from Albertine Rift Refuga to the Eastern Arc and East African coastal forest in the Yellow-streaked greenbul complex.

REFERENCES

REFERENCES

Akaike, H., 1973. Information theory as an extension of the maximum likelihood principle. In: Second International Symposium on Information Theory. Petrov B. N. and Csaki, F. (eds). Budapest: Akademiai Kiado.pp. 267-281.

Alexio, A., 2002. Molecular systematics and the role of the "varzea-terra-firma" ecotone in the diversification of Xiphorhynchus woodcreepers (Aves: Dendrocolaptidae). *Auk* . **119**: 621-640.

Allen, E. S. and Omland, K. E., 2003. Novel intron phylogeny supports plumage convergence in *Orioles icterus*. *Auk* **120:** 961-969.

Alström, P., Ericson, P. G. P., Olsson, U. and Sundberg, P. 2006. Phylogeny and classification of the avian superfamily Sylvioidea. *Molecular Phylogenetics and Evolution* **38:** 381-397.

Arbogast, B. S., Drovetski, S. V., Curry, R. L., Boag,P. T., Seutin, G., Grant, P. R., Grant. B. R. and Anderson, D., 2006. The origin and diversification of Galapagos mockingbirds. *Evolution*. **60**: 370-382.

Avilla-Pires, T. C. S., 1995. Lizards of the Brazilian Amazon (Reptillia: Squamata). Zoologische Verhandelingen **299**: 1-709.

Avise, J. C., 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York.

Avise, J. C., 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos* **63**: 62-76.

Avise, J. C., 2000. Phylogeography, Cambridge University Press, London.

Axelrod, D. I. and Raven, P. H., 1978 Late Cretaceous and Tertiary vegetation history of Africa. In Biogeography and Ecology of Southern Africa. Werger, M. J. A. (ed.) pp. 77-130.The Hague: Dr W. Junk Publications.

Backström, N., Karaiskou, N., Leder, E. H., Gustafsson, L., Primmer, C. R., Qvarnström, A. and Ellegren, H., 2008. A gene-based genetic linkage map of the collared flycatcher (*Ficedula albicollis*) reveals extensive synteny and gene-order conservation during 100 million years of avian evolution. *Genetics*, **179:**1479-1495.

Ballard J. W. O. and Whitlock M. C., 2004. The incomplete natural history of mitochondria, *Molecular Ecology* **13** :729-744.

Barker, F. K., Barrowclough, G. F. and Groth, J. G., 2002. A phylogenetic hypothesis for passerine birds: taxonomic and biogeographic implications of an analysis of nuclear DNA sequence data. Proc. R. Soc. Lond. B **269**: 295-308.

Barker, F. K., Cibois, A., Schikler, P., Feinstein, J. and Cracraft, J., 2004. Phylogeny and diversification of the largest avian radiation. Proc. Natl. Acad. Sci. USA **101**: 11040-11045.

Bartholomew, G. A. and J. B. Birdsell J. B., 1953, Ecology and the protohominids. *American Anthropologist* **55:**481-498.

Barrowclough, G.F. 1992. Systematics, biodiversity and conservation biology. In Eldredge, N.(Ed.) Systematics, Ecology and the Biodiversity Crisis, New York: Columbia University Press. pp. 121-143.

Beerli, P. and Felsenstein, J., 1999. Maximum likelihood estimation of a migration matrix and eVective population sizes in two populations using a coalescent approach. *Genetics* **152**: 763-773.

Beerli, P. and Felsenstein, J., 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proc. Natl. Acad. Sci. USA **98**: 4563-4568.

Belfiore N.M., Liu L. and Moritz, C., 2008. Multilocus phylogenetics of a rapid radiation in the genus Thomomys (Rodentia: Geomyidae). Syst. Biol. **57**:294-310.

Beresford, P., Fjeldså, J. and Kiure, J., 2004. A new species of Akalat (Sheppardia) narrowly endemic in the Eastern Arc of Tanzania. *Auk* **121**: 23-34.

Beresford, P., Barker, F.K., Ryan, P.G. and Crowe, T.M., 2005. African endemics span the tree of songbirds (Passeri): Molecular systematics of several evolutionary 'enigmas'. Proc. Roy. Soc. Lond. B. **272**: 849-858.

Bernatchez, L., Glemet, H., Wilson, C. C., and Danzmann, R. G., 1995. Introgression and fixation of arctic char (*Salvelinus alpinus*) mitochondrial genome in an allopatric population of brook trout (*Salvelinus fontinalis*). *Canadian Journal of Fisheries and Aquatic Sciences*, **52**: 179-185.

Birky, C, W,. 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annual Review of Genetics*, **35**: 125-148.

Bonnefille, R., 1993. Evidence for a cooler and drier climate in the Ethiopian uplands towards 2.5 Myr ago. *Nature* **303**: 487-91

Bowie, R. C. K, Bloomer, P., Clancey, P. A. and Crowe, T. M., 2003. The Karoo Trush (Turdus smithi Bonaparte 1850), a southern African endemic. *Ostrich*, **74 (1and2):** 1-7.

Bowie, R. C. K. and Fjeldså, J., 2005. Genetic and morphological evidence for two species in the Udzungwa forest partridge *Xenoperdix udzungwensis*. *Journal of East African Natural History* **94(1):** 191-205.

Bowie, R. C. K., Fjeldså, J., Hackett, S. J. and Crowe, T. M., 2004a. Molecular evolution in space and through time: mtDNA phylogeography of the Olive Sunbird (*Nectarinia olivacea/obscura*) throughout continental Africa. *Molecular Phylogenetics and Evolution* **33:** 56-74.

Bowie, R. C. K., Fjeldså, J., Hackett, S. J., and Crowe, T. M., 2004b. Systematics and biogeography of double-collared sunbirds from the Eastern Arc Mountains, Tanzania. *The Auk*, **121**: 660-681.

Bowie, R. C. K., Fjeldså, J., Hackett, S. J., Bates, J. M. and Crowe, T. M., 2006. Coalescent models reveal the relative roles of ancestral polymorphism, vicariance, and dispersal in shaping phylogeographical structure of an African montane forest robin. *Molecular Phylogenetics and Evolution* **38:**171-188.

Bowie, R. C. K., Fjeldså, J. and Kiure, J., 2009. Multilocus molecular DNA variation in Winifred's Warbler *Scepomycter winifredae* suggests cryptic speciation and the existence of a threatened species in the Rubeho-Ukaguru Mountains of Tanzania. *Ibis*. **151**: 709-719.

Bowie, R. C. K., Voelker, G., Fjeldså, J., Lens, L., Hackett, S. J. and Crowe, T. M., 2005. Systematics of the olive thrush Turdus olivaceus complex with reference to the taxonomic status of the endangered Taita thrush *T. helleri. Journal of Avian Biology*. **36**: 391-404.

Brooks, T., Balmford, A., Burgess, N., Fjeldså, J., Hansen, L. A., Moore, J., Rahbek, C. and Williams, P., 2001. Towards a blueprint for conservation in Africa. *BioScience* **51**: 613-624.

Brown, J. H. and Lomolino, M. V., 1998. Biogeography. 2nd Ed. P 692, Sinauer Associates Sunderland, Massachusetts, MA

Bryant, D. and Moulton V., 2004. Neighbor-Net: an agglomerative method for construction of phylogenetic networks. Mol. Biol. Evol. **21**: 255-265.

Burgess, N. D., Butynski, T. M., Cordeiro, N. J., Doggart, N. H., Fjeldså, J., Howell, K. M., Kilahama, F. B., Loader, S. P., Lovett, J. C., Mbilinyi, B., Menegon M, Moyer, D. C., Nashanda E., Perkin, A., Rovero, F., Stanley, W. T. and Stuart, S. N., 2007 The biological importance of the Eastern Arc Mountains of Tanzania and Kenya. Biol Conserv. **134:**209-231.

Burgess, N. D., Clarke, G. P. and Rodgers, W. A., 1998. Coastal forests of eastern Africa: status, endemism patterns and their potential causes. Biol. J. Linn. Soc. Lond , **64**:337.

Burgess, N. D. and Clarke, G. P., (eds), 2000. The coastal forests of Eastern Africa. IUCN, Gland, Switzerland and Cambridge, UK

Carstens, B. C. and Knowles, L. L., 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from Melanoplus grasshoppers. Syst. Biol., **56**: 400-411.

Chapman, F. M., 1939. The riddle of Oxyruncus. Am. Mus. Novitat. **1047**: 1-4.380.

Chen, F. C. and Li. W. H., 2001. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. Am. J. Hum. Genet. **68**:444- 456.

Chikuni, K., Minaka, N. and Ikenaga, H., 1996. Molecular phylogeny of some Passeriformes, based on cytochrome b sequences. J. Yamashina Inst. Orn. **28**: 1-8.

Chorowicz J., 2005, The East African rift system, *Journal of African Earth Sciences* **43**: 379-410.

Cibois, A., Slikas, B., Schulenberg, T. S. and Pasquet, E., 2001. An endemic radiation of Malagasy songbirds revealed by mitochondrial DNA sequence data. *Evolution*, **55:** 1198-1206.

Clancey, P. A., 1975. On the edemic birds of the montane evergreen forest biome of the Transvaal. Durban Mus. Novit., **10:** 151-180. See p. 164.

Clarke, G. P., 2000. Defining the coastal forests. In: Burgess, N. D. and Clarke, G. P (eds), The coastal forests of Eastern Africa. IUCN, Gland, Switzerland and Cambridge, UK, pp 9-26

Clement, M., Posada, D. and Crandall, K. A. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**:1657-1660.

Coppes,Y., 1994 East side story: the origin of humankind. Sci. Am. **270**:62-69.

Cotterill, F. P. D., 2004. Drainage evolution in south-central Africa and vicariant speciation in Swamp-dwelling weaver birds and Swamp flycatchers. *The Honeyguide* **B(1):** 7-25.

Cotterill, F. P. D., 2003. Geomorphological influences on vicariant evolution of some African mammals in the Zambezi Basin: Some lessons for conservation. In: Proceedings of the Ecology and Conservation of Mini-antelope: An International Symposium on Duiker and Dwarf Antelope in Africa. Plowman. A. B. (Ed.). p.11-58.

Cotterill, F. P. D., 2005. The Upemba Lechwe, Kobus anselli: an antelope new to science emphasizes the conservation importance of Katanga (south-central Africa). J. Zool. London **265**: 113-132.

Cotterill, F. P. D., 2006. Taxonomic status and conservation importance of the avifauna of Katanga (Southeast Congo Basin) and its environs. *The Ostrich* **77:** 1-21.

Cracraft, J., 1983. Species concepts and speciation analysis. pp. 159-188 in Johnston, F. (ed.) Current ornithology, vol. 1. Plenum Press, New York

Crowe, T. M. and Crowe, A. A., 1982. Patterns of distribution, diversity and endemism in Afro-tropical birds. *Journal of Zoology* **198**: 417 - 442.

Crozier, R. H. and Kusmierski, R. M., 1994. Genetic distances and the setting of conservation priorities. In: Conservation Genetics (eds Loeschcke V, Tomiuk J, Jain SK), pp. 227-237. Birkhäuser Verlag, Basel.

Currie, D. J., Francis, A. P. and Kerr, J. T., 1999. Some general propositions about the study of spatial patterns of species richness. *Ecoscience*. **6**: 392-399.

Daniels, S. R., N. Cumberlidge, Pérez-Losada, M., Marijnissen, S. A. E. and Crandall, K. A., 2006. Evolution of Afro-tropical freshwater crab lineages obscured by morphological convergence. *Molecular Phylogenetics and Evolution* **40:** 227-235.

Dart, Raymond A., 1925 *Australopithecus africanus*: The Man-Ape of South Africa. *Nature* **115**:195.

Deacon, H. J., 1983. Another look at the Pleistocene climates of South Africa. S. Afr. J. Sci, **79:** 325-328.

Deacon, J. and Lancaster, N., 1988. Late Quaternary palaeoenvironmenits of southern Africa, 225 pp. Clarendon Press, Oxford.

Degnan J. H. and Rosenberg N. A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. **2:**762-768.

de Klerk, H. M., 1998. Biogeography and conservation of terrestrial Afrotropical. Birds. Ph.D. thesis, University of Cape Town, South Africa.

Delacour, J., 1943. A revision of the genera and species of the family Pycnonotidae (Bulbuls). *Zoologica* **28:** 17-28.

deMenocal, P. B., 1995. Plio-Pleistocene African climate. *Science* **270:** 53-59.

deMenocal, P. B., 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters* **220**:3-24.

Denys, C., Chorowicz, J. and Jaeger, J. J., 1985. Tectoge´ ne` se et e´ volution des faunes de rongeurs et autres mammife` res du Rift Est-Africain au Mio-Plioce`ne. Bulletin Socie´ te´ Ge´ ologique de France **8 (I, 3):** 381-389.

Denys, C., Chorowicz, J. and Tiercelin, J. J., 1986. Tectonics and environmental control on rodent diversity in the Plio-Pleistocene sediments of the African Rift System. *Geological Society Special Publication* **25**: 363-372.

Desjardins, P., Morais, R., 1990. Sequence and gene organisation of the chicken mitochondrial genome. A novel gene order in higher vertebrates. J. Mol. Evol. **212**: 599-634.

Diamond, A. W. and Hamilton, A. C., 1980. The distribution of forest passerine birds and quaternary climate change in tropical Africa. J. Zool. **191:**379-402.

Dickinson, E. C. 2003. The Howard and Moore Complete Checklist of the Birds of the World. London: Christopher Helm.

Dijkstra, K. D. B., 2007. Gone with the wind: westward dispersal across the Indian Ocean and island speciation in Hemicordulia dragonflies (Odonata: Corduliidae). *Zootaxa* **1438**: 27-48.

Dolman, G. and Moritz, C., 2006. A multilocus perspective on refugial isolation and divergence in rainforest skinks (Carlia). *Evolution* **60**: 573-582.

Dowsett, R. J. and Dowsett-Lamarie. 1980. Systematic status of some Zambian birds. *Gerfaut* **70**: 151-199.

Dowsett, R. J., Olson, S. L., Roy, M. S. and Dowsett-Lemaire, F., 1999. Systematic status of the Black-collared Bulbul *Neolestes torquatus*. *Ibis*, **141**: 22-28.

Dowsett-Lemaire, F., 1989. Ecological and biogeographical aspects of forest bird communities in Malawi. *Scopus* **13**:1-80.

Doyle, J. J., 1992. Gene trees and species trees: molecular systematics as one character taxonomy. Syst. Bot. **17:**144-63

Drummond, A. J., Rambaut, A., Shapiro, B. and Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol. **22:**1185-1192.

Drummond, A. J., Nicholls, G. K., Rodrigo, A. G. and Solomon. W., 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* **161:** 1307-1320.

Drummond, A. J., Ho, S. Y. W., Phillips, M. J. and Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biology 4: e88.

Drummond, A. J. and Rambaut. A., 2007. BEAST v1.4.8, Available from http ://beast.bio.ed. ac.uk/.

Dupont, L. M. and Hooghiemstra, H.,1989. Variation of the African trade wind regime during the last 140,000 years: changes in pollen flux evidenced by marine sediment records. In Paleo-climatology and Paleometeorology: Modern and Past Patterns of Global Atmospheric Transports (M. Leinen and S. Sarnhein, eds) pp. 733-770. Nordecht, Netherlands: Kluver.

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

Edwards S. V., Liu L. and Pearl D. K., 2007. High-resolution species trees without concatenation. Proc. Natl. Acad. Sci. USA. **104:**5936-5941.

Endler, J. A., 1977. Geographic variation, speciation, and clines. Princeton University Press, Princeton, New Jersey.

Endler, J. A., 1982a. Pleistocene forest refuges: fact or fancy. In: Biological Diversification in the Tropics (G. T. Prance, ed.), pp. 641-657. Columbia University Press, New York.

Endler, J. A., 1982b. Problems in distinguishing historical from ecological factors in biogeography. Am. Zool**. 22:** 441-452.

Ericson, P. G. P. and Johansson, U. S., 2003. Phylogeny of Passerida (Aves: Passeriformes) based on nuclear and mitochondrial sequence data. Mol. Phylogenet. Evol. **29:** 126-138.

Erwin, T. L., 1991, How many species are there? Revisited, *Conservation Biology*. **5(3)** 330-333.

Evanno G., Regnaut S. and Goudet J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611-2620.

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

Excoffier, L., 2004, Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. Mol Ecol, **13:**853-864.

Excoffier, L., Smouse, P. E. and Quattro, J. M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes-application to human mitochondrial-DNA restriction data. *Genetics* **131:**479-491.

Faith, D. P.,1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**: 1-10.

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

Fishpool, L. D. C. and Evans, M. I. eds. 2001. *Important bird areas in Africa and associated islands: priority sites for conservation.* Newbury and Cambridge, U.K. Pisces Publications and BirdLife International.

Fjeldså, J. and Bowie, R. C. K., 2008. New perspectives on the origin and diversification of Africa's forest avifauna. *African Journal of Ecology* **46:** 235-247.

Fjeldså, J., Bowie R. C. K. and Kiure, J. 2006. The Forest Batis, *Batis mixta*, is two species: description of a new, narrowly distributed Batis species in the Eastern Arc biodiversity hotspot. *Journal of Ornithology* **147:** 578-590.

Fjeldså, J., Goodman, S. M., Schulenberg, T. S. and Slikas, B., 1999. Molecular evidence for relationships of Malagasy birds. Pp. 3084-3094 in N. Adams and R. Slotow, eds. proceedings of the 22nd international ornithological congress. Univ. of Natal, Durban.

Fjeldså, J. and Lovett, J. C., 1997. Geographical patterns of old and young species in African forest biota: the significance of specific montane areas as evolutionary centres. Biodiversity Conserv. **6:** 325-346.

Fjeldså, J., 2003. Patterns of endemism in African birds: how much does taxonomy matter? *Ostrich*, **74**: 30-38.

Fjeldså, J., Johansson, U. S., Lokugalappatti, L. G. S., Bowie, R. C. K., 2007. Diversification of African greenbuls in space and time: linking ecological and historical processes. *Journal of Ornithology* **148**: 359-367.

Fjeldså, J. and Burgess, N. D., 2008, The coincidence of biodiversity patterns and human settlement in Africa, *African Journal of Ecology*. **46(1)** 33-42.

Fjeldså, J. and Krabbe, N. 1990. Birds of the high Andes. Zool. Mus., Univ. of Copenhagen and Apollo Books, Svendborg. "http://cms.ku.dk/nat-sites/snm-sites/snm-faelles/" **,** http://cms. ku.dk/nat-sites/snm-sites/snm-faelles/ english/ staffsnm/ vip/ publication detail/?id=1097048

Fjeldså, J. and Rahbek, C. 2006., Diversification of tanagers, a species rich bird group, from lowlands to montane regions of South America. Integr Comp Biol **46:**72-81.

Ford, J., 1983. Speciation in the ground-thrush complex *Zoothera dauma* in Australia. *Emu* **83**: 141-151.

Freitag, S. and Robinson, T. J., 1993. Phylogeographic patterns in mitochondrial DNA of the Ostrich (*Struthio camelus*). *The Auk*, **110:** 614-622.

Fu, Y. X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147:**915-925.

Fuchs, J., Bowie, R. C. K., Fjeldså, J. and Pasquet, E., 2004. Phylogenetic relationships of the African bush-shrikes and helmet-shrikes (Passeriformes: Malaconotidae). Mol Phylogenet Evol **33:**428-439.

Fuchs, J., Fjeldså, J., Bowie, R. C. K., Voelker, G. and Pasquet, E., 2006. The African warbler genus *Hyliota* as a lost lineage in the Oscine songbird tree: molecular support for an African origin of the Passerida. *Molecular Phylogenetics and Evolution* **39:** 186-197.

Fuchs J., Ericson P.G.P., Pasquet E., 2008. Mitochondrial phylogeographic structure of the white-browed piculet (Sasia ochracea): cryptic genetic differentiation and endemism in Indochina. *Journal of Biogeography* **35:** 565-575.

Garcia-Moreno, J. and Fjeldså, J., 2000. Chronology and mode of speciation in the Andean avifauna. In: Isolated Vertebrate Communities in the Tropics (G. Rheinwald, ed.), pp. 25-46. IV International Symposium, Zoologisches Forshungs institute and Museum, A. Koenig, Bonn.

Gascon, C., Williamson, G. B. and Fonseca, G. A. B., 2000, Receding Forest Edges and Vanishing Reserves, *Science*, **288(5470):** 1356-1358

Gasse, F., Chalie´, F., Vincens, A., Martin A. J., Williams, M. A. J. and Williamson, D., 2008. Climatic patterns in equatorial and southern Africa from 30,000 to 10,000 years ago reconstructed from terrestrial and near-shore proxy data, *Quaternary Science Reviews* **27:** 2316-2340.

Geldenhuys, C. J., 1992. Richness composition and relationship of the flora of selected forest in southern Africa. *Bothalia*. **22**:205-233

Glenn, T. C. and Schable, N. A., 2005. Isolating microsatellite DNA loci. *Methods in Enzymology*, **395**: 202-222.

Good, J. M., Hird, S., Reid, N., Demboski, J. R., Steppan, S. J., Martin-Nims, T. R. and Sullivan, J., 2008. Ancient hybridization and mitochondrial capture between two species of chipmunks. *Molecular Ecology*. **17**: 1313-1327.

Gould, S. J. and Johnston R. E., 1972. Geographic variation. *Annual Review of Ecology and Systematics* **3:**457-498.

Griffiths, C. J., 1993. The geological evolution of East Africa. pp 23-32., In Biogeography and Ecology of the Rain Forests of Eastern Africa (Lovett, J. C and Wasser, S. K., Eds). Cambridge University Press, Cambridge.

Grunewald, S., Forslund, K., Dress, A. and Moulton, V., 2007. QNet: An Agglomerative Method for the Construction of Phylogenetic Networks from Weighted Quartets, Mol. Biol. Evol. **24(2):** 532-538. 2007.

Haffer, J., 1969. Speciation in Amazonian forest birds. *Science* **165:**131-37.

Haffer, J., 1992. On the "river effect" in some forest birds of southern Amazonia.Bol. Mus. Para. E. Goldi S´er. Zool. **8**:217-45.

Haffer, J., 1997. Alternative models of vertebrate speciation in Amazonia: an overview. Biodivers. Conserv. **6:**451-77.

Haffer, J., 1978. Distribution of Amazon forest birds. Bonner zool. Beiträge **29**: 38-78.

Haffer, J., 1993. Time's cycle and time's arrow in the history of Amazonia. *Biogeographica* **69:**15-45.

Hall, B. P. and Moreau, R. E., 1970. 'An Atlas of speciation in African Passerine birds', British museum (Natural History) London.

Hamilton, A. C., 1976. The significance of patterns of distribution shown by forest plants and animals in tropical Africa for the reconstruction of Upper Pleistocene palaeoenvironments: a review. *Palaeoecology of Africa, the Surrounding Islands and Antarctica* **9**: 63-97.

Hamilton, A. C., 1982. Environmental History of East Africa. A Study of the Quaternary. London and New York: Academic Press.

Hamilton, A. C. and Taylor, D., 1991. History of climate and forests in tropical Africa during the last 8 million years. Clim. Change, **19**: 65-78.

Hanotte O., Knox A.G., Prigogine A., 1987. The taxonomic status of Nicator: Evidence from feather protein electrophoresis, *Biochemical Systematics and Ecology*. **15(5):** 629-634.

Harpending, H. C., .1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* **66:** 591-600.

Harpending, H. C., Batzer, M. A., Gurven, M., Jorde, L. B.,. Rogers A. R and Sherry, S. T.,1998. Genetic traces of ancient demography. Proc. Nat. Acad. Sci. USA **95**: 1961-1967.

Hare, M. P., 2001. Prospects for nuclear gene phylogeography. *Trends in Ecology and Evolution*, **12:** 700-706.

Hasegawa M., Kishino H. and Yano T. 1985. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. **22:**160-174.

Hawkins, B. A. and Diniz-Filho, J. A. F., 2006. Beyond Rapoport's rule: evaluating range size patterns of New World birds in a two-dimensional framework. Global Ecol Biogeogr. **15:**461-469.

Hawkins, B. A., Diniz-Filho, J. A. F. and Soeller, S. A., 2005. Water links the historical and contemporary components of the Australian bird diversity gradient. J Biogeogr **32:**1-8.

Hayes, F. E., 2001. Geographic variation, hybridization, and the leapfrog pattern of evolution in the suiriri flycatcher (*Suiriri suiriri*) complex. *Auk* **118:** 457-471.

Hey, J. and Nielsen. R., 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167:**747-760.

Hey, J. and Nielsen, R., 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proceedings of the National Academy of Sciences, USA, **104:** 2785-2790.

Hershkovitz, P., 1977. Living New World monkeys. Vol. 1. University of Chicago Press, Chicago, Illinois.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405: 907-913.

Hewitt, G. M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol.J.linn.Soc.**58**, 247-276.

Hewitt, G.M., 2004. Genetic consequences of climatic oscillations in the Quaternary. Philos. Trans. R. Soc. **B 359**: 183-195.

Ho, S. Y. W. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. *J. Avian Biol.* **38:** 409-414.

Hockey, P. A. R., Dean, W. R. J. and Ryan, P.G., (eds). 2005. Roberts - birds of Southern Africa, 7th edn. The Trustees of the John Voelcker Bird Book Fund, Cape Town.

Honda, M. and Yamagishi, S., 2000. A molecular perspective on oscine phylogeny, with special reference to inter-familial relationships. Jpn J. Ornithol. **49**: 175-184.

Hovanitz, W., 1940. Ecological color variation in a butterfly and the problem of "protective coloration". *Ecology* **21:** 371- 380.

Huelsenbeck, J. P. and Ronquist, F., 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics* **17**: 754-755.

Humphries, C. J., Williams P. H. and Vane-Wright R. I,. 1995. Measuring Biodiversity Value for Conservation, *Annual Review of Ecology and Systematics*, **26:** 93-111

Hudson, R. R., Kreitman, M. and Aguadé, M.,1987. A test of neutral molecular evolution based on nucleotide data. *Genetics*.**116:** 153-159.

Huson, D. H. and Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. **23:** 254-267.

Jetz, W. and Rahbek, C., 2002. Geographic range size and determinants of avian species richness. *Science*, **297:** 1548-1551.

Johansson, U. S., Fjeldså, J., Lokugalappatti, L. G. S and Bowie, R. C. K. 2007. A nuclear DNA phylogeny and proposed taxonomic revision of African greenbuls (Aves, Passeriformes, Pycnonotidae). *Zoologica Scripta*, **36**: 417-427.

Johansson, U.S., Fjeldså, J. and Bowie, R.C.K. 2008. Phylogenetic relationships within Passerida (Aves: Passeriformes): a review and a new molecular phylogeny based on three nuclear intron markers. *Molecular Phylogenetics and Evolution*. **48:** 858-876.

Johnson, N. K. 2002. Leapfrogging revisited in Andean birds: geographical variation in the Tody-tyrant superspecies *Poecilotriccus ruficeps* and *P. luluae*. *Ibis* **144:** 69-84.

Joly, S., Bruneau, A. 2006. Incorporating allelic variation for reconstructing the evolutionary history of organisms from multiple genes: an example from Rosa in North America. Syst. Biol. **55:**623-636.

Jensen, F. P. and S. N. Stuart.,1982. New subspecies of forest birds from Tanzania. Bull. Brit. Orn. Cl., **102:** 95-99. See p. 96.

Kahindo, C., Bowie, R. C. K. and Bates, J. M., 2007. The relevance of data on genetic diversity for the conservation of Afro-montane regions. Biol. Conserv. **134:** 262-270.

Kaliba, P.M., 2006. Faunal turnover between east and southern Africa birds and small mammals: is Malawi the geographical break? MSc thesis. University of Cape Town, South Africa.

Keith, S., 1992. Pycnonotidae, bulbuls. In S. Keith and E. K. Urban and C. H. Fry (Eds) The Birds of Africa IV. London: Academic Press, pp. 1-459.

Kerr JT, Currie DJ (1999) The relative importance of evolutionary and environmental controls on broad-scale patterns of species richness in North America. *Ecoscience* **6:**329-337.

Kimball RT, Braun EL, Barker FK et al.: A well-tested set of primers to amplify regions spread across the avian genome. Molecular Phylogenetics and Evolution 2009, **50:** 654-660.

King, B.C., 1978. Structural and volcanic evolution of the Gregory Rift Valley. In: Bishop, W.W. (Ed.), Geological Background to Fossil Man. Scottish Academic Press, London, pp. 29-54

King, L., 1978. A geomorphology of central and southern Africa. Biogeography and ecology of southern Africa. Monographiae Biologicae 31 (ed. by M. J. A. Werger), pp. 1-17. Junk, The Hague.

Klicka, J. and Zink. R. M., 1999. Pleistocene phylogeographic effects on avian evolution. Proceedings of the Royal Society of London, Series B. 266:695-700.

Kluge A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among Epicrates (Boidae, Serpentes). Syst. Zool. 38:7-25.

Knowles, L. L. and Maddison, W. P., 2002. Statistical phylogeography. Mol. Ecol. **11:** 2623-2635.

Kolaczkowski, B. and Thornton, J. W., 2004. Performance of maximum parsimony and maximum likelihood phylogenetics when evolution is heterogeneous. *Nature*, **431:** 980-984.

Kubatko L. S. and Degnan J. H. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. **56:** 17-24.

Lawes, M. J., 1990. The distribution of the samango monkey (*Cercopithecus mitis erythrarchus* Peters, 1852 and *Cercopithecus mitis labiatus* I. Geoffroy, 1843) and forest history in southern Africa. *Journal of Biogeography* **17:** 669-680.

Lawes M. J., 2004. The guenons: diversity and adaptation in African Monkeys, Developments in Primatology: Progress and Prospects, 2004, Part IV, 375-392.

Lawes, M. J., Eeley, H. A. C., Findlay, N. J. and Forbes, D., 2007. Resilient forest faunal communities in South Africa: a legacy of palaeoclimatic change and extinction filtering? *Journal of Biogeography*, **34:** 1246-1264.

Laurent, R. F., 1973. A parallel survey of equatorial amphibians and reptiles in Africa and South America. In Tropical Forest Ecosystems in Africa and South America (B.J. Meggers, E.S. Ayensu and W.D. Duckworth, eds) pp. 256-66. Washington: Smithsonian Institution Press.

Leache, A. D., 2009. Species Tree Discordance Traces to Phylogeographic Clade Boundaries in North American Fence Lizards (*Sceloporus*) Syst. Biol. **58(6):**547-559.

Linder, H. P., 2001. Plant diversity and endemism in sub-Saharan tropical Africa. J. Biogeogr., **28:** 169-182.

Linnen, C. R. and Farrell, B. D., 2007. Mitonuclear discordance is caused by rampant mitochondrial introgression in Neodiprion (Hymenoptera:Diprionidae) sawflies. *Evolution*, **61:** 1417- 1438.

Livingstone, D. A., 1975. Late Quaternary climatic change in Africa. Ann. Rev. Ecol. Syst. **6**:249-80.

Livingstone, D. A., 1993. Evolution of African climate. In: Biological Relationships between Africa and South America. P. Goldblatt (ed.). New Haven, Yale University Press. p. 455-72.

Liu L., Pearl D. K. 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. Syst. Biol. 56:504-514.

Liu L., Pearl D. K., Brumfield R. T. and Edwards S. V., 2008. Estimating species trees using multiple-allele DNA sequence data. *Evolution*. **62:**2080-2091.

Lorenzen, E. D., Arctander, P. and Siegismund, H. R., 2006b, Regional genetic structuring and evolutionary history of the impala *Aepyceros melampus*. *Journal of Heredity*, **97**: 119-132.

Lorenzen, E. D., Arctander, P. and Siegismund, H. R., 2008. Three reciprocally monophyletic mtDNA lineages elucidate the taxonomic status of Grant's gazelles. *Conservation Genetics*, **9**: 593-601.

Lorenzen, E. D., Masembe, C., Arctander, P. and Siegismund, H. R. (2010), A long-standing Pleistocene refugium in southern Africa and a mosaic of refugia in East Africa: insights from mtDNA and the common eland antelope. *Journal of Biogeography*, **37**: 571-581

Lorenzen, E. D., Simonsen, B. T., Kat, P. W., Arctander, P. and Siegismund, H. R., 2006a Hybridization between subspecies of waterbuck (*Kobus ellipsiprymnus*) in zones of overlap with limited introgression. *Molecular Ecology*, **15**: 3787-3799.

Lovett, J. C., 1993. Eastern Arc moist forest flora. In: Biogeography and ecology of the rain forests of eastern Africa. Lovett, J. C. and Wasser, S. K. (eds).p. 33-55. Cambridge University Press, New York.

Lovette, J. C., and Wassre. S. K., 1993. Biogeography and Ecology of the Rain Forests of Eastern Africa. Cambridge University Press, Cambridge, United Kingdom.

Maddison, W. P. and Knowles, L. L., 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. **55**:21-30.

Maddison, W. P., 1997. Gene trees in species trees. Syst. Biol. **46**:523-536.

Maley, J., 1989. Late Quaternary climatic changes in the African rain forest: forest refugia and the major role of sea surface variations. In: Paleoclimatology and paleometeorology: modern and past patterns of global atmospheric transport. Leinen, M. and M. Sarnthein (eds). Kluwer Academic Publishers, Dordrecht.p.585-616.

Maley, J., 1991.The African rain forest vegetation and palaeoenvironments during the Late Quaternary. *Climatic Change* **19**: 79-98.

Maley, J., 1996. The African rain-forest - main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. Essays on the ecology of the Guinea-Congo rain forest (ed. by Alexander, I. J. et al.). Proc R Soc Edinburgh **104B**:31-73.

Maley, J., 2001. The impact of arid phases on the African rain forest through geological history. In: African rain forest ecology and conservation. Weber, W., White, L. J. T., Vedder A. and Naughton-Treves L. (eds). Yale University Press, New Haven and London. p. 68-87.

Mantel, R. M., 1967. The detection of disease clustering and a general regression approach. Cancer Res. **27**: 209-220.

Mark, B. D., 2010. Are lowland rainforests really evolutionary museums? Phylogeography of the green hylia (*Hylia prasina*) in the Afrotropics, *Molecular Phylogenetics and Evolution* **55**: 178-184

Marshall, T. C., Slate, J., Kruuk, L. E. B. and Pemberton, J. M.,1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*,**7**: 639-655.

Matsumura, S. I., Yokoyama, J., Tateishi, Y. and Maki, M., 2006. Intraspecific variation of flower colour and its distribution within a sea lavender, *Limonium wrightii* (Plumbaginaceae), in the northwestern Pacific Islands. J. Plant. Res. **119:** 625-632.

Mayr, E., 1963. Animal Species and Evolution. Harvard University Press, Cambridge, MA.

Mayr, E. and O'Hara, R. J., 1986. The biogeographical evidence supporting the Pleistocene forest refuge hypothesis. *Evolution* **40**: 55-67.

McCormack, J. E., Huang, H. and Knowles, L. L. 2000. Maximum likelihood estimates of species trees, hoe accuracy of phylogenetic inference depends upon the divergence history and sampling design. Syst. Biol. **58**:501-508

McGuire, G. and Wright, F., 2000. TOPAL 2.0: improved detection of mosaic sequences within multiple alignments. *Bioinformatics*. **16**:130-134.

Melo-Ferreira, J., Boursot, P., Suchentrunk, F., Freeland, N. and Alves, P. C., 2005. Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in Northern Iberia. *Molecular Ecology*, **14**: 2459.

Michael, S. Roy., Renate Sponer, and Jon Fjeldså, 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.

Moll, E. J. and White, F., 1978. The Indian Ocean Coastal Belt. In: Werger, M.J.A. (ed.), Biogeography and ecology of Southern Africa, pp. 561-598. Dr. W.Junk, The Hague.

Moodley, Y. and Bruford, M. W., 2007 Molecular biogeography: towards an integrated framework for conserving pan-African biodiversity. PLoS ONE **2**:454

Moreau, R. E., 1966. The Bird Faunas of Africa and its Islands. Academic Press, New York.

Moritz, C. and Faith, D. P.,1998. Comparative phylogeography and the identification of genetically divergent areas for conservation. *Molecular Ecology*, **7:** 419-429.

Moritz, C., Patton, J. L., Schneider, C. J. and Smith, T. B., 2000. Diversification of rainforest faunas: an integrated molecular approach. Annu. Rev. Ecol. Syst. **31:** 533-563.

Moritz, C., 2000. Strategies to Protect Biological Diversity and the Evolutionary Processes That Sustain It, *Systematic Biology*, **51**(2):238-254.

Moore, A. E., Cotterill, F. P. D., Main, M. P. L. and Williams, H. B. 2007. The Zambezi river. In: Large Rivers: Geomorphology and Management, Gupta, A. (ed.) p. 311-333. John Wiley and Sons, Ltd.

Mossel, E. and Vigoda, E., 2005. Phylogenetic MCMC algorithms are misleading on mixtures of trees. *Science*, **309:** 2207-2209.

Moyle, R. G. and Marks, B. D., 2006. Phylogenetic relationships of the bulbuls (Aves: Pycnonotidae) based on mitochondrial and nuclear DNA sequence data. *Molecular Phylogenetics and Evolution*, **40**: 687-695.

Nei, M. and Li, W., 1979. Mathematical model for studying genetics variance in terms of restriction endonucleases. *Proceedings of the National academy of Science* USA, **76**:5269-5273.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press. New York, NY.

Neigel J. E. and Avise J. C., 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In Evolutionary Processes and Theory, ed. E Nevo, S Karlin, pp. 515-34. New York: Academic

Nelson, G. and Platnick, N. I., 1981. Systematics and Biogeography. Columbia University Press, New York, NY.

Nicholson, S., 1994. Recent rainfall fluctuations in Africa and their relationship to past conditions over the continent. *Holocene* **4:**121-31.

Nichols R., 2001. Gene trees and species trees are not the same. Trends Ecol. Evol. **16**:358-64l.

Noonan, B. P., and Gaucher, P., 2006. Refugial isolation and secondary contact in the dyeing poison frog *Dendrobates tinctorius*. Mol. Ecol. 15: 4425-4435.

Norman, J. A., Christidis, L., Westerman, M. and Hill, A. R., 1998. Molecular analysis confirms the species status of the Christmas Island hawk-owl (*Ninox natalis*). *Emu* **98**: 197- 208.

Njabo, N. Y., Bowie, R. C. K. and Sorenson, M. D., 2008. Phylogeny, biogeography and taxonomy of the African wattle-eyes (Aves: Passeriformes: Platysteiridae). Mol. Phylogenet. Evol. **48**:136-149.

Nguembock, B., Cibois, A., Bowie, R. C. K., Cruaud, C. and Pasquet, E., 2009. Phylogeny and biogeography of the genus Illadopsis (Passeriformes: Timaliidae) reveal the complexity of diversification of some African taxa, J. Avian Biol. **40**: 113__125.

Nyblade, A. A. and Robinson ,S. W., 1994. The African Superswell, Geophys. Res. Lett., **21**(9): 765-768.

Nylander, J. A. A., 2004. MrModeltest, a program to evaluate the fit of several models of evolution to a given data and unrooted tree (version 2.2). Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Sweden. Available from: <http://www.abc.se/~nylander/>.

Nylander, J. A. A., Wilgenbusch, J. C., Warren D. L. and Swofford, D. L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*. **24**(4): 581-583.

Oberholser, H. C., 1905. The avian genus Bleda Bonaparte and some of its allies. Smiths. Misc. Coll. **48:** 149-172.

Oliveros, C. H. and Moyle, R. G., 2010. Origin and diversification of Philippine bulbuls. *Molecular Phylogenetics and Evolution*. **54:** 822-832.
Olson, S. L., 1989. Preliminary systematic notes on some old world passerines. Rivista Italiana di Ornitologia, Milano, **59:** 183- 195.

Outlaw, R. K., Voelker, G. and Outlaw, D. C., 2007. Molecular systematics and historical biogeography of the rock-thrushes (Muscicapidae: Monticola). *The Auk,* **124**: 561-57.

Outlaw, D. C. and Voelker, G., 2008. Pliocene climatic change in insular Southeast Asia as an engine of diversification in Ficedula flycatchers. J. Biogeogr. **35:** 739-752.

Palumbi, S. R. and Baker, S. C., 1994. Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molecular ecology and evolution*. **11(3):** 426-435.

Palumbi, S. R. and Baker, S. C., 1996. Nuclear genetic analysis of population structure and genetic variation using intron primers. In: Molecular genetic approaches in conservation. T.B. Slith and R.K. Wayne (eds). Oxford University Press, pp25-37.

Pamilo P. and Nei M., 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. **5:**568-583.

Partridge, T. C., Wood, B. A. and deMenocal, P. B., 1995. The influence of global climatic change and regional uplift in large-mammalian evolution in East and Southern Africa. In: Vrba E.S., Denton G.H., Partridge T.C., Burckle L.H. (eds). Paleoclimate and evolution, with emphasis on Human origins. Yale University Press, New Haven. p. 331-355.

Partridge, T. C., deMenocal, P. B., Lorentz, S. A., Parker, M. J. and Vogel, J. C., 1997. Orbital forcing of climate over South Africa: a 200 000 year rainfall record from the Pretoria Saltpan. *Quaternary Science Reviews,* **16**: 1-9.

Pasquet, E., Han, L. X., Khobkhet, O. and Cibois, A., 2001. Towards a molecular systematics of the genus Criniger, and a preliminary phylogeny of the bulbuls (Aves: Passeriformes: Pycnonotidae). *Zoosystema*, **23:** 857-863.

Patton, J. L. and Smith, M. F., 1992. mtDNA phylogeny and Andean mice: a test of diversification across ecological gradients. *Evolution*, **46:** 174-183.

Patton, J. L, Da Silva, M. N. F. and Malcolm, J. R., 1994, Gene Genealogy and Differentiation Among Arboreal Spiny rats (Rodentia: Echimyidae) of the Amazon Basin: A Test of the Riverine Barrier Hypothesis. *Evolution*,**48**(4): 1314-1323

Pavlova, A., Zink, R. M., Drovetski, S. V., Redkin, Y. and Rower, S., 2003. Phylogeography patterns in *Motacilla flava* and *Motacilla citreola*: species limits and population history. *The Auk*, **120**:744-758.

Per Alström, Per G. P. Ericson, Urban Olsson and Per Sundberg., 2006. Phylogeny and classification of the avian superfamily Sylvioidea, Mol. Phylogenet. Evol. **29**: 126-138.

Pereira, S. L and Baker, A. J., 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. Mol. Biol. Evol. **23(9)**: 1731-1740.

Peters, J. L., Zhuravlev, Y., Fefelov, I. Anne Logie. and Kevin E. Omland., 2007. Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (Anas spp.). *Evolution*, **61**: 1992-2006.

Plumptre, A. J., Behangana, M., Davenport, T. R. B., Kahindo, C., Kityo, R., Ndomba, E., Nkuutu, D., Owiunji, I., Ssegawa, P. and Eilu, G., 2003. The Biodiversity of the Albertine Rift. Albertine Rift Technical Reports No. 3: 105. Available from: <www.albertinerift.orgnarift-publications>.

Plumptre, A. J., Davenport, T. R. B., Behanganac, M., Kityoc, R., Eiluc, G., Ssegawac, P., Ewangoa, C., Meirted, D., Kahindo, C., Herremansd, M., Peterhanse, J. K., Pilgrimg, J. D., Wilsonh, M., Languyi, M. and Moyer, D., 2007.The biodiversity of the Albertine Rift, *Biological Conservation*, **134:**178 -194.

Posada, D., Crandall, K. A. and Templeton, A. R., 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Mol. Ecol. **9**: 487-488.

Posada, D. and Crandall, K., 2001. Intraspecific gene genealogies: trees grafting into networks. Trends Ecol Evol. **16**:37-45.

Poynton, J. C. and Boycot, R.C., 1996. Species turnover between Afromontane and eastern African lowland faunas: patterns shown by amphibians. *Journal of Biogeography* **23**: 669-680

Prell, W. L., Hutson, W. H., Williams, D. F., Bé, A. W. H., Geitzenauer, K. and Molfino, B.,1980. Surface circulation of the Indian Ocean during the last glacial maximum, approximately 18,000 yr B.P. *Quaternary Research*.**14 (3):** 309-336.

Prigogine, A., 1975. Etude taxinomique [sic] des populations de *Phyllastrephus flavostriatus* de l'Afrique Centrale et description de deux nouvelles races de la République du Zaïre. Gerfaut, **63:** 219-234.

Prigogine, A., 1987. Disjunctions of montane forest birds in the Afrotropical region. Bonn. Zool. Beitr. **38**:195-207.

Prigogine, A.,1988. Speciation pattern of birds in the Central African Forest Refugia and their relationship with other refugia. Acta XIX Congr. Int.Orn. 2537-46.

Prigogine, A., 1969. Trois nouveaux Oiseaux du Katanga. Rev. Zool. Bot. Afr., **79**: 110-116.

Primmer, C. R., Borge, T., Lindell J. and Sætre, G. P., 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.

Pritchard, J. K., Stephens, M. and Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics*. **155**. 945-959.

Rambaut, A. and Drummond, A. J., 2007. Tracer v 1.4. Available from: http://beast.bio.ed.ac.uk/Tracer.

Rannala, B. and Yang, Z., 2008. Phylogenetic inference using whole genomes. Annu. Rev. Genomics Hum. Genet. **9**:217-231.

Raymond, M. L. and Rousset, F., 1995. An exact test for population differentiation. *Evolution*. **49**: 1280-1283.

Remsen, J. V. Jr., 1984. High incidence of "leapfrog" pattern of geographic variation in Andean birds: implications for the speciation process. *Science* **224**: 171-173.

Rice, W. R., 1989 Analyzing tables of statistical tests. *Evolution*. **43**: 223-225.

Rice, W. R. and Hostert, E. E., 1993. Perspective: laboratory experiments on speciation: what have we learned in forty years. *Evolution* **47:**1637-53.

Robertson, A. V., Ramsden, C., Niedzwiecki, J., Jinzhong, Fu. and James P. Bogart. 2006. An unexpected recent ancestor of unisexual Ambystoma. *Molecular Ecology*. **15:** 3339-3351.

Robinson, T. J. and Matthee, C. A., 1999, Molecular genetic relationships of the extinct ostrich, Struthio camelus syriacus: consequences for ostrich introductions into Saudi Arabia, *Animal Conservation* **2:**165-171.

Roff, D. A., Bentzen, P., 1989. The statistical analysis of mitochondrial DNA polymorphisms: chi-squared and the problem of small sample sizes. Mol. Biol. Evol. **6:** 539-545.

Rokas A., Williams B. L., King N. and Carroll S. B., 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature*. **425:** 798-804.

Ronquist, F. and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:** 1572-15.

Ronquist, F., 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Syst Biol **46:**195-203.

Rosendahl, B. R., 1987, Architecture of Continental Rifts with Special Reference to East Africa, *Annual Review of Earth and Planetary Sciences*. **15**:445

Roy, M., 1997. Recent diversification in African greenbuls (Pycnonotidae: Andropadus) supports a montane speciation model. Proc. R. Soc. Lond. **264**: 1337-1344.

Roy, M. S, Arctander, P. and Fjeldså, J,. 1998. Speciation and taxonomy of montane greenbuls of the genus *Andropadus* (Aves: Pycnonotidae), *Steenstrupia*, **24** :51-66.

Roy, M. S., Cardoso da Silva, J. M., Arctander, P., García-Moreno, J. and Fjeldså, J., 1997.The speciation of South American and African birds in montane regions. In Avian molecular evolution and systematics Mindell D.P 325-343. Eds. San Diego:Academic Press.

Roy, M. S., Sponer, R. and Fjeldså, J., 2000. 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.

Ruvolo, M., 1997. Molecular phylogeny of hominoids: inference from muliple independent DNA sequence data sets. Mol. Biol. Evol. **14**:248-265.

Ryan, P. G. and Bloomer, P.,1999. The Long-Billed Lark complex: A species mosaic in southwestern Africa. *Auk* **116**:194-208.

Ryan, P. G., Hood, I., Bloomer, P., Komen, J. and Crowe, T., 1999. Barlow's Lark: a new species in the Karoo Lark *Certhilauda albescens* complex. *Ibis* **140**: 605-619.

Sæther B. E., Russell Lande, R., Engen, S., Weimerskirch, H., Lillegård, M., Altwegg R., Becker P. H., Bregnballe T., Brommer J. E., McCleery R. H., Merilä, J., Nyholm, E., Rendell, W., Robertson R. R., Tryjanowski, P. and Visser M. E., 2005. Generation time and temporal scaling of bird population dynamics. *Nature*, **436**: 99-102.

Sambrook, J., Fritsch, E. F. and Manitas. T., 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press.

Sánchez-González, L. A., Navarro-Sigüenza, A. G., Peterson, T. A. and García-Moreno, J., 2007. Taxonomy of Chlorospingus ophthalmicus in Mexico and northern Central America, Bull. B.O.C. 127(1)

Schierup, M.H. and Hein J., 2000. Consequences of recombination on traditional phylogenetic analysis. *Genetics*. **156**: 879-89

Schiùtz, A., 1967. The tree frogs (Rhacophoridae) of West Africa. Spolia Zool. Mus. Haun. **25**: 1-346.

Schneider, C., Smith, T. B., Larison, B. and Moritz, C., 1999. A test of alternative models of diversification in tropical rainforests: ecological gradients vs. rainforest refugia. Proc. Natl. Acad. Sci. U.S.A. **96**: 13869-13873.

Schneider, C. J. and Moritz, C., 1999. Rainforest refugia and evolution in Australia's Wet Tropics. Proc. R. Soc. Lond. Ser. B **266**: 191-196.

Schneider, S., Excoffier L., 1999, Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* **152**: 1079-1089.

Schneider, S., Roessli, D. and Excoffier, L., 2000. Arlequin ver. 2.000: A Software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva.

Schneider, S., Roessli, D. and Excoffier, L., 2005. Arlequin: A Software for Population Genetics Data Analysis, Ver. 3.1. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Schodde, R. and Calaby, J. H.,1972. The biogeography of the Australo-Papuan bird and mammal faunas in relation to Torres Strait. In Bridge and barrier: the natural and cultural history of Torres Strait (ed. D. Walker). pp255-300.

Schulenberg, T. S., 1993. Phylogeny of the Vangidae: inference from mitochondrial DNA. Proceedings of the VIII Pan-African Ornithological Congress, 23-28.

Sgariglia, E. A. and Burns, K.J., 2003. Phylogeography of the California Thrasher (*Toxotoma redivivum*) based on nested-clade analysis of Mitochondrial DNA variation. *The Auk*. **120**: 346-361.

Sibley, C. G., 1970. A comparative study of the egg-white proteins of passerine birds. Bull. Peabody Mus. Nat. Hist., Yale Univ. **32**: 1-131.

Sibley, C. G., and Ahlquist,J. E., 1990. "Phylogeny and Classification of the Birds of the world. Yale University Press, New Haven, CN.

Sibley, C. G. and Monroe, B. L.1996. Distribution and Taxonomy of Birds of the World. Yale University Press, New Haven.

Sheldon, F. H. and Gill, F. B., 1996. A Reconsideration of Songbird Phylogeny, with Emphasis on the Evolution of Titmice and their Sylvioid Relatives, *Systematic biology*, **45** (4): 473-495

Smith T. B., Bruford M. W. and Wayne R. K.,1993. The Preservation of Process: The Missing Element of Conservation Programs, *Biodiversity Letters*, **1**(6): 164-167.

Smith, T. B., Wayne, R. K., Girman, D. J. and Bruford, M. W. 1997. A role for ecotones in generating rainforest biodiversity. *Science* **276**: 1855-1857.

Smith, T. B., Calsbeek, R., Wayne, R. K., Holder, K. H., Pires, D. and Bardeleben. C., 2005.Testing alternative mechanisms of evolutionary divergence in an African rainforest passerine bird. *Journal of Evolutionary Biology* .**18**: 257-268.

Sinclair, I and Ryan, P., 2003. Birds of Africa south of the Sahara. Princeton University Press, New Jersey.

Sinclair, I. and Ryan, P., 2004. Birds of Africa South of the Sahara: A Comprehensive Illustrated Field Guide. Struik Publishers (Pty) Ltd.

Snow, D. W., (ed.). 1978. An Atlas of Speciation in African Non-passerine Birds. British Museum (Natural History), London, U.K.

Spicer, G. S. and Dunipace, L., 2004. Molecular phylogeny of songbirds (Passeriformes) inferred from mitochondrial 16S ribosomal RNA gene sequences. Mol. Phylogenet. Evol. **30**: 325-335.

Spinks, P. Q., Shaffer, H. B. 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.

Spinks, P. Q., Thomson, R. C. and Shaffer, H. B., 2010. Nuclear gene phylogeography reveals the historical legacy of an ancient inland sea on lineages of the western pond turtle, *Emys marmorata* in California, *Molecular Ecology*.**19:** 542-556.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* **22**: 2688-2690.

Stamatakis, A., Hoover, P. and Rougemont, J,. 2008. A rapid bootstrap algorithm for the RAxML web-servers. Syst. Biol. **75**: 758-771.

Stephens, M. and Donnelly, P., 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. **73**:1162-1169.

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

Stankiewicz, J. and De Wit. M., 2006. The east flowing Congo river. J. Afr. Earth Sci. **44**: 75-84.

Stattersfield, A. L., Crosby, M. J., Long, A. J., and Wege, D. C. 1998. Endemic bird areas of the worls: priorities for biodiversity conservation. Birdlife international. Cambridge, UK.

Swofford, D. L., 2002. PAUP*: Phylogeny Analysis using Parsimony (*and other methods), Version 4.0b10. Sinauer Association, Sutherland, MA.

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G. and Cosson, J. F., 1998. Comparative phylogeography and postglacial routes in Europe. Mol. Ecol. **7**: 453 - 464.

Tajima, F., 1989. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics*. **123**: 585-595.

Templeton, A. R., 1998. Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*. **7:** 381-397.

Templeton, A. R., 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*. **13**: 789-809.

Templeton, A. R., Routman, E. and Phillips, C. A., 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondial DNA haplotypes in the tiger salamander, *Ambystoma tigrinu*. *Genetics*. **140**: 767-782.

Templeton, A. R., Sing, C. F., 1993. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* **134**: 659-669.

Templeton, A. R., 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*. **13**: 789-809.

Timberlake, J. R.,1998. Biodiversity of the Zambezi Basin wetlands: review and preliminary assessment of available information - Phase 1. Consultancy report for IUCN ROSA, Harare.

Occasional Publications in Biodiversity No. 3. The Zambezi Society/Biodiversity Foundation for Africa, Harare/Bulawayo.

Tribsch, A., Schonswetter, P., 2003. Pattern of endemism and comparative phylogeography confirm paleo-environmental evidence for Pleistocene refugia in the eastern Alps. *Taxon* **52**: 447-497.

Tyson, P. D., 1986. Climatic change and variability in southern Africa, 220 pp. Oxford University Press, Cape Town.

Urban, E. K., Fry, C. H. and Keith, S., 1997. The birds of Africa vol. V. London: Academic Press.

Uys, C., Hamer, M. and Slotow, R., 2009.Turnover in flightless invertebrate species composition over different spatial scales in Afrotemperate forest in the Drakensberg, South Africa. *African Journal of Ecology*. **47**: 341-351.

Van Oosterhout C. V., Hutchinson, W. F., Wills D. P. M. and Shipley, P., 2004. micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology*. **4:** 535-538.

Van Zinderen, Bakker, E. M., 1978. *Quaternary vegetation changes in southern Africa*, p. 133-143. In: M. J. A. Werger (ed.). Biogeography and ecology of Southern Africa. Dordrecht. Dr. W. Junk Publishers.

Vanzolini, P. E., Williams, E. E., 1981. The vanishing refuge: a mechanism for eco-geographic speciation. Pap. Avulsos. Zool. **34**: 251-55.

Voelker. G., 1999a. Dispersal, Vicariance, and Clocks: Historical Biogeography and Speciation in a Cosmopolitan Passerine Genus (Anthus: Motacillidae). *Evolution*. **53**: 1536-1552

Voelker, G., 1999b. Molecular evolutionary relationships in the avian genus Anthus (Pipits: Motacillidae). *Molecular Phylogenetics and Evolution*. **11**: 84-94.

Voelker, G., 2002. Systematics and historical biogeography of wagtails (Aves: Motacilla): Dispersal versus vicariance revisited. *Condor* **104**: 725-739.

Voelker, G., Outlaw, R. K. and Bowie, R.C.K., 2010. Pliocene forest dynamics as a primary driver of African bird speciation. *Global Ecology and Biogeography*. **19**: 111-121.

Wallace, A. R., 1852. On the monkey of the Amazon. Proc. Zool. Soc. London **20**: 107-110.

Warren, B. H., Bermingham, E., Bowie, R. C. K., Prys-Jones, R. P. and The´ baud, 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. H., Bermingham, E., Prys-Jones, R. P. and Thebaud, C., 2005. Tracking island colonization history and phenotypic shifts in Indian Ocean bulbuls (Hypsipetes: Pycnonotidae). *Biological Journal of the Linnean Society*. **85**: 271-287.

Weir, J.T. and Schluter, D., 2008. Calibrating the avian molecular clock. Mol. Ecol. **17**: 2321-2328.

White, F., 1979. The Guineo-Congolian Region and its relationships to other phytochoria. Bulletin Jardin Botanique Nationale Belgtque. **49**: 11-55.

White, F., 1983. The Vegetation of Africa. Paris: UNESCO.

Wiens, J. J. and Donoghue, M. J., 2004. Historical biogeography, ecology, and species richness. Trends Ecol. Evol. **19**: 639 - 44.

Wiley, E. O., 1988. Vicariance Biogeography. Annu Rev Ecol Syst **19**:513-542.

Wilgenbusch, J. C., Warren, D. L. and Swofford, D. L., 2004. AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.

Zhang, D. X. and Hewitt, G. M., 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*. **12**: 563-584.

Zink, R. M. and Remsen, J. V. JR., 1986. Evolutionary processes and patterns of geographic variation in birds. *Current Ornithology* **4**:1-69.

Zuccon, D. and Ericson Per G. P., 2010. The phylogenetic position of the Black-collared Bulbul *Neolestes torquatus*. *Ibis*.**152**: 386-392.

Zwickl, D., 2006. Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biologiical Sequence Datasets under the Maximum Likelihood Criterion. PhD thesis, University of Texas at Austin.

Zwickl, D., 2008. GARLI, a program that performs phylogenetic searches on aligned sequence datasets using the maximum-likelihood criterion (version 0.96). Available from: <http://garli.nescent.org/> (accessed 13.18.2009).

APPENDIX

Appendix 3.1. Detailed information of the samples used in chapter 3. FMNH = Field Museum of Natural History, Illinois, Chicago, USA; MOM = Museums of Malawi, Blantyre, Malawi; ZMUC – Zoological Museum of University of Copenhagen, Copenhagen, Denmark; MVZ = Museum of Vertebrate Zoology, University of California Berkeley: RMCA =Royal Museum of Central Africa.

Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Sample Type
LGS030	Malawi		FMNH	MLW1013	Tissue
LGS031	Malawi		FMNH	MLW0815	Tissue
LGS033	Malawi		MOM	2005.1.185	Tissue
LGS034	Malawi		FMNH	MLW1108	Tissue
LGS035	Malawi		FMNH	MLW1040	Tissue
LGS036	Malawi		FMNH	MLW0782	Tissue
LGS037	Malawi		FMNH	MLW1148	Tissue
LGS038	Malawi		MOM	2005.1.21	Tissue
LGS042	Malawi	Misuku Hills	FMNH	439207	Tissue
LGS043	Malawi	Misuku Hills	FMNH	439208	Tissue
LGS044	Malawi	Misuku Hills	FMNH	439209	Tissue
LGS047	Malawi	Rumphi, Nyika	FMNH	440579	Tissue
LGS048	Malawi	Rumphi, Nyika	FMNH	440580	Tissue
LGS049	Malawi	Misuku Hills	MOM	2003.2.93	Tissue
LGS056	Malawi		MOM	2005.1.117	Tissue
LGS407	Malawi	Mt. Mulanji	FMNH	447394	Tissue
LGS408	Malawi	Mt. Mulanji	FMNH	447395	Tissue
LGS409	Malawi	Mt. Mulanji	FMNH	447396	Tissue
LGS410	Malawi	Mt. Mulanji	FMNH	447397	Tissue
LGS411	Malawi	Mt. Mulanji	FMNH	447398	Tissue
LGS902	Malawi	Mt. Mulanji	MVZ	RCKB1337 (uncat.)	Tissue
LGS903	Malawi	Ntchisi Forest	MVZ	RCKB1170 (uncat.)	Tissue
LGS904	Malawi	Ntchisi Forest	MOM	2007.2.335	Tissue
LGS905	Malawi	Ntchisi Forest	MOM	2007.2.204	Tissue
LGS906	Malawi	Mt. Mulange	MVZ	RCKB1339 (uncat).	Tissue

Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
LGS045	Mozambique	Zambezia,Murabue	FMNH	438556	Tissue
LGS046	Mozambique	Zambezia,Murabue	FMNH	438557	Tissue
LGS140	Tanzania	Madizini	ZMUC	119699	Tissue
LGS141	Tanzania	Madizini	ZMUC	119677	Blood
LGS142	Tanzania	Madizini	ZMUC	119767	Tissue
LGS143	Tanzania	Kahe forest	ZMUC	120004	Blood
LGS144	Tanzania	Uluguru north	ZMUC	122562	Blood
LGS146	Tanzania	Chita	ZMUC	120948	Tissue
LGS147	Tanzania	Magambo	ZMUC	120929	Blood
LGS148	Tanzania	SE Iringa	ZMUC	123503	Blood
LGS041	Zambia	Kafunglia river, Luapula	FMNH	205629	Blood
LGS042	Zambia	Kafunglia river, Luapula	FMNH	205635	Blood
LGS755	Zambia	Nkanka, Chingi Stream, East	FMNH	205629	Toe-pad
LGS756	Zambia	Lundazi, Mukutu, East	FMNH	205635	Toe-pad
LGS757	Zambia	Lundazi, Kalubi, East	FMNH	205633	Toe-pad
LGS758	Zambia	Lundazi,Boma, East	FMNH	205631	Toe-pad
LGS759	Zambia	Lundazi	FMNH	205630	Toe-pad
LGS760	Zambia	Mwinilunga, NW	FMNH	219789	Toe-pad
LGS761	Zambia	Lundazi,Mwanda	FMNH	205626	Toe-pad
LGS762	Zambia	Lundazi,Boma, East	FMNH	205632	Toe-pad
LGS763	Zambia	Nkanka, Chingi Stream, East	FMNH	205634	Toe-pad
LGS673	DRC	Kinshasa, Kisanga	RMCA	A.116237	Toe-pad

Appendix 4.1 Detailed information of the samples used in chapter 4. FMNH = Field Museum of Natural History, Illinois, Chicago, USA; ZMUC– Zoological Museum of University of Copenhagen, Copenhagen, Denmark; MVZ = Museum of Vertebrate Zoology, University of California Berkeley: Durban =Durban Museum of Natural History.

Extraction no	Subspecies	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
LGS020	*A. i. insularis*	Kenya	Sokoke Forest	RB	K129	Tissue
LGS634	*A. i. insularis*	Kenya	Kilifi	ZMUC	117563	Blood
LGS635	*A. i. insularis*	Kenya	Kilifi	ZMUC	117564	Blood
LGS636	*A. i. insularis*	Kenya	Kilifi	ZMUC	117565	Blood
LGS637	*A. i. insularis*	Kenya	Kilifi	ZMUC	117566	Blood
LGS638	*A. i. insularis*	Kenya	Kilifi	ZMUC	117567	Blood
LGS639	*A. i. insularis*	Kenya	Kilifi	ZMUC	117568	Blood
LGS640	*A. i. insularis*	Kenya	Kilifi	ZMUC	117569	Blood
LGS646A	*A. i. insularis*	Kenya	Kilifi	ZMUC	117562	Blood
LGS780	*A. i. insularis*	Kenya	Kilifi	FMNH	369836	Toe pad
LGS783	*A. i. insularis*	Kenya	Kilifi	FMNH	369835	Toe pad
LGS792	*A. i. insularis*	Kenya	Kilifi	FMNH	369834	Toe pad
LGS793	*A. i. insularis*	Kenya	Kilifi	FMNH	369837	Toe pad
LGS776	*A. i. insularis*	Kenya	Lamu,Manda Id	FMNH	197010	Toe pad
LGS789	*A. i. insularis*	Kenya	Lamu,Manda Id	FMNH	197009	Toe pad
LGS777	*A. i. insularis*	Kenya	Tana River	FMNH	197047	Toe pad
LGS848	*A. i. insularis*	Kenya	Machakcos	FMNH	197049	Toe pad
LGS642	*A. i. insularis*	Tanzania	Kahe forest	ZMUC	120011	Toe pad
LGS778	*A. i. insularis*	Tanzania	Zanzibar	FMNH	196980	Toe pad
LGS781	*A. i. insularis*	Tanzania	Zanzibar	FMNH	196981	Toe pad
LGS771	*A. i. insularis*	Somalia	Serenli	FMNH	196987	Toe pad
LGS772	*A. i. insularis*	Somalia	Serenli	FMNH	196993	Toe pad
LGS774	*A. i. insularis*	Somalia	Serenli	FMNH	196991	Toe pad

Extraction no	Subspecies	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
LGS779	*A. i. insularis*	Somalia	Serenli	FMNH	196992	Toe pad
LGS782	*A. i. insularis*	Somalia	Serenli	FMNH	196986	Toe pad
LGS784	*A. i. insularis*	Somalia	Serenli	FMNH	196985	Toe pad
LGS796	*A. i. insularis*	Somalia	Serenli	FMNH	196984	Toe pad
LGS791	*A. i. insularis*	Somalia	Serenli	FMNH	196989	Toe pad
LGS022	*A. i. importunus*	South Africa	Kei Motuh, EC	MVZ	Uncat.	Tissue
LGS023	*A. i. importunus*	South Africa	Kei Mouth, EC	MVZ	Uncat.	Tissue
LGS024	*A. i. importunus*	South Africa	Kei Mouth, EC	MVZ	Uncat.	Tissue
LGS025	*A. i. importunus*	South Africa	Greaf Reinet, EC	MVZ	Uncat.	Tissue
LGS026	*A. i. importunus*	South Africa	Greaf Reinet, EC	MVZ	Uncat.	Tissue
LGS027	*A. i. importunus*	South Africa	Greaf Reinet, EC	MVZ	Uncat.	Tissue
LGS028	*A. i. importunus*	South Africa	Greaf Reinet, EC	MVZ	Uncat.	Tissue
LGS029	*A. i. importunus*	South Africa	Kei Mouth, EC	MVZ	Uncat.	Tissue
LGS050	*A. i. importunus*	South Africa	Port Elizabeth	MVZ	Uncat.	Blood
LGS051	*A. i. importunus*	South Africa	Port Elizabeth	MVZ	Uncat.	Blood
LGS052	*A. i. importunus*	South Africa	Port Elizabeth	MVZ	Uncat.	Blood
LGS053	*A. i. importunus*	South Africa	Port Elizabeth	MVZ	Uncat.	Blood
LGS054	*A. i. importunus*	South Africa	Port Elizabeth	MVZ	Uncat.	Blood
LGS055	*A. i. importunus*	South Africa	Port Elizabeth	MVZ	Uncat.	Blood
LGS079	*A. i. importunus*	South Africa	Port Elizabeth	MVZ	Uncat.	Blood
LGS080	*A. i. importunus*	South Africa	Port Elizabeth	MVZ	Uncat.	Blood
LGS081	*A. i. importunus*	South Africa	Hulhluwe	MVZ	Uncat.	Blood
LGS086	*A. i. importunus*	South Africa	Hulhluwe	MVZ	Uncat.	Blood
LGS087	*A. i. importunus*	South Africa	Eshowe	MVZ	Uncat.	Tissue
LGS088	*A. i. importunus*	South Africa	Eshowe	MVZ	Uncat.	Tissue
LGS612	*A. i. importunus*	South Africa	Retirement F, KZN	MVZ	Uncat.	Blood
LGS613	*A. i. importunus*	South Africa	Greaf Reinet	MVZ	Uncat.	Blood

Extraction no	Subspecies	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
LGS614	*A. i. importunus*	South Africa	Kei Mouth	MVZ	Uncat.	Blood
LGS775	*A. i. importunus*	Swaziland	Balegane	FMNH	285330	Toe pad
LGS785	*A. i. importunus*	Swaziland	Balegane	FMNH	285329	Toe pad
LGS786	*A. i. importunus*	Swaziland	Balegane	FMNH	285328	Toe pad
LGS788	*A. i. importunus*	Swaziland	Balegane	FMNH	285327	Toe pad
LGS790	*A. i. importunus*	Swaziland	Balegane	FMNH	285326	Toe pad
LGS801	*A. i. importunus*	South Africa	EntabeneFR, Louis Trichard, LP	Durban	29948	Toe pad
LGS806	*A. i. importunus*	South Africa	Downs-Tzaneen, NE Transval	Durban	29947	Toe pad
LGS807	*A. i. importunus*	South Africa	Noordhoek, PE	Durban	31140	Toe pad
LGS814	*A. i. importunus*	South Africa	Westville, KZN	Durban	34417	Toe pad
LGS823	*A. i. importunus*	South Africa	RedH, PE	Durban	34001	Toe pad
LGS824	*A. i. importunus*	South Africa	Nkandhla, KZN	Durban	19587	Toe pad
LGS829	*A. i. importunus*	South Africa	Newington, AE Transval	Durban	6588	Toe pad
LGS830	*A. i. importunus*	South Africa	Nkandhla, KZN	Durban	19585	Toe pad
LGS831	*A. i. importunus*	South Africa	Nkandhla, KZN	Durban	19586	Toe pad
LGS840	*A. i. importunus*	South Africa	EntabeneFR,Louis Trichard, LP	Durban	29852	Toe pad
LGS841	*A. i. importunus*	South Africa	EntabeneFR,Louis Trichard, LP	Durban	29853	Toe pad
LGS842	*A. i. importunus*	South Africa	WoodbushFR Tzaneen, NE Transval	Durban	29854	Toe pad
LGS843	*A. i. importunus*	South Africa	Downs-Tzaneen, NE Transval	Durban	29945	Toe pad
LGS845	*A. i. importunus*	South Africa	Ballito, KZN,	Durban	37003	Toe pad
LGS850	*A. i. importunus*	South Africa	Gram'stown, EC	JF	518	Blood
LGS851	*A. i. importunus*	South Africa	Gram'stown, EC	JF	523	Blood
LGS854	*A. i. importunus*	South Africa	Kasuoga, EC	JF	600	Blood
LGS855	*A. i. importunus*	South Africa	Kasuoga, EC	JF	615	Blood
LGS856	*A. i. importunus*	South Africa	Kasuoga, EC	JF	616B	Blood
LGS857	*A. i. importunus*	South Africa	Kasuoga, EC	JF	616	Blood
LGS858	*A. i. importunus*	South Africa	Kasuoga, EC	JF	625	Blood

Extraction no	Subspecies	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
LGS859	*A. i. importunus*	South Africa	Kasuoga, EC	JF	639	Blood
LGS860	*A. i. importunus*	South Africa	Kasuoga, EC	JF	640	Blood
LGS861	*A. i. importunus*	South Africa	Kasuoga, EC	JF	643	Blood
LGS862	*A. i. importunus*	South Africa	Kasuoga, EC	JF	647	Blood
LGS863	*A. i. importunus*	South Africa	Kasuoga, EC	JF	655	Blood
LGS865	*A. i. importunus*	South Africa	Morgan bay, EC	JF	723	Blood
LGS866	*A. i. importunus*	South Africa	Morgan bay, EC	JF	735	Blood
LGS867	*A. i. importunus*	South Africa	Morgan bay, EC	JF	747	Blood
LGS868	*A. i. importunus*	South Africa	Morgan bay, EC	JF	758	Blood
LGS641	*A. i. hypoxanthus*	Tanzania	Nambiga F	ZMUC	136535	Blood
LGS773	*A. i. hypoxanthus*	Mozambique	Dondo F	FMNH	282962	Toe pad
LGS800	*A. i. hypoxanthus*	Mozambique	Chiroma, Shire R S	Durban	6566	Toe pad
LGS802	*A. i. hypoxanthus*	Mozambique	Chiroma, Shire R S	Durban	6567	Toe pad
LGS808	*A. i. hypoxanthus*	Mozambique	Chiroma, Shire R S	Durban	6559	Toe pad
LGS812	*A. i. hypoxanthus*	Mozambique	Chiroma, Shire R S	Durban	6560	Toe pad
LGS813	*A. i. hypoxanthus*	Mozambique	Chiroma, Shire R S	Durban	6563	Toe pad
LGS819	*A. i. hypoxanthus*	Mozambique	Mandigo	Durban	19889	Toe pad
LGS833	*A. i. hypoxanthus*	Mozambique	Mandigo	Durban	19888	Toe pad
LGS835	*A. i. hypoxanthus*	Mozambique	Inhaminga	Durban	23619	Toe pad
LGS836	*A. i. hypoxanthus*	Mozambique	Gorangosa	Durban	24775	Toe pad
LGS799	*A. i. oleaginous*	Mozambique	Inhaca Island, Delagoa Bay	Durban	6579	Toe pad
LGS804	*A. i. oleaginous*	South Africa	Koni Bay, KZN	Durban	6584	Toe pad
LGS816	*A. i. oleaginous*	Mozambique	Sul Do Save	Durban	6574	Toe pad
LGS821	*A. i. oleaginous*	Mozambique	Inhambane	Durban	26746	Toe pad
LGS822	*A. i. oleaginous*	South Africa	Koni Bay, KZN	Durban	6585	Toe pad
LGS827	*A. i. oleaginous*	Mozambique	Sul Do Save	Durban	6570	Toe pad
LGS828	*A. i. oleaginous*	Mozambique	Sul Do Save	Durban	6576	Toe pad

Subspecies	Extraction no	Country of origin	Collection Locality	Source	Museum catalog number	Type of sample
LGS837	*A. i. oleaginous*	Mozambique	Inhambane	Durban	26743	Toe pad
LGS838	*A. i. oleaginous*	Mozambique	Inhambane	Durban	26744	Toe pad
LGS839	*A. i. oleaginous*	Mozambique	Inhambane	Durban	26745	Toe pad
LGS846	*A. i. oleaginous*	South Africa	KZN, South Africa	Durban	38709	Toe pad

Appendix 5.1 Detailed information of the samples used in chapter 5. FMNH = Field Museum of Natural History, Illinois, Chicago, USA; ZMUC = Zoological Museum of University of Copenhagen, Copenhagen, Denmark; MVZ = Museum of Vertebrate Zoology, University of California Berkeley; Durban =Durban Museum of Natural History; RMCA =Royal Museum of Central Africa; MOM = Museums of Malawi, Blantyre, Malawi.

Subspecies	Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
Phyllastrephus f. flavostriatus	LGS001	South Africa	Ngoye Forest	MVZ	BE1814	Blood
Phyllastrephus f. flavostriatus	LGS004	South Africa	Ngoye Forest	MVZ	BD27615	Blood
Phyllastrephus f. flavostriatus	LGS005	South Africa	Ngoye Forest	MVZ	BD17014	Blood
Phyllastrephus f. flavostriatus	LGS077	South Africa	Ngoye Forest	MVZ	RSA####	Tissue
Phyllastrephus f. flavostriatus	LGS646	Zimbabwe	Lunitu river	Durban	29189	Toe pad
Phyllastrephus f. flavostriatus	LGS647	Zimbabwe	Seapleford F.	Durban	21476	Toe pad
Phyllastrephus f. flavostriatus	LGS649	South Africa	NE Transvaal	Durban	30015	Toe pad
Phyllastrephus f. flavostriatus	LGS650	South Africa	NE Transvaal	Durban	30014	Toe pad
Phyllastrephus f. flavostriatus	LGS651	South Africa	NE Transvaal	Durban	29828	Toe pad
Phyllastrephus f. flavostriatus	LGS652	South Africa	Port St Johns, Eastern Cape	Durban	6516	Toe pad
Phyllastrephus f. flavostriatus	LGS666	South Africa	Ngoye Forest	Durban	19450	Toe pad
Phyllastrephus f. flavostriatus	LGS671	South Africa	Ngoye Forest	Durban	19449	Toe pad
Phyllastrephus f. flavostriatus	LGS679	Zimbabwe	Selinda	FMNH	265386	Toe pad
Phyllastrephus f. flavostriatus	LGS680	Zimbabwe	Selinda	FMNH	265387	Toe pad
Phyllastrephus f. flavostriatus	LGS681	Zimbabwe	Selinda	FMNH	269003	Toe pad
Phyllastrephus f. flavostriatus	LGS742	Zimbabwe	Lunitu river	Durban	29188	Toe pad
Phyllastrephus f. flavostriatus	LGS744	Zimbabwe	Seapleford F.	Durban	21474	Toe pad
Phyllastrephus f. flavostriatus	LGS746	Zimbabwe	Banei	Durban	29191	Toe pad
Phyllastrephus f. flavostriatus	LGS749	Zimbabwe	Selinda	Durban	29176	Toe pad
Phyllastrephus f. flavostriatus	LGS750	Zimbabwe	Selinda	Durban	29177	Toe pad
Phyllastrephus f. vincenti	LGS309	Mozambique	Namuli	ZMUC	132356	Blood

Subspecies	Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
Phyllastrephus f. vincenti	LGS370	Mozambique	Zambezia	FMNH	438568	Tissue
Phyllastrephus f. vincenti	LGS412	Malawi	Mulanji Forest reserve	FMNH	447437	Tissue
Phyllastrephus f. vincenti	LGS413	Malawi	Mulanji Forest reserve	FMNH	447438	Tissue
Phyllastrephus f. vincenti	LGS414	Malawi	Mulanji Forest reserve	FMNH	447439	Tissue
Phyllastrephus f. vincenti	LGS415	Malawi	Mulanji Forest reserve	FMNH	447440	Tissue
Phyllastrephus f. vincenti	LGS416	Malawi	Mulanji Forest reserve	FMNH	447441	Tissue
Phyllastrephus f. vincenti	LGS417	Malawi	Mulanji Forest reserve	FMNH	447442	Tissue
Phyllastrephus f. vincenti	LGS418	Malawi	Mulanji Forest reserve	FMNH	447443	Tissue
Phyllastrephus f. vincenti	LGS419	Malawi	Mulanji Forest reserve	FMNH	447444	Tissue
Phyllastrephus f. vincenti	LGS907	Malawi	Mulanji Forest reserve	MVZ	RB1370	Tissue
Phyllastrephus f. uzungwensis	LGS149	Tanzania	Ndundulu Mt.	ZMUC	118794	Tissue
Phyllastrephus f. uzungwensis	LGS150	Tanzania	Ndundulu Mt.	ZMUC	118795	Tissue
Phyllastrephus f. uzungwensis	LGS166	Tanzania	Uhafiwa 1390m	ZMUC	121636	Blood
Phyllastrephus f. uzungwensis	LGS167	Tanzania	Uhafiwa 1390m	ZMUC	131638	Blood
Phyllastrephus f. uzungwensis	LGS169	Tanzania	Udzungwa scarp	ZMUC	122599	Blood
Phyllastrephus f. uzungwensis	LGS177	Tanzania	Uhafiwa 1390m	ZMUC	122599	Blood
Phyllastrephus f. uzungwensis	LGS252	Tanzania	Ndundulu Mt.	ZMUC	118799	Blood
Phyllastrephus f. uzungwensis	LGS271	Tanzania	Uhafiwa 1390m	ZMUC	121647	Blood
Phyllastrephus f. uzungwensis	LGS272	Tanzania	Uhafiwa 1390m	ZMUC	121664	Blood
Phyllastrephus f. uzungwensis	LGS279	Tanzania	Udzungwa scarp	ZMUC		Blood
Phyllastrephus f. uzungwensis	LGS282	Tanzania	Udzungwa scarp	ZMUC		Blood
Phyllastrephus f. uzungwensis	LGS283	Tanzania	Udzungwa scarp	ZMUC	122670	Blood
Phyllastrephus f. uzungwensis	LGS291	Tanzania	Udzungwa scarp	ZMUC	123652	Blood
Phyllastrephus f. uzungwensis	LGS292	Tanzania	Uhafiwa	ZMUC	124074	Blood
Phyllastrephus f. uzungwensis	LGS293	Tanzania	Uhafiwa	ZMUC	124094	Blood
Phyllastrephus f. uzungwensis	LGS708	Tanzania	Udzungwa Mt.	ZMUC	124111	Blood
Phyllastrephus f. uzungwensis	LGS709	Tanzania	Udzungwa Mt.	ZMUC	122959	Blood

Subspecies	Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
Phyllastrephus f. uzungwensis	LGS710	Tanzania	Udzungwa Mt.	ZMUC	124051	Blood
Phyllastrephus f. uzungwensis	LGS711	Tanzania	Udzungwa Mt.	ZMUC	124109	Blood
Phyllastrephus f. uzungwensis	LGS712	Tanzania	Udzungwa Mt.	ZMUC	122962	Blood
Phyllastrephus f. uzungwensis	LGS713	Tanzania	Udzungwa Mt.	ZMUC	124113	Blood
Phyllastrephus f. uzungwensis	LGS714	Tanzania	Udzungwa Mt.	ZMUC	121649	Blood
Phyllastrephus f. uzungwensis	LGS715	Tanzania	Udzungwa Mt.	ZMUC	121668	Blood
Phyllastrephus f. uzungwensis	LGS716	Tanzania	Udzungwa Mt.	ZMUC	121673	Blood
Phyllastrephus f. uzungwensis	CVV604	Tanzania	Wanganemo, Iringa	ZMUC	133048	Blood
Phyllastrephus f. tenuirostris	LGS151	Tanzania	Mount Kanga 600m	ZMUC	116261	Blood
Phyllastrephus f. tenuirostris	LGS152	Tanzania	Mount Kanga 600m	ZMUC	119272	Blood
Phyllastrephus f. tenuirostris	LGS154	Tanzania	Kwizu	ZMUC	119353	Blood
Phyllastrephus f. tenuirostris	LGS159	Tanzania	Amani	ZMUC	120075	Blood
Phyllastrephus f. tenuirostris	LGS160	Tanzania	Mkubego coastal F	ZMUC	121046	Blood
Phyllastrephus f. tenuirostris	LGS161	Tanzania	Mkubego coastal F	ZMUC	121093	Blood
Phyllastrephus f. tenuirostris	LGS162	Tanzania	Lutindi Mt, Nilo	ZMUC	121126	Blood
Phyllastrephus f. tenuirostris	LGS163	Tanzania	Lutindi Mt, Nilo	ZMUC	121172	Blood
Phyllastrephus f. tenuirostris	LGS164	Tanzania	Mafi forest reserve	ZMUC	121200	Blood
Phyllastrephus f. tenuirostris	LGS165	Tanzania	Mafi forest reserve	ZMUC	121213	Blood
Phyllastrephus f. tenuirostris	LGS170	Tanzania	Nguru North Nguu Mts	ZMUC	122688	Blood
Phyllastrephus f. tenuirostris	LGS175	Tanzania	Lulago F, Nguu	ZMUC	123568	Blood
Phyllastrephus f. tenuirostris	LGS176	Tanzania	Lulago F, Nguu	ZMUC	123571	Blood
Phyllastrephus f. tenuirostris	LGS181	Tanzania	Nguru North Nguu Mts	ZMUC	122693	Blood
Phyllastrephus f. tenuirostris	LGS182	Tanzania	Nguru South F, Morogoro	ZMUC	122741	Blood
Phyllastrephus f. tenuirostris	LGS183	Tanzania	Nguru South F, Morogoro	ZMUC	132741	Blood
Phyllastrephus f. tenuirostris	LGS184	Tanzania	Pugu hills	ZMUC	117622	Blood
Phyllastrephus f. tenuirostris	LGS185	Tanzania	Pugu hills	ZMUC	117623	Blood
Phyllastrephus f. tenuirostris	LGS191	Tanzania	Kilindi F Ng Tanga	ZMUC	137656	Blood

Subspecies	Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
Phyllastrephus f. tenuirostris	LGS193	Tanzania	Mtiniko F, Tanga	ZMUC	137685	Blood
Phyllastrephus f. tenuirostris	LGS253	Tanzania	Mhonda 700 m	ZMUC	119224	Blood
Phyllastrephus f. tenuirostris	LGS257	Tanzania	Mhonda 700 m	ZMUC	119225	Blood
Phyllastrephus f. tenuirostris	LGS258	Tanzania	Kwizu	ZMUC	119336	Blood
Phyllastrephus f. tenuirostris	LGS260	Tanzania	Ngarama Forest	ZMUC	119274	Blood
Phyllastrephus f. tenuirostris	LGS261	Tanzania	Kambai Forest	ZMUC	120037	Blood
Phyllastrephus f. tenuirostris	LGS263	Tanzania	Mkubege coastal forest	ZMUC	121064	Blood
Phyllastrephus f. tenuirostris	LGS266	Tanzania	Lutindi, Mt. Nilo	ZMUC	121171	Blood
Phyllastrephus f. tenuirostris	LGS267	Tanzania	Mafi Forest Reserve	ZMUC	121262	Blood
Phyllastrephus f. tenuirostris	LGS268	Tanzania	Mafi Forest Reserve	ZMUC	121266	Blood
Phyllastrephus f. tenuirostris	LGS269	Tanzania	Mafi Forest Reserve	ZMUC	121277	Blood
Phyllastrephus f. tenuirostris	LGS284	Tanzania	Nguru North, Nguu Mts.	ZMUC	122696	Blood
Phyllastrephus f. tenuirostris	LGS285	Tanzania	Nguru North, Nguu Mts.	ZMUC	122698	Blood
Phyllastrephus f. tenuirostris	LGS286	Tanzania	Nguru North, Nguu Mts.	ZMUC	122699	Blood
Phyllastrephus f. tenuirostris	LGS287	Tanzania	Gombero Forest, Nguu	ZMUC	123554	Blood
Phyllastrephus f. tenuirostris	LGS290	Tanzania	Gombero Forest, Nguu	MVZ	B7782	Blood
Phyllastrephus f. tenuirostris	LGS299	Tanzania	Kungwe Forest	ZMUC	129522	Blood
Phyllastrephus f. tenuirostris	LGS300	Tanzania	Chamanyani Mvuha Forest Reserve	ZMUC	129609	Blood
Phyllastrephus f. tenuirostris	LGS302	Tanzania	Chamanyani Mvuha Forest Reserve	ZMUC	129612	Blood
Phyllastrephus f. tenuirostris	LGS303	Tanzania	Kungwe Forest	ZMUC	131593	Blood
Phyllastrephus f. tenuirostris	LGS304	Tanzania	Kungwe Forest	ZMUC	131601	Blood
Phyllastrephus f. tenuirostris	LGS305	Tanzania	Kungwe Forest	ZMUC	131603	Blood
Phyllastrephus f. tenuirostris	LGS306	Tanzania	Kungwe Forest	ZMUC	131605	Blood
Phyllastrephus f. tenuirostris	LGS311	Tanzania	Kanga Forest Reserve	ZMUC	132706	Blood
Phyllastrephus f. tenuirostris	LGS312	Tanzania	Kanga Forest Reserve	ZMUC	132707	Blood
Phyllastrephus f. tenuirostris	LGS314	Tanzania	Nguru South Forest Reserve	ZMUC	132756	Blood
Phyllastrephus f. tenuirostris	LGS315	Tanzania	Kanga Forest Reserve	ZMUC	134281	Tissue

Subspecies	Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
Phyllastrephus f. tenuirostris	LGS316	Tanzania	Udaha Camp	ZMUC	134358	Tissue
Phyllastrephus f. tenuirostris	LGS317	Tanzania	Udaha Camp	ZMUC	134362	Tissue
Phyllastrephus f. tenuirostris	LGS318	Tanzania	Udaha Camp	ZMUC	134363	Blood
Phyllastrephus f. tenuirostris	LGS320	Tanzania	Nguru South Forest Reserve	ZMUC	136155	Tissue
Phyllastrephus f. tenuirostris	LGS321	Tanzania	Nguru South Forest Reserve	ZMUC	136156	Tissue
Phyllastrephus f. tenuirostris	LGS322	Tanzania	Nguru South Forest Reserve	ZMUC	136157	Tissue
Phyllastrephus f. tenuirostris	LGS323	Tanzania	Kilindi Forest, Nguu Mts	ZMUC	137644	Blood
Phyllastrephus f. tenuirostris	LGS324	Tanzania	Kilindi Forest, Nguu Mts	ZMUC	137655	Blood
Phyllastrephus f. tenuirostris	LGS367	Tanzania	Tanga	FMNH	356724	Tissue
Phyllastrephus f. tenuirostris	LGS691	Tanzania	Nguru South Forest Reserve	ZMUC	132671	Blood
Phyllastrephus f. tenuirostris	LGS692	Tanzania	Nguru South Forest Reserve	ZMUC	132770	Blood
Phyllastrephus f. tenuirostris	LGS693	Tanzania	Nguru South Forest Reserve	ZMUC	132951	Blood
Phyllastrephus f. tenuirostris	LGS694	Tanzania	Nguru South Forest Reserve	ZMUC	132981	Blood
Phyllastrephus f. tenuirostris	LGS695	Tanzania	Nguru South Forest Reserve	ZMUC	136154	Blood
Phyllastrephus f. tenuirostris	LGS696	Tanzania	Nguru South Forest Reserve	ZMUC	136204	Blood
Phyllastrephus f. tenuirostris	LGS697	Tanzania	Nguru South Forest Reserve	ZMUC	136206	Blood
Phyllastrephus f. tenuirostris	LGS698	Tanzania	Nguru South Forest Reserve	ZMUC	136217	Blood
Phyllastrephus f. tenuirostris	LGS699	Tanzania	Nguru South Forest Reserve	ZMUC	136222	Blood
Phyllastrephus f. tenuirostris	LGS700	Tanzania	Nguru South Forest Reserve	ZMUC	136233	Blood
Phyllastrephus f. tenuirostris	LGS701	Tanzania	Nguru South Forest Reserve	ZMUC	135235	Blood
Phyllastrephus f. tenuirostris	LGS702	Tanzania	Nguru South Forest Reserve	ZMUC	136238	Blood
Phyllastrephus f. tenuirostris	LGS703	Tanzania	Nguru South Forest Reserve	ZMUC	132747	Blood
Phyllastrephus f. tenuirostris	LGS704	Tanzania	Nguru South Forest Reserve	ZMUC	132766	Blood
Phyllastrephus f. tenuirostris	LGS705	Tanzania	Nguru South Forest Reserve	ZMUC	132795	Blood
Phyllastrephus f. tenuirostris	LGS706	Tanzania	Nguru South Forest Reserve	ZMUC	132965	Blood
Phyllastrephus f. tenuirostris	LGS707	Tanzania	Nguru South Forest Reserve	ZMUC	137458	Blood
Phyllastrephus f. tenuirostris	CVV566	Tanzania	Mazumbai	ZMUC	120128	Blood

Subspecies	Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
Phyllastrephus f. tenuirostris	CVV577	Tanzania	Mafi Forest Reserve	ZMUC	121317	Blood
Phyllastrephus f. tenuirostris	CVV584	Tanzania	Nguru North, Nguu Mts.	ZMUC	122709	Blood
Phyllastrephus f. tenuirostris	CVV586	Tanzania	Gombero Forest, Nguu	ZMUC	123559	Blood
Phyllastrephus f. tenuirostris	CVV587	Tanzania	Gombero Forest, Nguu	ZMUC	123565	Blood
Phyllastrephus f. tenuirostris	CVV593	Tanzania	Chamanyani Mvuha Forest Reserve	ZMUC	129604	Blood
Phyllastrephus f. tenuirostris	CVV606	Tanzania	Kanga Forest	ZMUC	134302	Blood
Phyllastrephus f. tenuirostris	CVV607	Tanzania	Udaha Camp	ZMUC	134356	Blood
Phyllastrephus f. tenuirostris	CVV608	Tanzania	Udaha Camp	ZMUC	134364	Blood
Phyllastrephus f. tenuirostris	CVV609	Tanzania	Udaha Camp	ZMUC	134372	Blood
Phyllastrephus f. tenuirostris	CVV610	Tanzania	Udaha Camp	ZMUC	134385	Blood
Phyllastrephus f. tenuirostris	CVV619	Tanzania	Kilindi Forest, Nguu Mts.	ZMUC	137657	Blood
Phyllastrephus f. tenuirostris	CVV620	Tanzania	Mtiniko Forest	ZMUC	137677	Blood
Phyllastrephus f. tenuirostris	CVV621	Tanzania	Mtiniko Forest	ZMUC	137686	Blood
Phyllastrephus f. alfredi	LGS172	Tanzania	Mbizi Forest	ZMUC	123316	Blood
Phyllastrephus f. alfredi	LGS173	Tanzania	Mbizi Forest	ZMUC	123402	Blood
Phyllastrephus f. alfredi	LGS222	Tanzania	Mbizi Forest	ZMUC	123291	Blood
Phyllastrephus f. alfredi	LGS223	Tanzania	Mbizi Forest	ZMUC	123298	Blood
Phyllastrephus f. alfredi	LGS224	Tanzania	Mbizi Forest	ZMUC	123318	Blood
Phyllastrephus f. alfredi	LGS225	Tanzania	Mbizi Forest	ZMUC	123334	Blood
Phyllastrephus f. alfredi	LGS226	Tanzania	Mbizi Forest	ZMUC	123388	Blood
Phyllastrephus f. alfredi	LGS227	Tanzania	Mbizi Forest	ZMUC	123391	Blood
Phyllastrephus f. alfredi	LGS228	Tanzania	Mbizi Forest	ZMUC	123403	Blood
Phyllastrephus f. alfredi	LGS229	Tanzania	Mbizi Forest	ZMUC	123404	Blood
Phyllastrephus f. alfredi	LGS326	Malawi	Chitipa,Zovo Chipolo Forest, Nyika NP	FMNH	440576	Tissue
Phyllastrephus f. alfredi	LGS327	Malawi	Chitipa,Zovo Chipolo Forest, Nyika NP	FMNH	440577	Tissue
Phyllastrephus f. alfredi	LGS328	Malawi	Chitipa,Zovo Chipolo Forest, Nyika NP	FMNH	440578	Tissue
Phyllastrephus f. alfredi	LGS329	Malawi	Ntchisi Forest Reserve	FMNH	444165	Tissue

Subspecies	Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
Phyllastrephus f. alfredi	LGS330	Malawi	Ntchisi Forest Reserve	FMNH	444166	Tissue
Phyllastrephus f. alfredi	LGS332	Malawi	Ntchisi Forest Reserve	FMNH	444168	Tissue
Phyllastrephus f. alfredi	LGS334	Malawi	Ntchisi Forest Reserve	FMNH	444170	Tissue
Phyllastrephus f. alfredi	LGS335	Malawi	Ntchisi Forest Reserve	FMNH	444171	Tissue
Phyllastrephus f. alfredi	LGS336	Malawi	Ntchisi Forest Reserve	FMNH	444172	Tissue
Phyllastrephus f. alfredi	LGS337	Malawi	Ntchisi Forest Reserve	FMNH	444173	Tissue
Phyllastrephus f. alfredi	LGS338	Malawi	Ntchisi Forest Reserve	FMNH	444174	Tissue
Phyllastrephus f. alfredi	LGS339	Malawi	Ntchisi Forest Reserve	FMNH	444175	Tissue
Phyllastrephus f. alfredi	LGS340	Malawi	Ntchisi Forest Reserve	FMNH	444176	Tissue
Phyllastrephus f. alfredi	LGS341	Malawi	Ntchisi Forest Reserve	FMNH	444177	Tissue
Phyllastrephus f. alfredi	LGS342	Malawi	Ntchisi Forest Reserve	FMNH	444178	Tissue
Phyllastrephus f. alfredi	LGS343	Malawi	Ntchisi Forest Reserve	FMNH	444179	Tissue
Phyllastrephus f. alfredi	LGS344	Malawi	Ntchisi Forest Reserve	FMNH	444180	Tissue
Phyllastrephus f. alfredi	LGS345	Malawi	Ntchisi Forest Reserve	FMNH	444181	Tissue
Phyllastrephus f. alfredi	LGS346	Malawi	Ntchisi Forest Reserve	FMNH	444182	Tissue
Phyllastrephus f. alfredi	LGS347	Malawi	Ntchisi Forest Reserve	FMNH	444183	Tissue
Phyllastrephus f. alfredi	LGS349	Malawi	Ntchisi Forest Reserve	FMNH	444185	Tissue
Phyllastrephus f. alfredi	LGS350	Malawi	Ntchisi Forest Reserve	FMNH	444186	Tissue
Phyllastrephus f. alfredi	LGS351	Malawi	Ntchisi Forest Reserve	FMNH	444187	Tissue
Phyllastrephus f. alfredi	LGS352	Malawi	Ntchisi Forest Reserve	FMNH	444188	Tissue
Phyllastrephus f. alfredi	LGS353	Malawi	Ntchisi Forest Reserve	FMNH	444189	Tissue
Phyllastrephus f. alfredi	LGS354	Malawi	Ntchisi Forest Reserve	FMNH	444190	Tissue
Phyllastrephus f. alfredi	LGS355	Malawi	Ntchisi Forest Reserve	FMNH	444191	Tissue
Phyllastrephus f. alfredi	LGS357	Malawi	Ntchisi Forest Reserve	FMNH	444193	Tissue
Phyllastrephus f. alfredi	LGS358	Malawi	Ntchisi Forest Reserve	FMNH	444198	Tissue
Phyllastrephus f. alfredi	LGS674	Zambia	Lundazi, E P	FMNH	205603	Toe pad
Phyllastrephus f. alfredi	LGS675	Zambia	Lundazi, E P	FMNH	205604	Toe pad

Subspecies	Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
Phyllastrephus f. alfredi	LGS676	Zambia	Lundazi, E P	FMNH	205605	Toe pad
Phyllastrephus f. alfredi	LGS677	Zambia	Lundazi, E P	FMNH	205606	Toe pad
Phyllastrephus f. alfredi	LGS678	Zambia	Lundazi, E P	FMNH	205607	Toe pad
Phyllastrephus f. alfredi	LGS908	Malawi	Ntchisi Forest Reserve	MVZ		Tissue
Phyllastrephus f. alfredi	LGS910	Malawi	Ntchisi Forest Reserve	MVZ	W4631/10/07	Tissue
Phyllastrephus f. alfredi	LGS911	Malawi	Ntchisi Forest Reserve	MVZ	RB1094	Tissue
Phyllastrephus f. alfredi	LGS914	Malawi	Ntchisi Forest Reserve	MVZ	RB1113	Tissue
Phyllastrephus f. alfredi	LGS915	Malawi	Ntchisi Forest Reserve	MVZ	RB1128	Tissue
Phyllastrephus f. alfredi	LGS916	Malawi	Ntchisi Forest Reserve	MVZ	RB1141	Tissue
Phyllastrephus f. alfredi	LGS921	Malawi	Ntchisi Forest Reserve	MVZ	MLW-B4	Tissue
Phyllastrephus f. alfredi	LGS922	Malawi	Ntchisi Forest Reserve	MOM	2007-2-206	Tissue
Phyllastrephus f. kungwensis	LGS319	Tanzania	Kasiha Forest, Mahale Mountains National Park	ZMUC	135427	Tissue
Phyllastrephus f. olivaceogriseus	LGS359	Burundi	Muramuya	FMNH	346351	Tissue
Phyllastrephus f. olivaceogriseus	LGS360	Burundi	Muramuya	FMNH	343653	Tissue
Phyllastrephus f. olivaceogriseus	LGS361	Burundi	Ndora, Cibitoke	FMNH	357986	Tissue
Phyllastrephus f. olivaceogriseus	LGS362	Burundi	Ndora, Cibitoke	FMNH	357987	Tissue
Phyllastrephus f. olivaceogriseus	LGS363	Burundi	Ndora, Cibitoke	FMNH	357989	Tissue
Phyllastrephus f. olivaceogriseus	LGS364	Uganda	Kisoro, Southern P	FMNH	384937	Tissue
Phyllastrephus f. olivaceogriseus	LGS365	Uganda	Kisoro, Southern P	FMNH	384938	Tissue
Phyllastrephus f. olivaceogriseus	LGS682	DRC	Kabobo Forest, SW Talama, South Kivu	FMNH	450460	Tissue
Phyllastrephus f. olivaceogriseus	LGS683	DRC	Kabobo Forest, SW Talama, South Kivu	FMNH	450461	Tissue
Phyllastrephus f. olivaceogriseus	LGS684	DRC	Kabobo Forest, SW Talama, South Kivu	FMNH	450462	Tissue
Phyllastrephus f. olivaceogriseus	LGS686	DRC	Kabobo Forest, SW Talama, South Kivu	FMNH	450464	Tissue
Phyllastrephus f. olivaceogriseus	LGS687	DRC	Kabobo Forest, SW Talama, South Kivu	FMNH	450465	Tissue
Phyllastrephus f. olivaceogriseus	LGS688	DRC	Kabobo Forest, SW Talama, South Kivu	FMNH	450466	Tissue
Phyllastrephus f. olivaceogriseus	LGS689	DRC	Kabobo Forest, SW Talama, South Kivu	FMNH	450467	Tissue

Subspecies	Extraction no	Country of origin	Collection locality	Source	Museum catalog number	Type of sample
Phyllastrephus f. olivaceogriseus	LGS690	DRC	Kabobo Forest, SW Talama, South Kivu	FMNH	459468	Tissue
Phyllastrephus f. olivaceogriseus	LGS735*	DRC	Kivu, Luiko	RMCA	91916	Toe pad
Phyllastrephus f. olivaceogriseus	LGS736*	DRC	Kivu, Luiko	RMCA	94657	Toe pad
Phyllastrephus f. olivaceogriseus	LGS764	Burundi	Kibira NP, Bubanza province	FMNH	427634	Toe pad
Phyllastrephus f. olivaceogriseus	LGS765	Burundi	Bururi FR, Bururi province	FMNH	364002	Toe pad
Phyllastrephus f. olivaceogriseus	LGS766	Burundi	Kibira NP, Bubanza province	FMNH	346334	Toe pad
Phyllastrephus f. olivaceogriseus	LGS767	Burundi	Kibira NP, Bubanza province	FMNH		Toe pad
Phyllastrephus f. olivaceogriseus	LGS768	Burundi	Kibira NP, Bubanza province	FMNH	346347	Toe pad
Phyllastrephus f. olivaceogriseus	LGS769	Burundi	Kibira NP, Bubanza province	FMNH	363999	Toe pad
Phyllastrephus f. olivaceogriseus	LGS770	Burundi	Kibira NP, Bubanza province	FMNH	357990	Toe pad
Phyllastrephus f. graueri	LGS366	DRC	Chashoga Swamp, Kahuzi-Biega National Park, South Kivu	FMNH	443817	Blood
Phyllastrephus f. graueri	LGS737	DRC	Ituri, Wago	RMCA	61306	Toe pad
Phyllastrephus f. graueri	LGS738	DRC	Kivu,Tohoba	RMCA	674-44-A876	Toe pad

** P. f. itombwensis*