# Ecological and evolutionary processes in two southern African endemic birds

**Ângela M. Ribeiro**



**Dissertation presented for the Degree of**
**Doctor of Philosophy**
**in the Department of Zoology**
**University of Cape Town**

**Cape Town, August 2011**

# DECLARATION

I, Ângela Maria Oliveira Ribeiro, hereby declare that: i) the work presented in this thesis is my own and, ii) neither the substance nor any part of this thesis has been submitted or is to be submitted for a degree at this or other University.

Cape Town, August 2011.

I chased lions, I climbed mountains, I dove deep.
I rambled the jungle, I lived in a tank, I flu with strings,
I had private clouds and self-sunrises.
I was a silent cold blue wind and a blooming spring flower.
In my world, in my scene, within me,
I recreated myself.
Now, in the mighty silence of my dreams,
in my scene, within me,
I see your smile.
I can swim, I can dance; we can fly
with no fears, with no strings, until the end.
And I still love.

*To the memory of Manuel Paço e Brito*

# CONTENTS

# PREFACE

## Sleepwalk typewriter

Oftentimes I find my memories sleepwalking to the imagination, where all the ephemerally disoriented memories unhappen in the same scene: a naked Baobab at the peaceful rhythm of the Indian Ocean waves; the mossy scent of the forest after a heavy rain with an eagerly roaring waterfall; the dazzling warm rays of a sunrise closely watched by beautiful Orion; the hollow sound of electric fences in the same tone of the cheerful carillon; a colourfully dressed woman carrying a tin of water, empress of the cool summer fog; the tart-sweet ripped guavas nurtured by the salty ocean smell; the shadow of blooming oaks over the wavy fiery red dunes; the beloved persons I lost for never in the same smile with the friends I gained for ever. All unhappening until the blue, red and green decide to sum up in the sky, and then shatter into the earth painting the most gorgeous and peaceful sunset one can witness. Memories; they all temporarily return to 'happened'. I find myself when I imagine this beautiful journey. Imagine.

# ACKNOWLEDGEMENTS

Completing my degree would not have been possible without the generosity of many people. It is a great pleasure to thank everyone, who have inspired, assisted and guided me in multiple ways.

Foremost, I want to thank Rauri Bowie, my advisor. Rauri, thank you so much for giving me the opportunity to be a member of your lab, at UC Berkeley. I appreciated the freedom you gave me to find my way. I am grateful for your affability, patience, support and guidance throughout this journey. My most sincere and endless thank you!

Penn Lloyd, my academic supervisor during the first two years of the programme, your help in the field was precious. I will never forget your hospitality when I first arrived in South Africa – thank you!

Phil Hockey, I am grateful for your willingness to act as my academic supervisor and your thoughtful comments when revising my chapters – thanks so much!

Ricardo Lopes, thanks for your support during the process of writing and submitting the FCT grant application and for joining me in the field.

I acknowledge the Fundação para a Ciência e Tecnologia (FCT, Portugal) and the Department of Science and Technology and National Research Foundation (NRF, South Africa) for providing funding to conduct my research. I am grateful to the Animal Ethics Committee at the University of Cape Town for approving my research program. And for granting the research and collecting permits I thank: SAFRING, the South African National Parks; Northern Cape Department of Nature and Environmental Conservation; Free State Department of Tourism, Environment and Economic Affairs; Eastern Cape Department of Agriculture, Environment and Tourism; and, the Department of Environment and Tourism Department of Namibia.

Hugh Chittenden, Mark Brown, Dawie de Swardt, Joseline and Trevor Smit, Adrian Craig, Niel MacGregor, Rod and Cathy Molteno, Larry King, Ronnelle and Stoffel Visagie, Sharon Louw, David Pretorius and, Pieter Strauss: thank you for being fantastic hosts, helping with logistics in the field and for sharing your knowledge about the region, wildlife and particularly Scrub-Robins.

I thank David Willard, Gary Voelker and Tom Gnoske (Field Museum, Chicago, USA), Dawie de Swardt (National Museum at Bloemfontein, Free State, South Africa), Joseph Heymans, Kobie Raijmakers, Craig Symes, and Peter Nupen, for providing biological samples.

Andrew, the mesmerizing; Jay, the adorable; Roberta, the heartfelt; José, the amusing; Ricardo, the epicurean; Aninha, the generous; Jérôme, the 'mais oui' jocular: to all of you my utmost, deep and sincere gratitude. Thanks for your endless help and support, for the company, for your thoughtful insights, for discussing and reading so carefully my manuscripts, for the heavy doses of optimism and mostly, for the memorable and marvellous moments together. I very much appreciated your friendship, kindness and generosity throughout this intense journey. I love you all very much dear friends!

Jérôme, Hanneline and Graeme, my field-office-labmates, I had a terrific time with you in the field! It was "baie, baie lekker!" Thanks very much for running for Scrub-Robins. Sandra, Marta, Buckley, Knud, Jon, Fujita, Mitzy, Kristal, Tânia, Michael, Luca, Potiphar, Sampath, Callan, Michael, Rowena: thanks for the wonderful times together!

I was officially a Ph.D student at the University of Cape Town in South Africa. However, I spent most of my time at the Museum of Vertebrate Zoology, University of California at Berkeley. Berkeley had/has an intriguing simultaneous duality, which makes it unique and unforgettable: innovative and traditional, friendly and aloof, inspiring and disheartening, fulfilling and overwhelming, challenging and monotonous. Being a student in Berkeley was such a rewarding experience: scientifically and personally! I acknowledge Craig Moritz, the director of the Museum, for allowing me to be part of the MVZ community. I extend my gratitude to everyone in the MVZ for welcoming and hosting me during these last four years – thanks very much! All the Bowie's lab members are thanked for tirelessly listen to, discussing and reading my ideias.

Cristina e Sara, obrigada por fazerem da distância geografica que nos separou um mero até logo. Encheu-me o coração cada uma das mensagens/presentes que me enviaram. Obrigada pela partilha!

Avó Maria e Ermelinda, embora a barreira tecnlogica tenha exacerbado a distância física, cada conversa foi fonte de alegria e felicidade. A vossa valentia foi e será sempre inspiradora e motivo de orgulho. Obrigada!

Mãe e pai, obrigada pela constante presença, pelo apoio infinito. Sem a vossa generosidade e amor nada seria possível. Adoro-vos!

Obrigada, Thank you, Dankie, Gracias, Merci!

August, 2011

# ABSTRACT

Understanding the continuous process of divergence that leads to the formation of species requires the simultaneous think about ecology, evolution and its intertwining relationship. In this dissertation, I propose a population-level framework that integrates ecology and population genetics to rethink the mechanisms that promote divergence. To accomplish that I studied two southern African endemic *Cercotrichas* scrub-robins: *C. coryphaeus* and *C. signata.* Despite its phylogenetic proximity the two study species occupy the two ends of the habitat spectrum available in southern Africa: semi-arid Karoo vs. Forest. First, I combined behavioural ecology with individual-based genetic data in the Karoo scrub-robin examined the role of social relationships in shaping mating patterns and dispersal behaviour and ultimately influence the evolution of dispersal. Second, using the knowledge obtained about the dispersal biology of the Karoo scrub-robin, I tested whether natural selection can counter-balance gene flow and hence promote local adaptation along the aridity gradient that characterizes the range of the species. The results suggested that selective pressures on physiology, mediated by the mtDNA genome, might facilitate the adaptation to new climatic conditions. Third, prompted by the previous findings, I examined the historical context of a spatial and demographic expansion to understand the role of natural selection in the colonization of a new environment, in the Karoo scrub-robin. Lastly, I evaluated how the eco-geographical context affects the divergence process in the forest-dweller *C. signata.*

# CHAPTER SUMMARY

The intellectual motivation for this dissertation was shed light on the mechanisms that promote population divergence and can ultimately lead to the formation of species, in an area of the world that is still understudied: the southern African subcontinent. My thesis consists of seven chapters. Chapter one is an introduction to the topics I proposed to address, and chapter seven the general conclusions from the research I developed and described in five interrelated chapters. Each of these chapters focuses on particular aspects, but collectively enabled me to present an integrated picture of the ecological and evolutionary mechanisms that intersect during the speciation process in southern African avifauna.

Chapter two is a technical note where I report the development of species-specific molecular genetic markers, microsatellites, required to address the questions in chapters three to five.

In chapter three, I combined behavioural ecology data with individual-based genetic data for the Karoo Scrub-Robin *Cercotrichas coryphaeus*, a facultative cooperatively breeding bird, to examine the role of social relationships in shaping mating patterns and dispersal behaviour and that ultimately influence the evolution of dispersal. The results revealed that males and females do not have symmetrical roles in structuring the population. The long-distance dispersal strategy of females ensured that Karoo Scrub-Robins do not pair with relatives thereby compensating for male philopatry caused by cooperation.

Chapter four builds on the knowledge I obtained about the dispersal biology of the Karoo Scrub-Robin and the increasing appreciation of the role of the mitogenome in physiology, to tested whether natural selection can counter gene flow and hence promote local adaptation along the aridity and seasonality gradient that characterizes the range of the species. The results suggest that the mitochondrial genome, given its physiological role, might mediate the adaptive response to the selective pressures exerted on physiology.

In chapter five, I examined the historical context of a spatial and demographic expansion in the Karoo Scrub-Robin in order to understand the role of natural selection in the colonization of a new environment. I used population genetic theory to test distinct demographic scenarios and compared these results to the observed pattern of variation at putative adaptive and non-adaptive variation.

Finally, in chapter six, I evaluated how the eco-geographical context affects the divergence process in the forest-associated *C. signata*. Including this study species, allowed me to gain some insights into the mechanisms that drive divergence in two phylogenetically related *Cercotrichas* species.

CHAPTER ONE

# Introduction

## THE PROCESS OF SPECIES FORMATION

Since *On the Origin of Species by Means of Natural Selection* (Darwin, 1859) the study of speciation has been pursued with great and enthusiastic debate. What is a species? Are they biological entities? How do species form? Is the speciation process punctuated or continuous? Addressing these fundamental questions has generated many philosophical essays and important theoretical models (e.g. Dobzhansky 1951; Mayr 1942, 1947, 1963; Endler 1977; Mallet 1995; Turelli et al. 2001; Kirkpatrick and Ravigné 2002; Wiens 2004, Rundle and Nosil 2005, Dieckmann et al. 2004, Mallet 2008).

During the *Modern Synthesis* movement, ecology and geography were central ideas in the development of models to explain the process of origin of species (e.g. Dobzhansky 1951, Mayr 1942, 1947). In the *geographical theory* of speciation, Mayr (1942) posited that a new species could only develop if a population that became geographically isolated acquired characters that maintained their isolation when the external geographical barrier disappeared. Mayr (1947) recognized also the importance of ecological factors, i.e. circumstances that affect population demographic dynamics, in the formation of species. However, he emphasised the importance of geographical isolation as the starting point of the speciation process.

Mayr's perspective about speciation became pervasive and overshadowed the models that proposed that local differences in the environment are the starting point for the formation of a new species (Thorpe 1945, Ehrlich and Raven 1969, Schluter 2000, 2001; Via 2001). Ehrlich and Raven (1969) challenged the Mayrian viewpoint. They argued that rather than focusing only on the factors responsible for interrupting gene flow, the disruptive and cohesive roles of selection and its influence on gene flow should be given more attention when studying speciation. Ehrlich and Raven's verbal argument was later formalized by Turelli et al. (2001). These authors demonstrated that despite gene flow being at its maximum strength, divergence could proceed within a single panmitic population. This signalled the development of a new paradigm in speciation studies: natural selection can overcome the homogenizing effects of gene flow and recombination, and lead to the development of distinct genetic clusters, and ultimately new species, even in the absence of any physical barrier to dispersal.

Studying speciation in a retrospective approach, i.e. considering a latest snapshot of the process to infer the mechanisms, and reducing speciation to geographical modes (allo-, para- or sympatric) or to pre- and post-mating isolation categories has been recognized as potentially overlooking important aspects along the continuum that is the process of species formation (Endler 1977, Schluter 2001, Coyne and Orr 2004, Price 2008, Fitzpatrick et al. 2009, Via 2009, Sobel et al. 2010). During the last decade, speciation studies have begun to make a transition from the before mentioned procedure to an approach where the

focus are populations that are not yet reproductively isolated. Besides providing a framework to understand the ecological and micro-evolutionary processes that interact to determine divergence, the population level approach also reduces the confounding effects of differences accumulated after speciation (Via 2009). The implementation of this population-level perspective, in several systems, have provided interesting insights about the evolution and the formation of species (e.g. *Mimulus* monkey-flowers: Bradshaw and Schemske 2003, Angert et al. 2008; *Anopheles* African mosquitoes: della Torre 2002, Turner et al. 2007, Costantini et al. 2009; *Rhagoletis* maggot flies: Feder et al. 1998, Feder et al. 2003; *Timena* walking-sticks: Nosil et al. 2004, 2008, 2009a; *Littorina* rough periwinkles: Rolán-Alvarez 1997, Grahame et al. 2006, Butlin et al. 2008; *Gasterosteus* sticklebacks Schluter 2000, Vines and Schluter 2006; African Cichlidae: Allender et al. 2003, Seehausen 2006; *Anolis* lizards: Thorpe et al. 2004, Thorpe et al. 2008; *Geospiza* Darwin's finches: reviewed in Grant and Grant 2008).

## THE INTERSECTION OF ECOLOGY AND EVOLUTION: DISPERSAL AND NATURAL SELECTION

Darwin (1859) was the first to appreciate that ecology and evolution are intertwined. In his formulation, population densities, that would otherwise growth exponentially, were controlled by factors that affected individual survival and reproduction. Only those individuals that could outcompete others would be favoured by natural selection, and would contribute to the following generation. Nearly a century later, during the development of the *synthetic theory* of evolution, Mayr (1963) defined evolution as a two-step process that involves both the origin and the change in frequency of a new variant in a population. Underlying Mayr's definition of evolution there is a self-evident ecological concept: demographic dynamics. The increase in frequency of new variants results from demographic changes in the population. Despite this long-recognised nexus between evolution and ecology (discipline that studies the distribution and demographic changes of populations) the two fields of biology have drifted for a long time.

Recently, multiple authors have been calling for the simultaneously consideration about ecological and evolutionary processes (Saccheri and Hanski 2006, Metcalf and Pavard 2007, Kokko and Lopéz-Sepulcre 2007, Schoener 2011). For instance, the fitness of members of a population that have a particular phenotype (e.g., behaviour, morphology), which is determined by a genotype, will affect the population demography. In turn, the demographic dynamics of the population will increase/decrease certain phenotypes and thus affect its evolution.

**Dispersal and Gene Flow**

Nearly a century ago, based on his field observations, Joseph Grinnell (1914, 1922) noted the effect of centrifugal dispersal from the *metropolis* (the centre of abundance) towards the *periphery* (the edge of the range of the species, typically an area of low densities). The population dynamics in the *metropolis* is dominated by density-dependent (biotic) factors, such as competition, whereas in the *periphery* density-independent (abiotic) factors, such as environmental factors, are the most important. Grinnell's original formulation led to multiple subsequent essays about dispersal and its key importance in the fields of ecology and evolution (e.g. Stenseth and Lidicker 1992, Clobert *et al.* 2001, Bowler and Benton 2005, Ronce 2007).

Dispersal is often defined as the change of spatial location over the course of an organism life (Stenseth and Lidicker 1992, Clobert *et al.* 2001). However, dispersal is best understood if it is defined as an individual-based behavioural response to local environmental and social constrains, shaped by natural selection to optimize individual fitness. The decision of leaving the natal site is commonly thought to be costly due to: i) the risks of moving through unknown environments, ii) the associated energetic expenditure of finding a suitable habitat and iii) the need to outcompete residents (for space, food or mates) in the new habitat (Bowler and Benton 2004). Nevertheless, dispersal is a common phenomenon in nature. Unravelling the proximate triggers of such behaviour requires investigating the patterns of sex asymmetry in dispersers, the timing that individuals decide to abandon the natal site and the distance moved. For instance, habitat quality, social context and/or population density may all be important cues to trigger dispersal (Ronce 2007). Different hypotheses have been proposed to explain the benefits of leaving the natal site, despite the associated costs: dispersal might have evolved as a life-history trait selected to avoid competition for mates (Hamilton and May 1977, Greenwood 1980), resources (Dobson 1982), and/or to reduce the risk of breeding with relatives (Moore and Ali 1984, Perrin and Mazalov 1999, Perrin and Goudet 2001).

Dispersal, although an individually based behaviour, is inherently an ecological concept because it influences the dynamics of populations (Hanski and Gilpin 1997). In addition, dispersal determines the underlying population genetics through the movement of genes, i.e. gene flow, as outlined by Sewall Wright (1931) and Gustave Malécot (1967). Both the demographic dynamics and the genetics of populations can reciprocally feed back into the evolution of the dispersal behaviour itself (Kokko and López-Sepulcre 2006, Ronce 2007).

The fact that species' ranges often span large geographical areas, relative to individual dispersal distances, prompted biologists to equate the importance of dispersal and gene flow in theories that aim to explain changes in the geographical distribution of species (Case and Taper 2000, Holt 2003, Goldberg

and Lande 2007), colonization of new habitats (Baker and Stebbins 1965, Keller and Taylor 2008), adaptation to local conditions (Antonovics 1976, Barton and Kirkpartrick 1997, Kawecki 2004), and speciation (Mayr 1963, Turelli et al. 2001, Coyne and Orr 2004, Rundel and Nosil 2005).

**Natural Selection**

Populations occupy spatially and temporally heterogeneous environments. This heterogeneity can be thought in the context of the presence of other individuals (homospecific or heterospecific) or landscape features. Local ecological conditions likely constrain individual fitness, i.e. the effective (genetic) contribution of individuals to the following generation, and therefore the frequency of phenotypes. The process whereby a phenotypic variant changes in frequency, given that it is polymorphic, inheritable and associated with mating success or fecundity is known as natural selection (Darwin 1859, Endler 1986). It is relevant to mention that the change in frequency of a given trait, regardless of its origin, does not imply natural selection; it can result from alternative evolutionary processes (drift).

Natural selection can drive the movement of a population towards a phenotypic state that best fits the present environment: the evolutionary process of adaptation (Fisher 1930). However, it is important to remember that natural selection does not equate to adaptation, but it rather explains the process through which a relative better adaptation can increase in frequency once it appears in the population (Endler 1986). Adapting to local conditions depends ultimately on the interaction between natural selection (increasing the frequency of the advantageous variant) and gene flow (spreading the adaptive trait). Therefore, studying local adaptation requires a good understanding about the biology of dispersal and quantitative estimates of gene flow.

Cases of adaptation to local conditions are common in nature (see examples in Endler 1986, Schluter 2000, Kawecki and Ebert 2004, Gavrilets and Losos 2009, Bicudo et al. 2010). One of the classical examples where natural selection results in adaptation to local ecological conditions are the *Geospiza* Darwin's finches (see Grant and Grant 2008 for a synthesis), where the availability of food resources pushes individuals to distinct phenotypic adaptive peaks (bill morphology). *Geospiza* finches are one of the few cases of adaptive evolution where the key trait and the underlying genes are known (Abzhanov et al. 2004, 2006). The genetic basis of adaptation has been uncovered slowly, as exemplified by studies that have revealed the role of hemoglobin in adaptation to hypoxia (e.g. Storz et al. 2007, Storz et al. 2009, McCracken et al. 2009) or pigmentation genes such as melanocortin-1 receptor in background matching (e.g. Hoekstra et al. 2006, Linnen et al. 2009, Rosenblum et al. 2010). Mitochondrial genes and the mito-nuclear interaction in the energy production system have also been shown to play an important role in local adaptation and potentially speciation (reviewed by Das 2006; Gershoni et al. 2009, Ballard

and Melvin 2010).

Adaptive evolution can play an important role in the colonization of new areas (e.g. Rosenblum et al. 2006, Kim and Gulisija 2010). During the ecological process of colonization of a new habitat dispersers are exposed to an environment that may differ dramatically from their natal habitat. The successful establishment of populations in a new habitat depends on the ability of dispersing individuals to reach patches of suitable habitat, and whether they can reproduce in a habitat where food resources, climatic conditions and predators and competitors differ. If migrants with a particular trait have an advantage in the new environment, the demographic dynamics of the population will change and the descendants of individuals that inherited the adaptive trait will eventually dominate the population. Hence, the new population will occupy a new phenotypic adaptive peak.

## SOUTHERN AFRICA: TWO CONTRASTING COASTS

The southern African subcontinent includes Namibia, South Africa, Botswana, Zimbabwe, Lesotho, Swaziland and Mozambique south of the Zambezi River. The region has exceptional species diversity (faunal and floral), with three biodiversity hotspots currently defined. However, that diversity is not evenly spread across the whole subcontinent; it tends to be concentrated on the coast (Figure 1). Two oceanic currents primarily influence the southern Africa climate: the tropical Agulhas current that flows southwestward along the eastern coast and in the Benguela current, creates the northward flux of cool south Atlantic waters along the western coast. These oceanic currents together with the Drakensberg Mountains (Figure 1) drive the west to east aridity gradient: from the sparsely vegetated arid west to the luxuriantly forested temperate and tropical east. While the Benguela current creates a very low precipitation regime and drought conditions, the Drakensberg mountains act as a barrier to the moist air masses moving from the Indian Ocean to the interior of the subcontinent.

The western side of southern Africa is characterised by two pronounced geographical rainfall gradients (Desmet and Cowling 1998). The first is a latitudinal aridity gradient with the overall amount of precipitation decreasing northwards into the southern Namib Desert. The second one is a longitudinal seasonality gradient (perpendicular to the previous) from a winter rainfall regime along the Atlantic coast to summer rainfall regime in the interior in a very small geographical scale (200 km). In the east, over less than 100 km, the relief changes dramatically from the sea level to the west highlands (Drakensberg Mountains; ~ 1200 m). As a consequence of this topography, the isotherms run parallel to the coast, such that minimum temperatures decline from the coast going inland and summer rainfall predominance

increases (Mucina and Rutherford 2006). Latitudinally, there is a clear south to north trend of seasonality, from all-year to predominately summer rainfall, and an increase in minimum temperature. This peculiar climatic setting has a dramatic effect on vegetation physiognomy, primary productivity and water availability across the landscape of southern Africa (Cowling et al. 1997, Dean and Milton 2004). While the xeric Karoo and Fynbos dominate the western part of the subcontinent, the eastern coast is dominated by temperate and subtropical vegetation (Mucina and Rutherford 2006).

The Karoo includes the semi-arid western coastal plain and the arid central plateau, and can be divided into two sectors according to the seasonality of rainfall: Succulent and Nama Karoo. The Succulent Karoo comprises the western coastal plains below 800 m above sea level from the southern part of the Namib Desert extending to meet the Fynbos in the Western Cape, and is bounded to the east by the Great Escarpment. Dwarf and succulent shrubs mainly dominate the flora of this semi-desert (Milton et al. 1997). Summers are mild and mean minimum temperatures in July (winter) are seldom below zero, preventing severe frosts. Most of the area receives rainfall during the winter and only a small portion may receive rain at any time of the year (Milton et al. 1997). The amount of precipitation received from rain is low (60 mm – 400 mm) but reliable, and is supplemented by precipitation from fog.

In the arid central plateau, east of the Great Escarpment, lies the Nama Karoo. This biome occurs at altitudes that range from 800 m to 1500 m above sea level (Palmer 1997) and is dominated by dwarf shrubs. Temperatures range from above 30°C in the summer to below freezing in the winter (July) causing severe frost. The amount of mean annual rainfall is comparable to the Succulent Karoo, but less reliable and concentrated in the summer (Dean and Milton 2004).

The southwest of the subcontinent harbours the Mediterranean-type Fynbos biome (Afrikaans word for fine leaved bush). This biome is an evergreen, fire-prone shrubland confined to the nutrient-poor soils of the southwestern part of the continent (Mucina and Rutherford 2006). The Fynbos occurs from the sea level to the highlands of the Cape Fold Belt, a mountain range where vegetation transitions to the drier Karoo flora (Figure 1). The temperature ranges are very similar to the Succulent Karoo due to the buffering effect of the Atlantic Ocean. Like in the adjoining Succulent Karoo, it rains during the winter, but in larger quantities (> 400 mm; Cowling et al. 1997).

Under the influence of the warm Agulhas current, the eastern coast supports a subtropical mosaic of savanna, grassland, thicket and forest biomes. Most of the coastal region is currently considered a biodiversity hotspot: Albany-Pondoland-Maputaland (Figure 1). It stretches along the coast from Port Elisabeth in the south to the Limpopo river mouth (Mozambique) and is bounded by the Drakensberg and Lebombo Mountains along the western side (Cowling and Hilton-Taylor 1997). This region supports forest in areas with high water abundance, either from precipitation, fog or water-holding capacities of

soils, such as high highlands, river margins, gorges, seaward-facing scarps and coastal dunes. The forest biome also extends inland into the highlands of Lesotho, Swaziland, Mpumalanga and Limpopo, in an archipelago-like distribution (Mucina and Rutherford 2006). Two main forest complexes can be defined according to the associated climate: Afromontane type in areas of warm and temperate climates and Indian Ocean type in the subtropical area (Midgley et al. 1997).

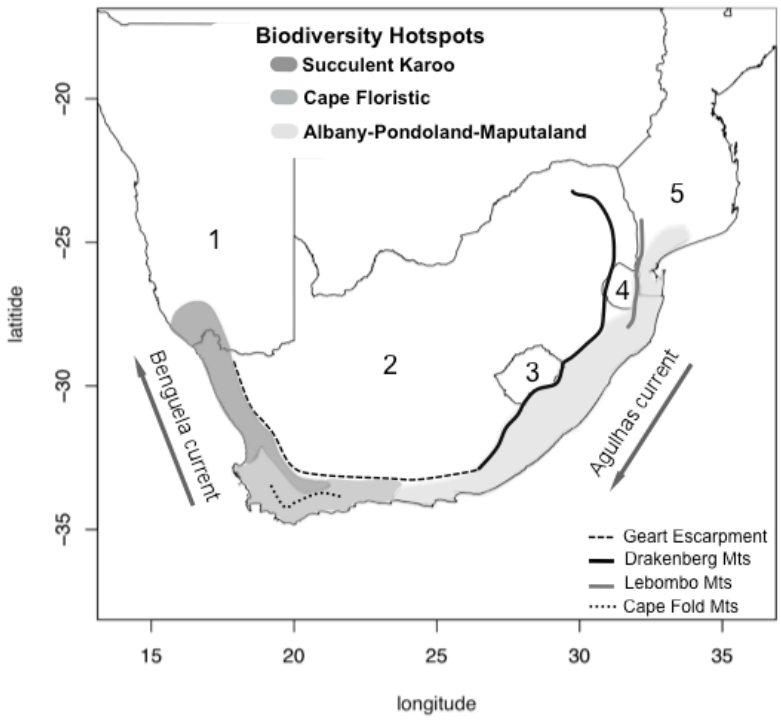


**Figure 1.** Location of the three Biodiversity hotspots recognized in southern Africa, and the main geographical features: Great Escarpment, Drakensberg Mountains, Lebombo Mountains and Cape Fold Mountains. 1. Namibia; 2. South Africa; 3. Lesotho; 4. Swaziland; 5. Mozambique.

Throughout the Miocene (25 to 5 million years – Myr - before present), west and east Africa were considerably warmer and wetter than at present (Cerling et al. 1997). At the late Miocene-Pliocene boundary conditions started to change with a general cooling of the global climate, that resulted in increased aridity causing floral extinctions as suggested by shrub-grassland-forest mosaic records (Tyson 1986). The origin and establishment of the early types of shrublands and succulent vegetation along the western coast and in the interior Karoo dates to the end of the Miocene (Scott et al. 1997). During the Pliocene a further opening of vegetation likely occurred and the Fynbos biome of southwestern Cape expanded along the Escarpment (Tyson 1986). Deep-sea sedimentary sequences reveal a major shift in

climatic variability *ca.* 2.8 Myr ago (first of three arid peaks; deMenocal 2004). The climate of the Pleistocene epoch was mainly characterized by cold arid glacial periods (two aridity peaks about 1.7 and 1.0 Myr before present) oscillating with warmer and wetter interglacial cycles (deMenocal 2004). Aridity was exacerbated during glacial periods when most of the subcontinent was drier and cooler (5 - 6°C lower than present) than today (Chase and Meadows 2007).

In contrast to the arid biomes in the western coast, unique due to its Gondwanaland affinities (Cowling et al. 1997), the forest biome has paleo-biogeographic affinities with forest outside of southern Africa. The Afrotemperate forest is part of the Afromontane region that encompasses the forests in the mountains of East and Central Africa, as well as isolated massifs in Angola and Cameron. The Indian Ocean forest is a southward extension of the coastal forest of Mozambique and Tanzania. The current distribution of forest along the eastern coast of southern Africa does not represent its the past distribution. Under the drier and colder conditions of the Last Glacial Maximum (24 000 - 18 000 years before present) there was a dramatic decrease in Afrotemperate forest in South Africa. These forest isolates are though to have merged recently, about 6 000 years ago (Eeley et al. 1999).

The peculiarities of this part of the globe, such as the historical geology and climatic events that have lead to the formation of the current landscape, makes it a very interesting system, although understudied, with which to address questions pertaining to the interaction between ecological and evolutionary processes in the formation of species, particularly with respect to the highly endemic avifauna.



## *CERCOTRICHAS* SCRUB-ROBINS

The Scrub-Robins are medium-sized, insectivorous birds of the genus *Cercotrichas*, Boie 1831. They are mainly African species of open woodland or shrub habitats (Hockey *et al*. 2005). The current taxonomy and subspecies definitions in *Cercotrichas* are based on plumage colouration, size variation and vocalizations. Ten species are currently defined, two of which are endemic to southern Africa and inhabit the two extremes in the habitat spectrum occupied by Scrub-Robins: the Karoo Scrub-Robin (*Cercotrichas coryphaeus,* Veillot 1817) in the arid/semi-arid shrublands and the Brown Scrub-Robin (*Cercotrichas signata,* Sundevall 1850) in the temperate and tropical evergreen forest. A recent molecular phylogeny revealed that these two species are closely related forming a lineage that is sister to a clade of five *Cercotrichas* species (G. Voelker and R. Bowie unpublished data). The correct spelling of the Latin name

of the Karoo Scrub-Robin, *Cercotrichas coryphaeus* or *Cercotrichas coryphoeus,* is an issue in dispute. For my thesis I decided for the most commonly used *Cercotrichas coryphaeus.*

The Karoo Scrub-Robin *Cercotrichas coryphaeus* is a medium sized (18 - 23g), ground-feeding insectivore (Oatley and Arnott 1988). Pairs are sedentary and defend year-round territories. Clutch-size varies from two to four eggs, with incubation lasting 13 - 15 days. Only females incubate. In some pairs, male offspring are retained on the natal territory until they recruit to a breeding vacancy (Lloyd et al. 2009). The Karoo Scrub-Robin spans the arid and semi-arid zones of the southern African subcontinent, from Namibia southward to the Cape Province and to east to the Free State (Hockey *et al.* 2005). Its range encompasses an area with steep rainfall and temperature gradients: a latitudinal gradient where precipitation decreases northward, from the Cape Province into the southern Namib Desert; and a longitudinal trend in seasonality extending from a winter rainfall regime along the western coast to a summer rainfall regime in the interior, with a narrow intermediate area with year-round rainfall (Desmet and Cowling 1998, Chase and Meadows 2007). Plumage colour differences are used as taxonomic diagnostic traits to define two subspecies: *C .c. cinerea* and *C. c. coryphaeus* (Collar 2005, Hockey et al. 2005). *Cercotrichas c. cinerea,* with greyish upper-parts, is found in the winter rainfall regime area (west and southern coast of South Africa) where the habitat is characterized by semi-arid shrubs and dense thickets. *Cercotrichas c. coryphaeus,* a brownish colour morph, occurs in the central arid area of the subcontinent (central South Africa and Namibia) where sparse and low shrubs dominate the vegetation, and rain falls mostly during the summer (Figure 2a).

The forest dweller Brown Scrub-Robin, *Cercotrichas signata*, spans a narrow and scattered ribbon of evergreen forest along the eastern coast of southern Africa, from the eastern Great Escarpment of Limpopo Province, northeastern Swaziland and northern KwaZulu-Natal to Eastern Cape (Hockey *et al.* 2005). Two subspecies/ecotypes are currently recognized: the dark olive-brown *C. s. signata* and the paler, smaller and shorter-billed *C. s. tonguensis* (Figure 2b) The latter also has narrower white markings on the face. While *C. c. signata* spans the temperate and subtropical forest of Pondoland and highlands of Swaziland, Mpumalanga and Limpopo (reaching 2000 m above sea level), *C. c. tonguensis* is restricted to the tropical lowland coastal forests of northern KwaZulu-Natal and southern Mozambique (Maputaland region). This medium sized (32 - 42g) elusive bird is resident to the understorey of the forest. It sings in the mid-strata of the forest for a short period of time at dawn and sunset. The Brown Scrub-Robin is a territorial bird, monogamous and a solitary nester that breeds from October to December (Collar 2005). Clutch-size varies from two to three eggs, with incubation lasting 14 days (Oatley and Arnott 1988). It feeds mainly on invertebrates and occasionally on seeds and fruit (Collar 2005).

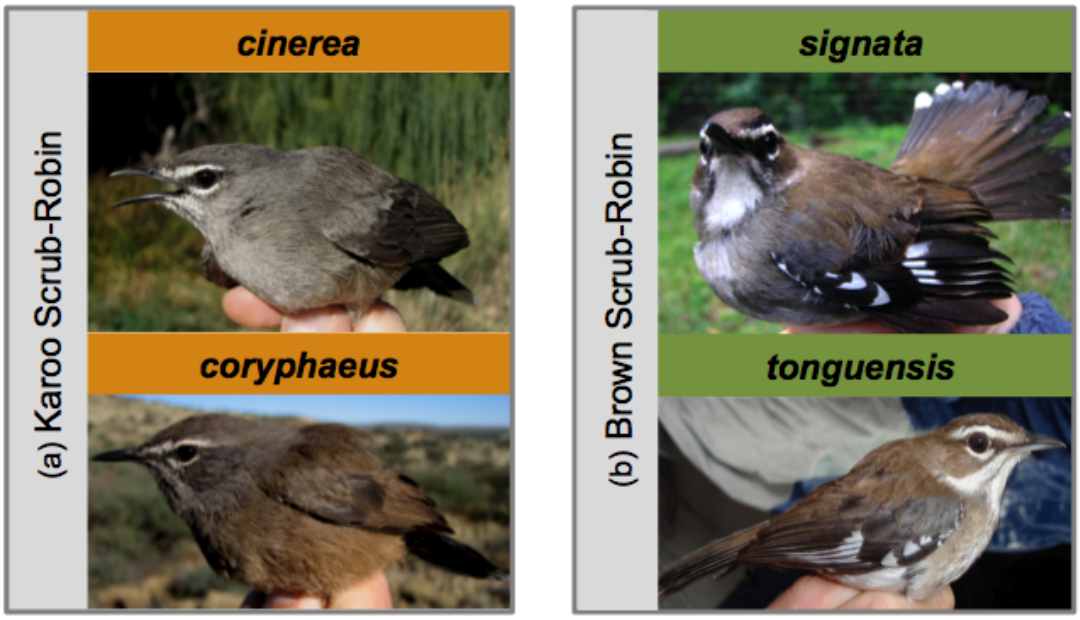


**Figure 2.** The two study species, Scrub-Robins, and susbpecies currently recognized. (a) Karoo Scrub-Robin *Cercotrichas coryphaeus*; (b) Brown Scrub-Robin *Cercotrichas signata*. All photographs by Ângela M. Ribeiro except *C. s. tonguensis* by Penn Lloyd.

CHAPTER TWO

# Microsatellites in the Karoo Scrub-Robin, *Cercotrichas coryphaeus*: isolation and characterization of 13 autosomal and two sex-linked loci



## INTRODUCTION

The Karoo Scrub-Robin *Cercotrichas coryphaeus* is endemic to semi-arid regions of southern African and is a facultative cooperative breeder, with strong territory and mate fidelity (Oatley 2005). Field observations suggest sex-biased natal dispersal, with males exhibiting greater natal philopatry (Lloyd et al. 2009). Thus, the identification of microsatellite markers would allow researchers to test predictions about paternity, relatedness and hence fine-scale genetic structure for this species.

## METHODS

Microsatellite markers were developed using an enrichment protocol (Glenn and Schable 2005). Approximately 4µg of genomic DNA (gDNA) from one individual was digested with BstUI and XmnI. SuperSNX24 linkers were ligated onto the ends of gDNA fragments and served as priming sites for subsequent PCRs. Five biotinylated tetranucleotide probes [$(ACAT)_8$, $(AAAT)_8$, $(AAGT)_8$, $(AACT)_8$, $(AGAT)_8$] were hybridized to cut gDNA. The biotinylated probe–gDNA complex was added to streptavidin-coated magnetic beads (Dynabeads M-280, Invitrogen) and washed twice with 2xSSC, 0.1%SDS, and four times with 1xSSC, 0.1%SDS at 52°C. Between washes, a magnetic particle-collecting unit was used to capture the magnetic beads. As streptavidin strongly binds to biotin, gDNA fragments containing the aforementioned tetranucleotide repeats are captured whereas other fragments are washed away. After the last wash, enriched fragments were removed from the biotinylated probe by denaturing at 95°C and precipitated with 95% ethanol and 3M sodium acetate. To increase the amount of enriched fragments, a 'recovery' PCR was performed in a 25µL reaction containing 1xPCR buffer (10mM Tris-HCl, 50mM KCl, pH 8.3), 1.5mM $MgCl_2$, 10xBSA, 0.16mM of each dNTP, 0.52µM of the SuperSNX24 forward primer, 1U Taq DNA polymerase and approximately 25ng enriched gDNA fragments. Thermal cycling was performed as follows: 95°C for 2min followed by 25 cycles of 95°C for 20s, 60°C for 20s and 72°C for 90s, and a final elongation step of 72°C for 30min. Subsequent PCR fragments were cloned using the TOPO-TA Cloning Kit following the manufacturer's protocol (Invitrogen). Bacterial colonies containing a vector with gDNA were used as a template for subsequent PCR in a 25µL reaction containing 1×PCR buffer (10mM Tris-HCl, 50mM KCl, pH 8.3), 1.5mM $MgCl_2$, 10xBSA, 0.12mM of each dNTP, 0.25µM of the M13 primers and 1U Taq DNA polymerase. Thermal

cycling was as follows: initial step of 95°C for 7min, followed by 35 cycles of 95°C for 20s, 50°C for 20s and 72°C for 90s. PCR products were cleaned using MultiScreen-PCR Filter Plates following the manufacturer's protocol (Millipore). DNA sequencing was performed using the BigDye v3.1 kit (Applied Biosystems, Forster City, CA, USA). Sequencing reactions were precipitated (ethanol and EDTA) and run on an ABI3730 DNA Analyzer. In total, 96 colonies were sequenced, of which 83 contained repeat elements of varying forms.

Sequences were edited using Sequencher v4.7 (Gene Codes Corporation 2008) and primers were designed by eye, based on flanking regions (Table 1). PCRs were performed in a final volume of 10µl containing: 10–50ng of genomic DNA, 0.5U of Taq polymerase (Roche), GeneAmp® 10xPCR Gold Buffer, 0.1mg BSA, 2.5mM $MgCl_2$, 0.4mM of each dNTP and 0.3µM of each primer. The PCR profile was conducted for 30 cycles, for all loci, as follows: 3min at 95°C; 30 cycles of 95°C for 45s, 52-54°C for 30s, 72°C for 45s and a final elongation at 72°C for 30 min. Allele sizes of fluorescently labelled fragments were determined using the size standard LIZ-500 (Applied Biosystems) on an ABI3730 DNA Analyzer followed by analysis with GeneMapper v4.0 (Applied Biosystems). All birds were sexed using the method described by Fridolfsson and Ellegren (1999).

Allelic richness was estimated in FSTAT (Goudet 2002), observed ($H_O$) and expected ($H_E$) heterozygosity for each population was computed using GeneAlEx v.6 (Peakall and Smouse 2006). Deviations from Hardy-Weinberg expectations ($F_{IS}$) across all loci and populations, by population and by gender were determined using the exact test (Guo and Thompson 1992) as implemented in GENEPOP 3.4 (Raymond and Rousset 1995). For sex-linked loci (indicated with § in Table I) $H_O$, $H_E$ and $F_{IS}$ were estimated only for males (n = 10).



## RESULTS AND DISCUSSION

Forty-two Scrub-Robins collected in the Nama Karoo (n = 20) and Koeberg (n = 22) were genotyped for 15 microsatellites. Since the Karoo Scrub-Robin can breed cooperatively, I included only one individual per territory thus avoid biasing the estimators of genetic diversity as a consequence of relateness. All loci were found to be polymorphic. I detected a highly significant departure ($p < 0.01$) from HWE at the Eco3 and Eco54 loci in females but not in males, suggesting that these two microsatellites conform to the expectations of Z-linked loci: all females (heterogametic sex in birds) screened were homozygous. A highly significant departure ($p < 0.01$) from HWE was found when pooling samples from two sites (Nama Karoo and Koeberg) suggesting

a non-panmitic population. A locus-by-locus analysis of autossomal loci revealed no significant deviations in the Great Karoo population as opposed to the Koeberg population where the Eco12 and Eco46 loci did depart from the expectations. $F_{IS}$ values for both loci (Eco12 and Eco46) indicated a deficit of heretozygotes (Table I), suggesting the presence of null alleles, inbreeding or selection. Locus Eco46 seemed to depart from neutrality (Beaumont and Nichols 1996 method). Linkage-disequilibrium was tested both within each population and across populations, using FSTAT (Goudet 2002): critical levels of significance were adjusted using the sequential Bonferroni correction (Rice 1989). The Eco16 and Eco68 loci were found to be in linkage in both analysed populations.

Excluding the markers that are Z-linked (Eco3 and Eco54), under possible positive selection (Eco46) or not in linkage equilibrium (Eo68), the 11 remaining loci had a high probability of identity of full siblings (PIDSibs, Waits *et al.* 2001) equal to $4.9x10^{-5}$ (Koeberg samples). Moreover, when one parent is known, as is often the case, the loci had a probability of exclusion of 0.999. This set of microsatellites is therefore appropriate for estimating extra-pair paternity and relatedness between individuals, measurements that are critical for testing hypotheses relating social structure and the evolution of dispersal in the Karoo Scrub-Robin.

**Table I.** Characterization and variability of 15 microsatellites for *Cercotrichas coryphaeus* from two localities, which are 320 km distant: Nama Karoo and Koeberg.

| Locus | GB | Primer sequence (5'-3') | Repeat motif | Ta (°C) | Size range (bp) | Nama Karoo (n = 20) | | | | Koeberg (n = 22) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | A | $H_E$ | $H_O$ | $F_{IS}$ | A | $H_E$ | $H_O$ | $F_{IS}$ |
| **Eco2** | FJ380188 | F: HEX-TGGCTTCATCCATCCATGAC<br>R: TCTGGAGTTTCCATGTATGC | (TACA)6 | 52 | 152-172 | 5 | 0.697 | 0.800 | -0.148 | 3 | 0.383 | 0.318 | 0.087 |
| **Eco3§** | FJ380189 | F: FAM-GTCTAATCAATACCTATATCTC<br>R: CCCTTAGACGTAAACATATGC | (TATC)7 | 52 | 158-174 | 5 | 0.689 | 0.800 | -0.143 | 3 | 0.444 | 0.364 | -0.220 |
| **Eco12** | FJ380190 | F: NED-<br>R: CGACTCTTGCTCGTCATAC | (AGAT)10 | 54 | 234-290 | 11 | 0.902 | 0.900 | -0.010 | 13 | 0.920 | 0.682 | 0.264* |
| **Eco13** | FJ380191 | F: HEX-CTAAGGATGGTAGGACTACG<br>R: TGGTCATTTACAGTTGGCTTG | (TAGA)21 | 52 | 172-204 | 11 | 0.912 | 0.947 | -0.025 | 10 | 0.886 | 0.773 | 0.123 |
| **Eco14** | FJ380192 | F: HEX-TTGAGTTCTAAGAGTTGCCAG<br>R: GTGTTGCTTCTGGCATGTC | (TAGA)11 | 52 | 229-273 | 10 | 0.817 | 0.650 | 0.295 | 8 | 0.735 | 0.636 | 0.09 |
| **Eco16** | FJ380193 | F: HEX- CTTCTCTGCCTTGGTAACAC<br>R:TCCTTAGCATCAGTCACTAC | (TCTA)5 | 54 | 265-289 | 6 | 0.668 | 0.600 | 0.172 | 7 | 0.698 | 0.590 | 0156 |
| **Eco17** | FJ380194 | F: FAM-CTTAATCCAGTTATAGGCAC<br>R:CTATGTGTTATGAAACTGTC | (TAGA)2 T (TAGA)12 | 52 | 161-181 | 8 | 0.695 | 0.550 | 0.404 | 7 | 0.739 | 0.500 | 0299 |
| **Eco30** | FJ380195 | F: FAM-AGCTTTAATCCGTGTATCAG<br>R: CTTACAGCTTTATCCAACCATC | (GTTT)3... (AAAT)7 | 52 | 231-259 | 8 | 0.849 | 0.700 | 0.203 | 5 | 0703 | 0.571 | 0.192 |
| **Eco32** | FJ380196 | F: FAM- GATACCTTAGAATCATACATG<br>R:CTATCTCTTGGATGTGGGAC | (TATG)7 | 54 | 233-261 | 10 | 0.878 | 0.850 | -0.054 | 8 | 0.850 | 0.863 | -0.011 |
| **Eco42** | FJ380197 | F: NED-TAGTTTGTACTTGCTTGTGG<br>R:GACCTGACTTCATAGACTTG | (TATG)5 | 52 | 229-237 | 3 | 0.626 | 0.450 | 0.242 | 3 | 0.554 | 0.591 | -0.068 |
| **Eco46** | FJ380198 | F: NED-TGATGCTGGGAGTACAGTTG<br>R: AGCAGGAGTGCAGATTGAAG | (TAGA)13 | 52 | 135-173 | 8 | 0.730 | 0.750 | 0.094 | 6 | 0.572 | 0.181 | 0.633* |
| **Eco54§** | FJ380199 | F: HEX-ACCACATAAGGATTGACAATC<br>R: CAAGGAATTTATGTGACATCC | (TATC)12 | 54 | 234-266 | 9 | 1.000 | 0804 | -0.163 | 5 | 1.000 | 0.758 | -0.397 |
| **Eco56** | FJ380200 | F: NED-AGTCACACTGGACAGAATCC<br>R: CAGAAGTCAATGTTGAGTTAG | (TAGA)2.. (TAGA)14 | 54 | 193-221 | 8 | 0.869 | 0.850 | -0.078 | 6 | 0.841 | 0.864 | -0.027 |
| **Eco66** | FJ380201 | F: FAM-GCTACAGATATGGCAATTCTC<br>R: TCAGATTCCTCCACTTCTGC | (ATGT)8 | 54 | 243-263 | 4 | 0.619 | 0.550 | -0.010 | 4 | 0.663 | 0.772 | -0.178 |
| **Eco68** | FJ380202 | F: HEX-CACATCTTATACTTTGCAGC<br>R: TTAGCATCAGTCACTACAGC | (TCTA)10 | 54 | 206-242 | 6 | 0.641 | 0.520 | 0.372 | 7 | 0.643 | 0.591 | 0.082 |
| **Overall** | | | | | | | 0.699 | 0.695 | 0.147 | | 0.670 | 0.570 | 0.147 |

GB: GeneBank Accession numbers: F: forward primer; R: reverse primer; HEX, FAM, NED: fluorescent labels; Ta: annealing temperature; bp: base pair; n: number of individuals scored; * significant departure from Hardy-Weinberg expectation, $\alpha$=0.01; § Z-linked loci.

CHAPTER THREE

# Microgeographic socio-genetic structure of an African cooperative breeding passerine revealed: integrating behavioural and genetic data



## INTRODUCTION

Natal dispersal is often defined as the movement of an organism from its birthplace to the site of its first reproduction, whereas breeding dispersal is described as the movement of individuals between subsequent breeding territories (Stenseth and Lidicker 1992, Clobert *et al.* 2001). Fundamentally, these movements are a behavioural response to environmental and social pressures on natal and breeding sites, molded by selection in order to optimize individual fitness.

Despite its risks, dispersal is a common phenomenon, which suggests that selective pressure maintains its prevalence (Ronce 2007). The two primary and non-mutually exclusive adaptive advantages of dispersal over philopatry are attributed to: (1) a reduction in competition between relatives (Hamilton and May 1977) and, (2) the lower likelihood of inbreeding depression (Bengtsson 1978, Perrin and Goudet 2001). Support for inbreeding avoidance comes from multiple studies, which suggest that sex biases in dispersal are common (e.g. Dobson 1982, Clutton-Brock 1989, Handley and Perrin 2007). In numerous species, members of one sex tend to disperse more frequently and further than individuals of the opposite sex; but predicting which sex is the disperser is not straightforward. Based on empirical data, Greenwood (1980) proposed a nexus between the selective pressures underlying the breeding systems and directionality of dispersal biases. He postulated that the philopatric sex should be the sex that is most involved with territory acquisition: usually females in mammals, and males in birds (Greenwood 1980, Greenwood and Harvey 1982).

Theoretical models seeking to explore the mechanisms underpinning the evolution of dispersal have also revealed that living in social groups plays a crucial role in promoting sex-bias in dispersal (Perrin and Mazalov 2000, Perrin and Goudet 2001). When cooperation develops among kin of the same sex, philopatry is reinforced, and the risk of inbreeding depression selects for dispersal in the opposite sex. Therefore, inbreeding avoidance might be the most important selective force promoting the development of asymmetric dispersal behaviour (Perrin and Goudet 2001, Handley and Perrin 2007). Thus, combining information about dispersal behaviour, the social and genetic mating system, kinship and gene flow is central to gaining insight into the processes that underlie the complex dispersal patterns in social-living species. Despite the obvious need for such data, there is a paucity of such studies due to: a) the difficulty of obtaining a direct measure of dispersal in natural populations (small mammals: e.g. Stenseth and Lidicker 1992; birds: Koenig and Dickinson 2004), b) the traditional focus of population geneticists on population dynamics over larger spatial scales (e.g. Slatkin 1987, Wilson and Rannala 2003), and c) the lack of attention given to the influence of social organization in generating nonrandom genetic patterns

(Epperson 2003). Consequently, to date, only a few studies of social birds have explicitly taken advantage of both behavioural and genetic data to investigate the effect of sex-biased dispersal on the social and genetic structure of a population, and hence the role of dispersal as an evolutionary mechanism with which to avoid inbreeding depression (e.g. Superb Fairy-Wren: Double *et al.* 2005; Apostole bird: Woxvold et al. 2006; White-breasted Thrasher: Temple *et al.* 2006/2009; Vinous-throated Parrotbill: Lee *et al.* 2010).

In this study I develop a framework that combines behavioural data with individual–based genetic analyses to explore the functionality of sex-biased dispersal in mediating the antagonistic effects of cooperation and inbreeding pressures in a social bird, the Karoo Scrub-Robin *Cercotrichas coryphaeus*. The Karoo Scrub-Robin is a facultative cooperative breeder (Oatley 2005): during the breeding season approximately 15% of breeding pairs are assisted by one or more retained male helpers (Lloyd *et al*. 2009). As in the majority of cooperatively breeding birds (Koenig and Dickinson 2004), Karoo Scrub-Robin helpers are offspring of the breeding/attending pair who have delayed dispersal and postponed reproduction (Lloyd *et al.* 2009). These social bonds may result in the spatial aggregation of male kin and, as a consequence, the clustering of alleles that are identical by descent (IBD, Malécot 1967). This microgeographic genetic structure will in turn affect the spatial social organization and ultimately the evolution of dispersal in the Karoo Scrub-Robin. Although I recognize that temporal heterogeneity of the environment can affect dispersal behaviour and its evolution, at this small temporal scale other proximate factors are thought to be more relevant in determining the intensity and direction of dispersal. Therefore, I focus on the feedback loop between dispersal behaviour and social interactions (Figure 1) in order to test whether female-biased dispersal has evolved to reduce inbreeding depression by limiting the spatial overlap with male relatives. Whereas female biased dispersal is expected to separate relatives spatially and thus reduce the probability of mating with kin, delayed natal dispersal and philopatry of males, induced by cooperation, in contrast, should increase the effect of the bias and cause the spatial aggregation of related alleles. To accomplish my aim, I quantified differences in dispersal distance between males and females; measured how genetic similarity (identity by descent) varies with geographical distance; estimated relatedness between mated pairs; and compared how close social and genetic mating systems match.

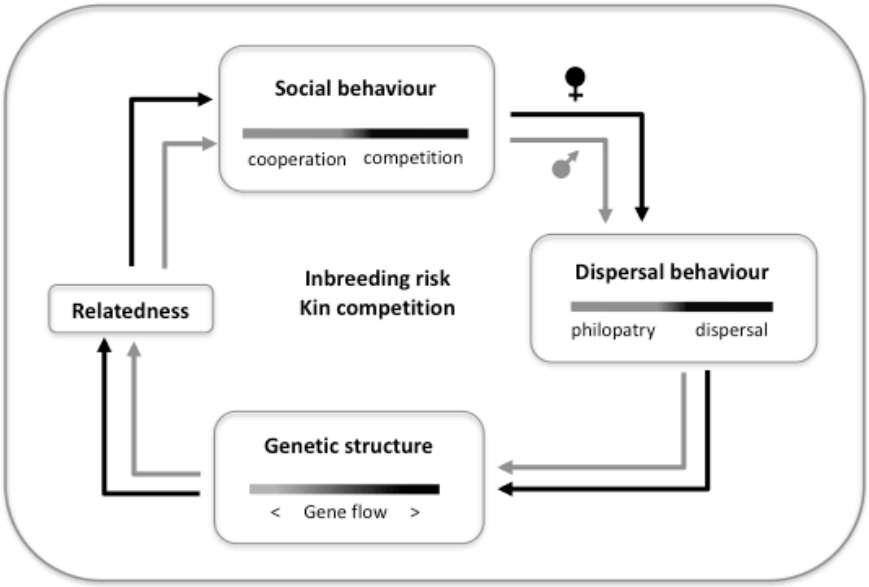


**Figure 1.** Conceptual framework with which to study dispersal behaviour in avian social systems: the balance between cooperation and competition in selecting for dispersal and thereby avoid the costs of inbreeding depression and reduce competition between relatives. Black and grey represent the expected behaviour and genetic structure for females and males, respectively.



## METHODS

### Study species and field study

The Karoo Scrub-Robin *Cercotrichas coryphaeus* is a medium sized (18 - 23g), ground-feeding insectivore endemic to the semi-arid regions of southern Africa (Oatley 2005). Pairs are sedentary and defend year-round territories. Clutch-size varies from two to four eggs, with incubation lasting 13 - 15 days and only females incubate. Male offspring appear to be retained on the natal territory until they recruit to a breeding vacancy; however, most of the pairs, 85±2%, breed without any assistance from helpers. Nestlings are fed by both parents and sometimes by attending helpers. The majority (93%) of female recruits bred in their first year (Lloyd et al. 2009), thus one year is assumed to be the generation time in the study population.

I studied a population of Karoo Scrub-Robins occupying up to 104 territories within a 260 ha area

(the study area) of the 2,900 ha Koeberg Nature Reserve (-33º 41′ S, 18º 26′ E), on the west coast of South Africa (Figure 2). The study population occupied an area dominated by sand-plain Fynbos and dune ticket, disrupted by some cleared areas. Territories were evenly distributed and the average territory size was 2.5 ha.

This population has been studied intensively during the breeding season, August to early November 2000-2008 (Lloyd *et al.* 2009). Banding of adults and all nestlings reaching nine days of age with a unique combination of three colour bands and a numbered metal band commenced in 2001, with 331 individual adults and 616 nestlings banded by the end of 2007. A small blood sample (20-60 µl) was taken via brachial venipuncture from captured adults and from nestlings that survived to three days of age. All birds occupying the study area were re-sighted during the breeding season each year, and the positions of all territories were mapped onto aerial photographs. To check for dispersal beyond the confines of the study area, I also re-sighted all birds within a three-territory radius beyond the entire perimeter of the study area.

For the genetic analyses, I genotyped a total of 238 individuals sampled from three breeding seasons (Table I). Different subsets of individuals were used according to the question being tested, as indicated in each section of the methods below. Individual sampling periods were not independent because some individuals were present across multiple years.




**Figure 2**. Map outline of the study area in the Koeberg Nature Reserve. Inset shows the geographical location of the study population on the west coast of South Africa.

**Table I.** Sample size, sex and breeding status of the Karoo Scrub-Robins used for genetic analyses in each breeding season.

| Sampling period | Females | | Males | | | Total | Breeding pairs |
|---|---|---|---|---|---|---|---|
| | dominant | offspring | dominant | helpers | offspring | | |
| 2003 | 8 | 15 | 8 | 3 | 18 | 52 | 8 |
| 2005 | 23 | 0 | 35 | 4 | 0 | 62 | 11 |
| 2007 | 44 | 4 | 60 | 9 | 7 | 124 | 32 |
| **Total** | 75 | 19 | 103 | 16 | 25 | 238 | 51 |

**Microsatellites genotyping, error rate and PCR-based sexing**

A total of 238 birds were successfully genotyped for eleven microsatellites specifically developed for the Karoo Scrub-Robin (chapter two). Microsatellites were multiplexed in four reactions as follows: Karooplex1 (Eco2 + Eco14 + Eco17), Karooplex2 (Eco13 + Eco30 + Eco32), Karooplex3 (Eco12 + Eco16 + Eco32) and Karooplex4 (Eco56 + Eco6). PCR amplifications were performed in a final volume of 10 µl containing: 10 – 50 ng of genomic DNA, 0.5 U of Taq polymerase (Roche), GeneAmp 10× PCR Gold Buffer, 2.0 - 2.5 mM $MgCl_2$, 0.3 mM of each dNTP and primer concentrations of 0.15 µM (Eco2, Eco12, Eco13, Eco14, Eco16, Eco32) and 0.2 µM (Eco17, Eco30, Eco56, Eco66). Thermal-cycling proceeded as follows: 3 min at 94°C; 30 cycles of 94°C for 45 s, 52 - 54°C for 30 s, and 72°C for 45 s with a final extension at 60°C for 20 min. Allele sizes of fluorescently-labelled fragments were determined using the size standard LIZ-500 on an ABI 3730 DNA Analyser followed by analysis with GeneMapper version 4.0 (Applied Biosystems, Forster city, CA, USA).

Studies that rely on the relationships of individual genotypes can be influenced by genotyping errors (Hoffman and Amos 2005). I randomly chose 10 (4%) of the 238 samples as 'blind' samples. These samples went throughout the entire process again (from extraction to scoring) accumulating all the possible causes of genotyping error. The observed mismatches were then used to calculate the error rate per allele as the ratio of total mismatches to total alleles typed. This error rate was then incorporated into the paternity analyses.

The sex of all birds included in the genetic analysis was determined using a PCR-based assay (Fridolfsson and Ellegren 1999). Thirty-two birds sexed in the field as breeding females, on the basis of nesting behaviour, were used to check the accuracy of the DNA assay.

**Genetic diversity**

Microsatellites were examined for deviations from linkage equilibrium and Hardy–Weinberg (HW) expectations using GENEPOP 3.4a (Raymond and Rousset, 1995). A set of 30 randomly selected breeding individuals from two breeding seasons (2005 and 2007) were used to independently implement the exact tests through a Markov chain (Guo and Thompson 1992) with 1000 demorizations, 400 batches and 1000 iterations per batch. I used GenAlEx v.6 (Peakall and Smouse 2006) to calculate the observed ($H_O$) and expected heterozygosity ($H_E$) per breeding season, as well as $F_{IS}$ values for breeders (n = 32) and offspring (n = 21). Given the aims of this study, I estimated the probability of two randomly selected siblings having the same multilocus genotype (PIDSibs, Waits *et al.* 1997), a measure of statistical power for the set of loci. Moreover, the *a priori* ability of the set of loci to detect paternity inconsistencies, measured as the probability of exclusion ($P_{EX}$), was computed using GenAlEx v.6 (Peakall and Smouse 2006).

**Philopatry and dispersal**

*Behavioural estimates*

I measured natal dispersal distances for 112 Karoo Scrub-Robins (82 males and 30 females) that were first banded as nestlings and were subsequently recorded as breeders within or adjoining the study area. Breeding dispersal occurred when an individual holding a breeding position was re-sighted holding a breeding position in another territory. I measured breeding dispersal distances across the years, for 332 males and 290 females. Divorce was defined as a particular case of breeding dispersal: when the breeder moved to pair with a new mate despite the survival of the disperser's previous mate. The Wilcoxon-Mann-Whitney test was used to examine the effects of sex on natal and breeding dispersal, as implemented in R (R Development Core Team 2010). Dispersal distance was estimated as the number of territories traversed and then from the known average territory area, converted into meters, assuming that territories are best described as a circular shape, as follows: $\sigma$ = distance dispersed per generation = [2 * √ (mean territory area/$\pi$) * median # of territories traversed] / generation time.

*Genetic estimates*

To examine the pattern of distribution of genetic variation at a fine spatial scale, I utilized a multivariate approach (spatial autocorrelation) as implemented in GenAlEx v.6 (Peakall and Smouse 2006). The two matrices I used to investigate spatial autocorrelation were a geographic distance matrix and a pairwise genetic distance matrix. I expressed geographic distance as the number of territories separating each mated pair (dominant territory holders) along a straight line separating them. These matrices were then used to estimate an autocorrelation coefficient, *r,* between pairs of individuals for a

given distance class. This coefficient ranges between –1 and +1 measuring the genetic similarity between individuals (Smouse and Peakall 1999). The null hypothesis (i.e. random distribution of alleles) for each spatial category was constructed by shuffling the observed genotypes 999 times among the different distance classes. After each permutation, a new *r*-value per lag (distance class) was obtained and the upper (97.5%) and lower (2.5%) bounds of the 95% confidence interval were constructed. Any observed value outside these limits would be considered statistically significant. When positive spatial genetic autocorrelation is detected, the distance class where *r* first intercepts the x-axis in the autocorrelation plots reveals the extent of non-random genetic structure (Peakall et al. 2003). Data from two breeding seasons (breeders and helpers only; $n_{2005}$ = 62 and $n_{2007}$ = 111) were explored independently and separate tests were implemented for males and females. Because of the smaller sample size of dominat individuals obtained for breeding season 2003, I excluded this from the autocorrelation analysis (see Epperson 2003 for a discussion on the importance of sample sizes).

**Relatedness among social mates**

To test the prediction that female-biased dispersal is an effective mechanism to avoid pairing with a related male, I compared the average genetic similarity between observed social mates (n = 32; breeding season 2007) with a simulated set of randomly mated pairs. I used a likelihood-based estimator of *R* (QG*R;* Queller and Goodnight 1989) to obtain a direct measure of relatedness. A Monte Carlo approach was used to construct the null distribution of the test statistic QG*R*. I randomized and re-sampled 64 adult birds (32 social pairs) 10 000 times, to create 50 potential pairs, but prevented the same pair from being sub-sampled twice for the same set of 50 pairs. The observed test statistic (QG*R*avg*SocialPairs*) was then compared to the distribution of the simulated values (QG*R*avg*Sim)* to examine if the observed value was significantly higher than expected for a random mating population. The estimated probability of obtaining a result that exceeded the observed value under the null hypotheses was estimated as $p$ = [(# QG*R*avg*Sim* ≥ QG*R*avg*SocialPairs*) / total # simulations]. All statistical analyses were performed using R (R Development Core Team 2010; scripts are available from AMR upon request).

**Kinship, parentage and extra-pair paternity**

Relatedness values were 'calibrated' by computing the QG*R* value between 27 same-clutch offspring dyads that had been tested and confirmed to be full siblings. The expected average *R*-value for a first-order relationship is 0.5 whereas the observed value was 0.494 ± 0.036. Therefore, pairs with *R*-values ≥ 0.458 were designated as first-order relatives (parent-offspring or full siblings).

Parentage was assessed for 33 chicks using a two-phase approach. First, parentage was confirmed

via Mendelian checks for all offspring-parent social dyads. The occurrence of up to one inconsistency across all loci was regarded as an occasional mismatch caused either by mutation or genotyping error. The estimated error rate for the data set was about 2%, within the range suggested by Ewen *et al.* (2000) to have little effect. In this situation, parentage was assigned to the putative parent. In all the other cases with two or more mismatches, I determined if one or both parents are excluded. Exclusion of the putative father implied an extra-pair offspring. Exclusion of the putative mother was taken as evidence of brood parasitism or sample handling error if the inconsistency remained after the genotypes were re-scored two further times. In the second step, any offspring whose putative father had been excluded was assigned a paternity using the likelihood approach implemented in CERVUS v3.0 (Kalinowski, Taper and Marshall 2007).

CERVUS v3.0 (Kalinowski, Taper and Marshall 2007) takes into account the proportion of candidate males sampled and allows for errors in the genetic data. Likelihood ratios (hypothesis of the alleged father being the true father *vs* the hypothesis that the candidate father is not the true father) were estimated for each candidate father given the offspring and mother's genotypes as well as the allele frequencies in the population. Simulations were then performed to determine the resolving power of the loci to assign parentage and calculate a statistic delta (D), the criterion that allows assignment of paternity to the most likely male with a known level of statistical confidence (Marshall *et al.* 1998). I simulated 10 000 offspring, 90 candidate males (reasoning that the actual sire is a male from the study area) allowing 35% of the population to be unsampled, specifying that 98% of the multilocus genotypes were complete and set an error rate of 0.02. LOD scores were calculated for each candidate parent; negative, zero, or positive LOD scores respectively imply that the alleged father is less likely, equally likely, or more likely to be the true father than any other individual randomly chosen from the population.

Paternity was only assigned when the two criteria suggested by Marshall *et al.* (1998) were attained: i) D > 3.0 and ii) strict confidence level = 95%. Otherwise, I concluded that the actual sire was an unsampled male. Whenever extra-pair young were detected, I compared QG*R* values between social and genetic parents to determine whether females sought extra-pair copulations (EPC) with less related males.



## RESULTS

### Field data on dispersal

*Natal dispersal*

There was a strong sex bias in natal dispersal distance (Wilcoxon-Mann-Whitney: W = 12.5 *P* =

0.016). Males showed a leptokurtic distribution of dispersal distances (kurtosis = 1.95) whereas the female dispersal distribution was platykurtic (kurtosis = 0.66). Male offspring appear to be retained on the natal territory until they recruit to a breeding vacancy on a nearby territory (median = 1 territory, range: 0 – 4, n = 82; Figure 3a). Although only 9% of males inherited their natal territory, 96% obtained a breeding vacancy within two territory distances of their natal territory, with most settling on a neighbouring territory that shared a border with their natal territory. The dispersal distance for males, $\sigma_{males}$, was 178 m/generation. Female offspring disperse as early as a month after fledging, and the ones that settled within the study area dispersed further than males (median = 5 territories, range = 0 – 14, n = 30; Figure 3a). Further, the smaller number of female recruits (27% of the total) suggests that most female offspring disperse beyond the confines of the study area. The median dispersal distance for females, $\sigma_{females}$, was 892 m/generation.

*Breeding dispersal*

Moving between breeding sites was uncommon. Overall, 8.6% of females dispersed to a new breeding territory between years (n = 290 between-year survival events), whereas just 2.4% of breeding males (n = 332) undertook breeding dispersal (Figure 3b); however, this asymmetrical trend was not significant (Wilcoxon-Mann-Whitney: W = 5, $P$ = 0.4218). Most breeding dispersal in both males and females was to a neighbouring territory, but some females did undertake longer-range breeding dispersal. Among 23 dispersing females with a known history (out of 27), two moved to an adjoining territory with their original mate, whereas dispersal coincided with divorce (n = 4) or the disaparence of their mate (n = 17) in the others. Of the latter, the dispersal of seven females coincided with a son, still present on the territory, inheriting its natal territory following the death of its father. Among seven dispersing males with a known history (out of eight), four moved to an adjoining territory with their original mate, two dispersed after the death of their mate, and one dispersed after being forcibly divorced from its mate by an intruding male.

**Genetic diversity**

Of the 11 microsatellites used, only one locus (Eco12) deviated from HW expectations. Eco12 had a significant excess of homozygotes across the two distinct sampling periods, suggesting the presence of null alleles, since no selective pressure was previously found at this locus, for this population (chapter two). I therefore excluded this locus, implementing all further analyses using data from the ten remaining loci. The average $H_O$ was 0. 644 ± 0.020, $H_E$ = 0.510 ± 0.011 and the number of alleles per locus ranged from three to 12. $F_{IS}$ values were not significantly different from zero ($F_{IS\ Breeders}$ = 0.09 and $F_{IS\ Offspring}$ = -

0.06). The ten loci delivered a PIDS = $1.6x10^{-4}$ and a $P_{EX}$ = 0.99 suggesting sufficient discriminatory power for discerning parent-offspring relatives. The molecular sexing matched the behavioural sexing of 32 females with 100% accuracy.

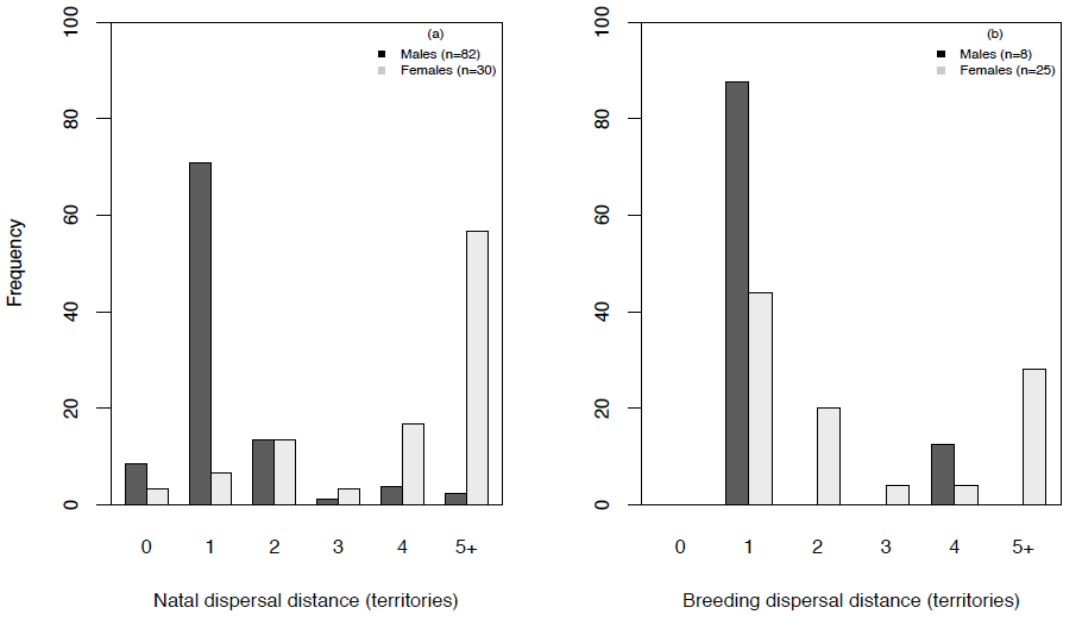


**Figure 3.** Natal (a) and breeding (b) dispersal of males (black) and females (grey), measured as the number of territories separating initial and settling sites. The y-axis represents the proportion (percentage) of individuals that settle at each territory class and was truncated at class 5. The class 5+ contains all individuals that dispersed 5 or more territories.

**Fine-scale spatial genetic structure**

Spatial autocorrelation analysis combining both sexes found no significant relationship between genetic similarity and distance ($r \sim 0.002$, outside 95% CI). To test for sex-specific patterns I split the data into males and females within each sampling period (Figure 4a-c). Partitioning the data revealed strong evidence of spatial genetic autocorrelation among males, but not females. The *r*-value, for males, was significantly positive within a one-territory radius ($r$ 2005 = 0.49 and $r$ 2007 = 0.32, $p < 0.05$) after which it decreased to become non-significant until it crossed the x-axis at four territories distance and became significantly negative. In contrast, for females, the correlation coefficient was significantly negative among neighbours and then increased to values not significantly different from zero ($r = 0$, the value expected for genetically unrelated individuals).

Removing the within-territory pairwise comparisons of breeding and helping males (13%) from the analysis did not change the result ($r$ 2005 = 0.062 and $r$ 2007 = 0.038, $p < 0.01$; Figure 4b). Overall, the patterns persisted across seasons. The results show that alleles are non-randomly distributed in space and the extent of non-random positive genetic structure is four territories (x-intercept in the autocorrelation plots).

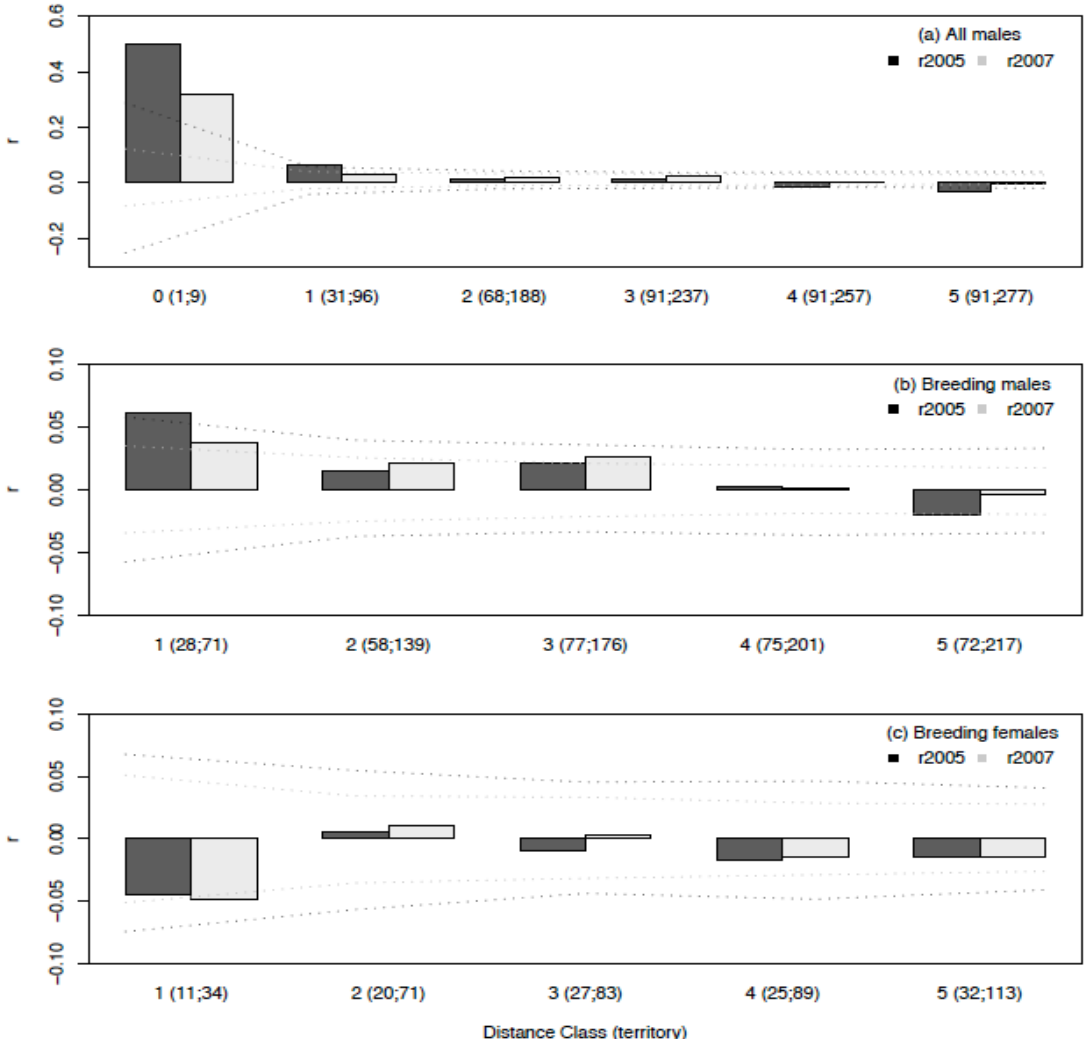


**Figure 4.** Spatial correlograms produced by plotting the autocorrelation coefficient $r$ as a function of distance for: (a) all males, (b) breeding males, and (c) breeding females. Black and grey refers to the 2005 and 2007 breeding seasons, respectively. The upper and lower bounds of the 95% confidence interval around the mean value ($r$ = 0) of the null distribution of a random distribution of alleles in space are represented as dashed lines (black for 2005 and grey for 2007). The number of comparisons at each distance class is reported below the x-axis (n for 2005; n for 2007).

**Relatedness, parentage and extra-pair paternity**

Overall, relatedness among social pairs was low, *R*avgSP = 0.039 ± 0.144 SD, but significantly higher than would be expected by chance ($p$ = 54 / 10,000 = 0.005; Figure 5). The averaged QG*R* between breeding males and attending helpers was 0.354 ± 0.04, slightly lower than expected for first-order relatives, according to my own 'calibration' (see methods).

After examining the possible mismatches of 33 nestlings from 12 broods, via Mendelian inheritance, all but one bird matched the female attending the nest (three allelic mismatches out of 10 loci). Despite social monogamy, paternity analysis revealed that females do sometimes engage in extra-pair copulations, but in all instances it was with a helper male when breeding in a cooperative group.

When implementing the analysis with the identity of the mother included, 30 nestlings fulfill the two criteria previously defined to accurately assign parentage, whereas the remaining three had relaxed confidence limits of assignment. Six offspring (18%; one entire brood of three nestlings and one nestling from each of three additional clutches) excluded their social father as the sire. These instances of extra-pair paternity occurred in pairs whose QG*R* < 0.144, and in all cases the helper was attending a first-order relative male (0.399 < R < 0.562). The remaning 27 had the social father assigned as the sire, mostly with high confidence limits of assignment as estimated in Cervus. Only once did the female mate with a helper more closely related to her than the social mate ($R_{Female\text{-}Helper}$= 0.26 *vs*. 0.064).

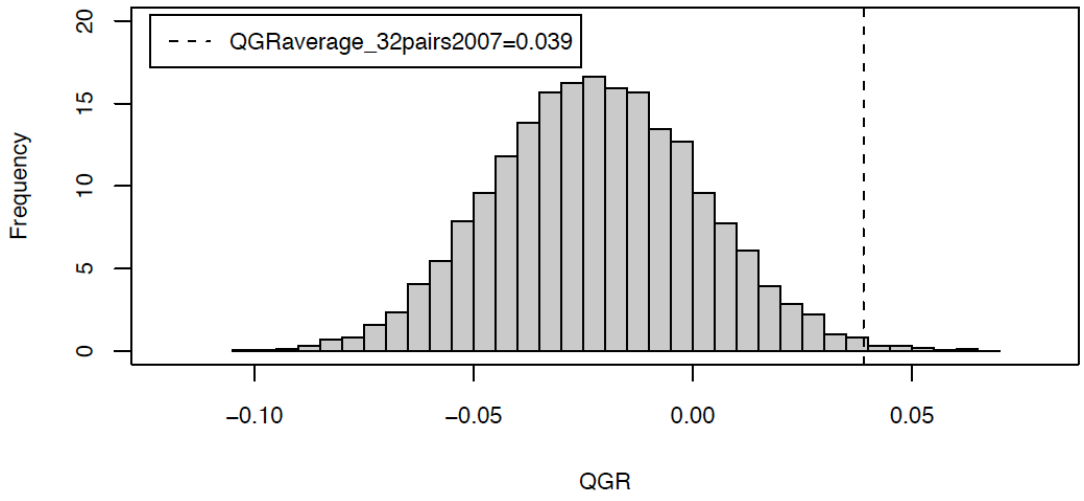


**Figure 5.** Distribution of 10 000 MC simulated values of QG*R* (Queller and Goodnight 1989 estimator). The observed average relatedness between social pairs (QG*R*avg*SocialPairs*) is plotted as dashed line on the distribution.

## DISCUSSION

Philopatry and dispersal are behaviours of individual organisms that have ecological and genetic consequences for the entire population. Understanding their complex patterns and consequences in socially organized species is only possible by placing observational field data into a population genetic framework (e.g. Harris *et al.* 2009). Whereas behavioural data provided information on the timing of dispersal, social organization and the possible proximate causes of dispersal, genetic data yield valuable insights into the spatial arrangement of kin, mating system and, the evolutionary causes of dispersal.

### Male philopatry: local relatedness and social bonds

The observed sex-biased dispersal behaviour in the Karoo Scrub-Robin indicates that the costs and benefits of leaving the natal site are different among the sexes. Males delay dispersal if they are unable to obtain a breeding vacancy in the close vicinity. Consequently, most males do not breed in their first year, remaining instead on the natal territory and helping the breeding pair as part of a cooperative family group (Lloyd et al. 2009). This strategy may avoid the costs associated with longer-distance dispersal while also providing opportunities for direct or indirect reproductive benefits. However, the interest of juvenile males and both breeding adults must coincide.

Recently, Lloyd et al. (2009) demonstrated that Karoo Scrub-Robin male helpers assist adults and effectively contribute to the overall increase of annual productivity of the breeding pair. The present study, by including relatedness estimates, enables me to also understand the reproductive benefits of such cooperative behaviour. The genetic data revealed that when living with non-relative females, helpers may obtain a share of reproduction, and thus besides gaining indirect fitness by providing care to half-sibs, helpers also accrue direct genetic benefits. Benefits from familiar-group membership, other than genetic, can also play an important role in the decision of a juvenile male to postpone dispersal and help. The fact that an individual, which delayed dispersal, stayed in the parental site, and not elsewhere, may reduce the need to outcompete the same cohort of males in order to obtain a territory, and concurrently ensure access to feeding sites (Ekman et al. 2004), and thus enhance its own reproductive success later in life (Covas and Greisser 2007).

Most of the male Karoo Scrub-Robins, 96%, gained their first breeding position within two territories of their natal site. Once a male fills a breeding vacancy, subsequent breeding dispersal is exceptional. A site change to an unpredictable area might be costly and so adult males only moved in the event of a breeding vacancy opening up in an adjoining territory. The social behaviour between males translated, as per my predictions, into a striking pattern of clustering of alleles identical by descent, within

the distance of one territory from the natal site. Such strong philopatry and the associated genetic consequences have been observed only in certain cooperatively breeding birds, such as the White-breasted thrasher (Temple *et al.* 2006), the Brown Tree-Creeper (Doerr and Doerr 2006) and the Superb Fairy-Wren (Cockburn et al. 2008).

**Female dispersal and breeding system**

Natal dispersal in Karoo Scrub-Robins is female biased, with females dispersing earlier and further than males. This is a common pattern in other cooperative breeding birds (Koenig and Dickinson 2004) and conforms to the expectations of the resource-defense hypothesis proposed by Greenwood (1980), where the sex that defends resources, in this case territory, benefits from philopatry and the opposite sex disperses. This greater propensity of females to disperse appears to remain throughout their life, with 10% of adult females changing breeding sites, as opposed to only 1% of adult males. Females always dispersed when a son inherited the territory, suggesting female breeding dispersal may often be driven by selection to avoid fitness consequences of breeding with kin.

As predicted, the long-distance dispersal strategy prevented the clustering of alleles that share the same recent ancestry. The negative values of genetic correlation among neighbouring females indicate a frequent introduction of new alleles by immigrant females, and corroborates the field data that showed that only a small portion of females recruit in the study area.

The socially monogamous breeding system of the Karoo Scrub-Robin observed was not reflected in the actual breeding system as revealed with the genetic data. Females do occasionally engaged in extra-pair copulations with subordinate helper males when breeding in a cooperative group. The observed extra-pair paternity rate (EPP; approximately 18%) is intermediate between levels reported in other cooperatively breeding birds (e.g. 76% in Superb Fairy-Wren, Double et al. 2005; 7% in White-breasted Thrasher, Temple et al. 2006). The detection of within-group paternities challenges the hypotheses of genetic benefits. These hypotheses postulate that females gain indirect benefits, by seeking good/compatible genes and thus enhance offspring's genetic quality (Petrie and Kempenaers 1998, Blomqvist et al. 2002). These hypotheses further predict that extra-pair males should be less related to the females than the social mate and also that EPP should correlate with the relatedness of a social pair. Yet, females sought extra-pair copulations from subordinates helper males within the group, and in one case, a female did breed with a related male, and in a second instance a female chose an extra-pair male that was more closely related (second-order relative) to her than her social mate. These observations are in line with reproductive skew theory (Clutton-Brock 1998), which predicts that dominant females and males make reproductive concessions to intra-group subordinate males, when they are closely related and its

prospect of independent reproduction is small, and by doing so can improve the group productivity. Besides an increase in the group productivity, Karoo Scrub-Robin females also gain direct benefits. Helper reduces female's mortality risk by allo-feeding, sentinel and mobbing behaviour (Lloyd et al. 2009). These results further suggest that female breeders significantly influence social structure in the Karoo Scrub-Robin.

**Adaptive bases of philopatry and dispersal in social species**

Living in a socially structured population can result in a prominent risk of mating with relatives (Koenig and Haydock 2004) unless there is a countervailing mechanism. The genetic results indicate the risk of cooperation in increasing coancestry between mated pairs. In agreement with formal predictions (Dobson 1982, Perrin and Mazalov 1999, Goudet and Perrin 2001), my study suggests that moderate inbreeding can induce a strong sex bias in dispersal. The observed asymmetry of dispersal towards females seems to have a phylogenetic component in birds (Greenwood 1980); however, a pattern of extreme philopatry in males is less common.

The extended social organization in familiar groups in the Karoo Scrub-Robin, which stems from differential dispersal behaviour between males and females, is a consequence of inbreeding avoidance and kin cooperation. Strong philopatry by males, induced by cooperation, creates a kin structure, which then prompts dispersal of females to avoid breeding with kin. The following lines of evidence suggest that the observed female bias in dispersal behaviour might be the predominant mechanism with which Karoo Scrub-Robins avoid inbreeding depression. First, the extensive field data revealed that natal dispersal in females occurred as early as one month old, soon after becoming independent. Because the breeding season in Koeberg is seasonal (August to October) competition for mates and resources should not be the cause of female biased dispersal. Second, the tendency and intensity of breeding dispersal was still more marked in females. Third, Karoo Scrub-Robin females actively avoided incest. Observational data revealed that females abandoned the breeding site soon after the vacancy as a breeder, left by the death of their social mate, had been taken by a direct male offspring. Fourth, despite the aggregation of males, the influx of non-related alleles by immigrant females and the reshuffling of alleles in subsequent generations appear to keep the population in a steady state where observed heterozygosity conforms to the theoretical expectations for a randomly mating population.

The formation of familiar groups and male-kin structure in the Karoo Scrub-Robin may be primarily under the control of breeding males. In the southern African semi-arid zone, food abundance is highly seasonal and associated with rainfall (du Plessis et al. 1995). It is the increase of insect biomass associated with rainfall that triggers insectivore breeding (Dean and Milton 2004). As arthropod

availability decreases, remaining in a familiar group until the acquisition of a territory may facilitate coping with food shortage during the driest part of the year and thus improve their survival. However, the presence of extra-pair males results in high susceptibility of paternity loss. Therefore, breeding males should only accept related males as helpers if they are not capable of raising a brood alone, thus risking being cuckolded but enhancing the survival of its own offspring. The cooperation between males goes beyond the natal site. When the habitat is locally saturated, as is the case in the Koeberg population, neighbours are more likely to compete than cooperate. Nevertheless, by establishing a territory in the vicinity, males may benefit from nepotistic assistance (familiar-based recognition has been proposed as a mechanism to enable discrimination of kin from non-kin in bird societies; Komdeur and Hatchwell 1999) or greater levels of tolerance as opposed to aggressive interactions with non-relatives (e.g. Watson 1994, male Red Grouse *Lagopus lagopus* are more aggressive to non-kin than to relatives). The prolonged parental care and the reduced aggressiveness towards related males may have an effect on fitness of both interacting males and thus be a selective factor for male philopatry.

**Dispersal behaviour and scaling-up to the landscape level**

In an ideal scenario where dispersal had no costs or limits, any organism experiencing spatial heterogeneity in the environment would maximize its fitness by simply tracking suitable habitat. However, as shown in this study, dispersal is a complex phenomenon that despite being conditioned by intrinsic traits, such as age and gender, also depends on individual level behaviour and social factors. Because dispersal affects a multiplicity of ecological and evolutionary processes, such as meta-population dynamics (Hanski and Gilpin 1997), gene flow (Wright 1943), local adaptation (Garant et al. 2007) and species range limits (Duckword and Badyaev 2007) I emphasize the importance of understanding individual differences in dispersal behaviour, its proximate and ultimate causes.

CHAPTER FOUR

# A tight balance between natural selection and gene flow in the southern African arid-zone endemic bird



## INTRODUCTION

Nearly a century ago, Joseph Grinnell (1924) underscored the importance of environmental-based selection in maintaining species adapted to their environments: *'The only mechanism of which I have ever heard, that I can reasonably conceive as operating to permit of the molding of species under the stress of environments is natural selection'*. Grinnell's verbal argument was later formalized to set the framework for subsequent research about the role of selection in promoting adaptation of populations to local conditions. Theoretical models predict that local adaptation can occur provided that selection is strong relative to migration. Under the *migration-selection model* (Levene 1953, Bulmer 1972, Endler 1977, Lenomard 2002) two populations (1 and 2) inhabit two environmental patches (X and Y) and exchange *m* migrants per generation. Alleles *A* and *a* have antagonistic fitness, whether they occur in patch X or Y. In this model, migration tends to limit or prevent differentiation among populations whereas natural selection removes less fit immigrants from the population. Traditionally, restricted gene flow has been seen as a requisite for local adaptation to develop. Nonetheless, theoretical (e.g. Endler 1977, Doebeli and Dieckmann 2003) and empirical evidence (e.g. Smith et al. 2001, Ogden and Thorpe 2002, Mullen and Hoekstra 2008, Nosil et al. 2009a) challenge this viewpoint: strong disruptive selection can maintain adaptation to different environments despite very high levels of gene flow, particularly if selection acts on one or a few traits with large effects on fitness.

Aside from selection, historical demographic processes (e.g. colonization of a new habitat) can also generate variation in allele frequencies and/or phenotypes (Endler 1977). Therefore, comparing loci that track demography (neutral) with presumed adaptive variation (genetic and/or phenotypic) is essential if we were to disentangle the effects of non-adaptive versus adaptive processes and hence gain an understanding of the mechanisms underlying the observed spatial pattern in phenotype and/or genotype across the landscape of interest (e.g. Rosenblum et al. 2007).

Many adaptive responses to local environmental conditions are thought to be mediated by physiological mechanisms (Ricklefs and Wikelski 2002). Compelling examples of local adaptation induced by selective pressures acting on molecular physiology have emerged from studies of Deer Mice *Peromyscus maniculatus* (Storz et al. 2007), Rufous-collared Sparrows *Zonotricha capensis* (Cheviron and Brumfield 2009), Yellow-billed Pintails *Anas georgica* (McCracken et al. 2009) and the Bar-headed Goose *Anser indicus* (Scott et al. 2010). Most of these studies have focused on mechanisms that enable individuals to persist along extreme hypoxic gradients. However, the mechanisms that facilitate adaptation to other extreme habitat types remain largely unknown. For instance, daily temperature fluctuations and,

the limited and unpredictable availability of water and food in arid environments might affect energy and water requirements in homeotherms (Lillywhite and Navas 2006). Taking advantage of well-developed population genetic theory, I integrate ecology, molecular genetics and morphology to study how the interplay between natural selection and gene flow affects the spatial structure across the entire range of a bird endemic to the south-western arid-zone of Africa (Figure 1a).

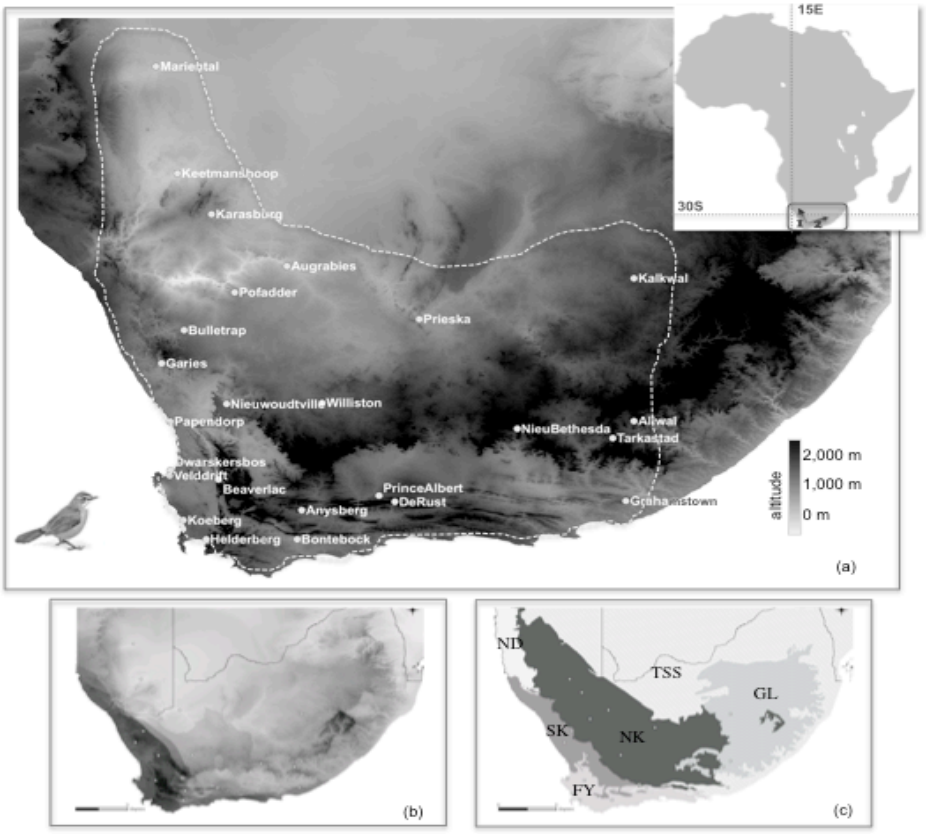


**Figure 1.** (a) Geographical range and the 24 sampling localities for the Karoo Scrub-Robin. Inset showing location in Africa and the two conceptual axes used to characterize climate across the species range. (b) Variation in mean annual temperature over southern Africa represented in shades of grey, lighter shade = larger temperature range. (c) Biomes in southern Africa: FY – Fynbos, GL – Grassland, ND – Namib Desert, NK - Nama Karoo, SK – Succulent Karoo, TSS – Tree and shrub savanna.

The Karoo Scrub-Robin, *Cercotrichas coryphaeus*, is a sedentary, medium sized, ground-feeding insectivorous bird (18 - 23g; Oatley and Arnott 1998) that exhibits female-biased dispersal (chapter three). Its geographical range encompasses an area with rainfall and temperature gradients, which result from two main processes: the cold Benguela Current running northward along the western coast of southern Africa and the *rain shadow effect* created by the Drakensberg Mountains in the eastern part of the sub-continent (Werger 1986). Thus, two conceptual axes can be used to characterize the overall climate within the range of the Karoo Scrub-Robin (Figure 1a and b): 1) a latitudinal gradient where precipitation decreases northward, from the Cape Province into the southern Namib Desert; and 2) a longitudinal trend in seasonality extending from a winter rainfall regime along the western coast to a summer rainfall regime in the interior, with a narrow intermediate area with a year-round rainfall regime (Desmet and Cowling 1998, Chase and Meadows 2007). This peculiar climatic setting was established during the Holocene (Chase and Meadows 2007) and has a dramatic effect on vegetation physiognomy (Figure 1c), primary productivity and drinking water availability across the landscape (Dean and Milton 2004). This provides an opportunity for directionality of natural selection to differ across the range and hence to promote adaptation to local conditions.

Should geographically varying selective pressures in the southern African arid-zone suffice to overcome the influx of non-adaptive traits/alleles, I predict (Figure 2): 1) geographical variation in environmental variables to create a non-random distribution of adaptive traits as opposed to neutral traits, which are expected to vary independently of environmental variables, 2) strongly differentiated populations coincident with spatial climatic structure within the range as opposed to a pattern of isolation by distance, which would only depend on effective dispersal distance (gene flow) and effective population size (genetic drift), and 3) contrasting levels of gene flow for adaptive and neutral loci.

Recent studies have shown that a combination of low basal metabolic rate (BMR) and evaporative water loss (EWL) is favoured in passerine birds from dry and hot environments (Williams and Tieleman 2005, Tieleman et al. 2003). Therefore, I tested the above predictions by quantifying two types of traits likely to mediate the adaptive response: morphology and physiological genetics. First, birds can reduce their EWL by decreasing the amount of water lost by evaporation through the skin (Williams and Tieleman 2005). One possible mechanism to reduce water loss though the skin is to minimize the surface area exposed. Thus, I examined evidence of a trend toward increased body size in response to greater aridity as a means to minimise cutaneous water loss. Second, mitochondria play a central role in energy and heat production (Hochachka and Somero 2002). Because of the uni-parental mode of inheritance of the mitochondrial genome, and the high mutation rate favouring the rapid expression of new advantageous alleles, several authors have proposed that mitochondria are likely to mediate an adaptive response to

extreme environmental challenges and dietary conditions (e.g. Mishmar et al. 2003, Ballard and Whitlock 2004, Gershoni et al. 2009, Ballard and Melvin 2010). Furthermore, Tieleman et al. (2009) recently demonstrated the functionality of the mitochondrial genome in the regulation of energetic metabolism in birds, particularly BMR. I examined the genetic variability underlying a key enzyme from the oxidative phosphorylation pathway (OXPHOS) with the rationale that climatic-based selection might act on amino acid substitutions in the mtDNA and thus impact metabolism.

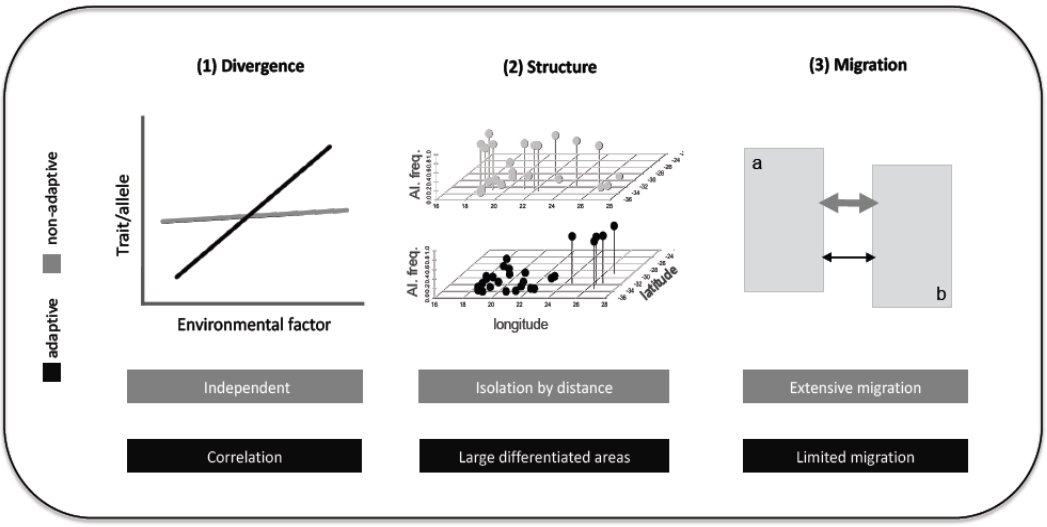


**Figure 2.** Schematic description of predicted patterns of: (1) divergence, (2) population structure and (3) migration rates for adaptive and non-adaptive traits/loci under a *migration-selection model*. 'a' and 'b' represent environmentally distinct patches. The width of the arrow in (3) is proportional to migration rates.



## METHODS

### Field Procedures

Four field expeditions, across different years, were conducted to capture birds from 24 localities in southern Africa using mist-nets and traps (Figure 1). This sampling scheme encompassed the entire distribution of the species. Captured birds were permanently banded with a uniquely numbered aluminum ring, weighed, measured and bled by puncture of the sub-brachial vein. Alternatively, a representative voucher specimen was collected. Tarsus-length was measured with a digital caliper to an accuracy of 0.1 mm and mass was measured with a 50 g Pesola spring scale to an accuracy of 0.5 g. Juvenile birds were distinguished on the basis of plumage features (Oatley and Arnott 1988).

**Environmental and morphological variation**

Eighteen bioclimatic variables at a resolution of 30 arc-seconds (1 $Km^2$ grid) were obtained from the WorldClim database (Hijmans *et al.* 2005). The BIOCLIM algorithm (Nix 1986) as implemented in DIVA-GIS v5.4 (Hijmans *et al.* 2004) was used to extract climatic data for each of my geo-referenced localities. Bioclimatic variables were used to characterize the climatic envelope by implementing a PCA. Because principal components (PC) are dependent on the units of the original variables, before compressing the data into new uncorrelated variables, the original variables were standardized by the maximum value. The PCs were extracted from a covariance matrix and the scores used to document patterns of environmental variation across the geographical range of the species, as well as to estimate environmental distances between populations along each resulting climatic vector.

Juvenile birds were excluded from the morphological analyses. Body mass (g) and tarsus-length (mm) of adult birds (n = 270) were used to quantify phenotypic variation across the entire range. In order to test the effect of climate (two factors: PC1 and PC2) on morphology (response variables: body mass and tarsus-length) I implemented a MANOVA on log10-transformed data that was previously tested for normality. I chose to use body mass and tarsus-length as a proxy for body-size, variables which are likely to reflect climatic-based selective pressures for reduced water loss. When the overall MANOVA was significant (*Wilks'* test) I next examined the univariate *F*-test for each variable to understand its respective effect. When the ANOVA was significant, a post-hoc Tukey HSD test was used to discriminate which populations occupied significantly different regions of the morphospace. All statistical analyses were performed using R (R Development Core Team 2010).

**Genetic data**

Genomic DNA was extracted from blood/tissue samples using DNeasy kits (Qiagen, Valencia, CA). The sex of all birds was determined using a PCR-based assay (Fridolfsson and Ellegren 1999). Of the 16 polypeptides that comprise the fifth complex of the OXPHOS pathway, the ATP synthase, I chose to examine the only two polypeptides that are mtDNA-encoded: ATPase subunits 6 and 8. Subunit 6 and 8 play an essential role in the translocation of protons across the inner membrane and in the assembly of the complex, respectively (Pedersen et al. 2000).

**1. DNA sequences**

Two mitochondrial protein-coding genes (ATP6 and ATP8), three nuclear autosomal loci (Gapdh-intron11, *Gallus gallus* chromosome 1; βFib-intron5, *Gallus gallus* chromosome 4; *15691 Gallus gallus* chromosome 5) and one Z-linked intron (BRM-intron15) were amplified by PCR using the following

primers: ATP6-8 (A8L8929 and A6H9929, E. Bermimgham Lab), Gapdh (Friesen et al. 1997), Fib5 (Fuchs et al. 2004, Kimball et al. 2009), *15691* (Backström et al. 2008) and BRM (Goodwin 1997). All reactions were performed in a total volume of 10 µl with 10-20ng of genomic DNA, 0.5U of Taq Polymerase (Roche), GeneAmp 10x PCR Gold Buffer, 2.0-2.5 mM $MgCl_2$, 0.3 mM of each dNTP and primer concentrations of 0.15 mM. The thermocycling profile consisted of: an initial denaturizing step at 95°C for 3 min followed by 35 cycles at 95°C for 30 seconds, and locus-specific annealing temperature (55°C - 60°C) for 30 seconds, and 72°C for 30 seconds, with a final extension step at 72°C for 7 min. Cycle-sequencing reactions were performed in both forward and reverse directions using the ABI BigDye Terminator Kit v3.1 (Applied Biosystems, Foster City, CA, USA). Cycle sequencing reaction products were purified using Sephadex columns and then analyzed on an ABI 3730 automated sequencer (Applied Biosystems).

Sequences were edited and aligned using CodonCode Aligner v3.5.2 (CodonCode Corporation 2009) and Geneious Pro v5.0 (Biomatters Ltd 2010). All mtDNA sequences were unambiguously translated into their amino acid sequence (no stop codons) and no ambiguous base calling was observed. Length polymorphisms were found in all nuclear introns. Because the evolution of *indels* is difficult to model and hence incorporate into coalescent-based analysis of sequence variation, I chose to remove them from the alignments prior to further analysis, after being resolved using an algorithm implemented in CodonCode Aligner. The exception was locus *15691,* for which I decided to truncate the sequence data at the first *indel* in the reverse direction given the presence of multiple length polymorphisms.

When the nuclear loci sequences were heterozygous at two or more nucleotide positions I use a two-step approach to determine the gametic phase of each sequence. I first analyzed the sequences of autosomal and Z-linked loci of each individual using PHASE 2.1 (Stephens et al. 2001, Stephens and Donnelly 2003). The PHASE algorithm was run twice for each data set ($10^4$ main iterations with a $10^3$ burn-in, -x100 option) from different starting points. The gametic phase of most of the genotypes was resolved with posterior probabilities greater than 0.80 for three introns (Gapdh 95%, *15691* 100% and BRM 97% of the sequences); however, for Fib5 several individuals had allele pair probabilities < 0.80. Thus, I designed allele-specific primers (3'-end matched one of the polymorphic sites in the template sequence) to amplify both alleles independently (see Bottema et al. 1993). I then employed PHASE once more, but with the additional –k option in effect, i.e. the known gametic phases for several individuals was *a priori* specified, thus improving the efficiency of the algorithm: 96% (n= 96) of the genotypes were resolved with posterior probability > 0.94.

In order to test for possible intralocus recombination events I used the four-gamete test (Hudson and Klapan 1985) as implemented in DnaSP v5.10.1 (Librado and Rozas 2009). If recombination was detected

at a given locus, all subsequent analyses were performed with the largest independently segregating block. I used DnaSP to estimate haplotype diversity (Hd, Nei 1987), nucleotide diversity (π, Nei 1987), the number of segregating sites (*S*, Waterson 1975), theta from *S* (θ, Nei 1987) and $K_{ST}$ values (haplotype based statistics similar to $F_{ST}$; Hudson et al. 1992).

**2. Microsatellites**

A total of 285 birds were genotyped for eleven autosomal nuclear microsatellites specifically developed for the Karoo Scrub-Robin and multiplexed in four reactions as described in chapter two. Allele sizes of fluorescently labelled fragments were determined using the size standard LIZ-500 on an ABI 3730 DNA Analyser followed by analysis with GeneMapper version 4.0 (Applied Biosystems). Linkage disequilibrium and deviations from Hardy–Weinberg (HW) expectations, for each locus and each locality, were tested using GENEPOP v3.4a (Raymond and Rousset 1995). Allelic diversity, Ho, He and $F_{ST}$ values were computed in GenAlEx v6.1 (Peakall and Smouse 2006).

**Population genetic structure**

Two distinct approaches were used to test for population structure using the genetic data. First, I used STRUCTURE v2.3 (Pritchard et al. 2000, Hubisz et al. 2009) to identify groups of randomly mating individuals with different microsatellite allele frequencies, by minimizing deviations from HW expectations and linkage disequilibrium. Analyses were implemented using the *admixture* model with *correlated allele frequencies*, including information about the sampling location as a prior (note that the LOCPRIOR option is different from implementing the USEPOPINFO flag; see Hubisz et al. 2009). I ran five pseudo-replicates with $10^6$ MCMC iterations following a burn-in of $10^5$. Second, to test the hypothesis that greater differentiation occurs among groups occupying distinct locals in environmental space than among groups within similar environmental niches, I used an analysis of molecular variance (AMOVA, Excoffier et al. 1992). Sequence variation (mtDNA and nuclear DNA) and allele frequency (microsatellites) were partitioned between ecologically defined groups and among populations within the previous groupings. The AMOVA was implemented in ARLEQUIN v3.1 (Excoffier et al. 2005) and the statistical significance of the *F* statistics was assessed using $1x10^5$ permutations.

**Signatures of selection in mitochondrial genes**

I contend (at least for the present) that evidence for selection within species may best be inferred by implementing polymorphism-based methods. As such, sequence polymorphism data was used to test for deviations from the neutral expectation of the number of mutations (i.e. allele frequency spectrum) by

estimating Tajima's *D* (Tajima 1989) and Fu's *Fs* (Fu 1997) statistics. The null distribution and statistical significance were determined by simulating the coalescence process $1x10^5$ times in DnaSP.

**Correlative approach: comparing environmental, mitochondrial genetics and morphological variation**

Mantel matrix correspondence analysis (Mantel 1967, Smouse et al. 1986) was used to test the hypothesis that populations that occupy environmentally similar sites tend to be similar in morphology and/or in physiological genetic variation implying local adaptation. Correlation tests were implemented as described below, using different matrices.

*The effect of climate in mtDNA variation*

I estimated the correlation between an environmental variation matrix produced by estimating the linear distances along the PC1 and PC2 axes across the geographical range and a matrix representing the population molecular differentiation at mtDNA (ATPase6 and APTase8) as measured by linearized $F_{ST}$ ($F_{ST}$ /1- $F_{ST}$; Rousset 1997).

*The effect of climate on morphology*

Morphological variation among populations was computed as the Euclidean distances between the residuals from the linear model fitted to tarsus-length and mass – body size. The association was tested with the climatic matrices described above.

Based on the current distribution of the species and the predicted historical ranges (Figure 3A and Annexe A) a historical scenario of range expansion following the Holocene Maximum (6 000 years before present) was determined to be plausible. Because population densities and migration tend to decrease towards the margin of the distribution (abundant core theory; Eckert et al. 2008) the dynamics of colonization of newly available habitat from marginal populations is expected to have a strong stochastic component and thus increase differentiation. Using the relative position of each sampled locality within the species' present distribution I expect that range expansion to have caused greater genetic differentiation among newly colonized patches than among localities sampled in the core of the species distribution.

*The effect of range expansion in the mtDNA variation*

The expected differentiation based on the last dramatic change of the species range (Holocene

Maximum) was coded in a trichotomous key (Figure 3B): '0' for differentiation between populations within the stable core, '1' representing divergence between the stable core and marginal population pairs and, '2' coding the expected greater differentiation between populations in the newly colonized area. The correspondence between the observed differentiation at mitochondrial genes and this hypothesis matrix was then tested.

*The effect of range expansion on morphology*

Correspondence between morphological variation (as described above) and the hypothesised differentiation matrix (Figure 3B) was tested.

All matrices were composed of the same 24 populations; nevertheless, the number of individuals used to estimate the mean population value varied according to the nature of the variable. The statistical significance of the correlation was assessed using permutations as implemented in the VEGAN package within R (Oksanen et al. 2010). The distribution of the coefficient values *r,* under the null hypothesis of '*no association',* was obtained by shuffling the matrix values $1x10^5$ times. The association was first tested between the two main matrices. I then used a partial matrix correspondence test to control for the effect of spatial proximity while still testing the possible correlation between the first two matrices. Geographical proximity between sites was measured as the Euclidean distances between log-transformed latitude/longitude coordinates, previously converted using Lambert's azimuthal equal-area projection (a projection that better represents area on a sphere, particularly important given the broad distribution of the Karoo Scrub-Robin).



**Coalescent estimates of gene flow**

To test the role of environmental selective pressures in constraining gene flow, I used the coalescent-based *isolation with migration* model (Wakeley and Hey 1998, Nielsen and Wakeley 2001, Hey and Nielsen 2007) to estimate migration rates (*m*) between ecologically defined population pairs. The model was applied to the data (only the independently segregating segments as mentioned above) as implemented in IMa (Hey and Nielsen 2007). IMa analyses were conducted to investigate the relative gene flow across marker type, under the full six-parameter model (effective population size $\theta_1$, $\theta_2$, $\theta_3$; migration $m_1$, $m_2$; time of divergence *t)*. After preliminary runs to determine the appropriate range of priors for each of the three parameters, as well as confirming proper MCMC mixing, the final runs were performed for $2 x10^7$ steps with 30 Markov-Monte-Carlo coupled chains with a geometric heating scheme (g1 = 0.3 and g2 = 0.9). The first $10^6$ steps were discarded as the burn-in. To enable comparisons of

migration rates between the three types of marker, which are characterized by different effective population sizes, I used an *inheritance scalar* to adjust the parameters in the model: (0.25 - mitochondrial, 0.75 - Z-linked, 1 - autosomal). I then implemented the 'Load-Trees' mode for the mtDNA and nDNA multilocus data set (introns and microsatellites separately) to test the full model against a model of divergence in isolation, i.e. $m_1$ and $m_2$ was restricted to zero. The significance of the test statistics (log-likelihood ratio) was assessed using a chi-squared test (Hey and Nielsen 2007). This method offers the best available option to disentangle the relative effects of lineage sorting and gene flow.

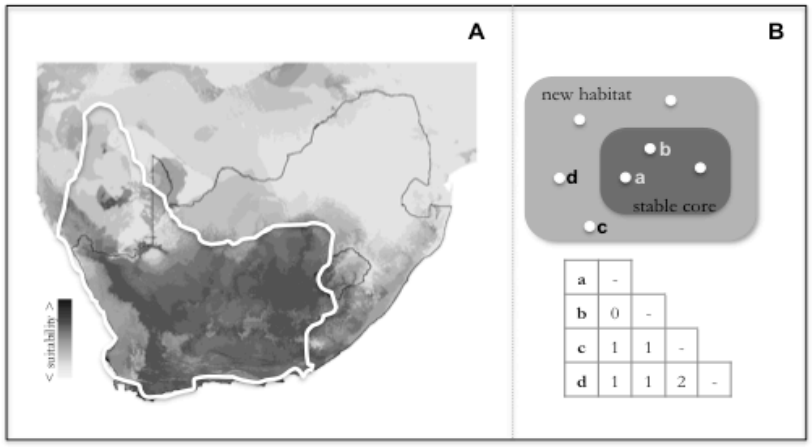


**Figure 3.** A) Predicted distribution range of the Karoo Scrub-Robin during the Holocene. Probabilities of occurrence are depicted in a continuous scale of grey: the darker the grey the higher the likelihood of occurrence. The region outlined in white shows the current distribution of the species. B) Colonization model and its predicted consequences for population differentiation. Dark grey represents the historical species range and light grey the current distribution. The values '0', '1' and '2' in the matrix represent the expected differentiation between populations within the stable core, divergence between the stable core and marginal population pairs, and differentiation between populations in the newly colonized area, respectively.



## RESULTS

### Environmental heterogeneity

The first and second principal components accounted for 78% of the variation in the data: PC1 explained 44% and PC2 34%. The variation in PC1 is explained by differences in rainfall regimes:

precipitation of the driest month (0.461) and precipitation of the driest quarter (0.452) loaded positively, whereas precipitation seasonality had a negative coefficient (-0.231). PC2 illustrates the effect of continentality: mean temperature of the coldest month (0.673) and mean temperature of the driest quarter (0.423) had the highest coefficients. A closer inspection by sampling locality revealed: i) increasing aridity as one moves northwards, with a clear distinction between the mesic Fynbos and the xeric Karoo; ii) a sharp contrast between localities along the western coast where climate is moderated by the Atlantic Ocean, and the interior and south-east localities which exhibit a greater annual range in temperature.

**Morphological divergence**

A significant overall MANOVA (*Wilk's* = 0.7214, $p < 0.001$) followed by univariate analysis indicated that environmental variables had an effect on morphology. Despite the significant effect of PC1 on mass ($F_{23,270} = 7.312$, $p < 0.001$) and tarsus-length ($F_{23,270} = 7.312$, $p < 0.001$), as well as PC2 on mass ($F_{23,270} = 4.168$, $p < 0.001$) and tarsus-length ($F_{23,270} = 4.168$, $p < 0.001$) the majority of the 24 sampled populations were morphologically indistinguishable in environmental space. The few exceptions were always eco-geographically peripheral populations (Table I). For instance birds from Grahamstown (GTN, n = 8) at the SE margin of the species range were heavier, whereas individuals from the Dwarskersbos population (DWK, n = 5) along the SW margin had the lowest body mass (see Figure 1). A similar trend was recovered for tarsus-length with birds from Koeberg (KBG, n = 14) in the SW part of the range, having longer tarsi than birds from Prieska (PRK, n = 14) in the north central part of the species range (see Figure 1).



**Table I.** Morphological variation across the aridity gradient as summarized using PCA: groups statistically distinguishable with post hoc Tukey HSD test.

| | Factor | df | *F* | *P* | Groups (n) |
|---|---|---|---|---|---|
| **Mass** | | | | | |
| | PC1 | 23, 270 | 7.3123 | <0.0001 | GTN (8); DWK (5); others |
| | PC2 | 23, 270 | 7.3123 | <0.0001 | GTN (8); DWK (5); others |
| **Tarsus-length** | | | | | |
| | PC1 | 23, 270 | 4.1682 | <0.0001 | KBG (14); PRK (14); others |
| | PC2 | 23, 270 | 4.1682 | <0.0001 | KBG (14); PRK (14); others |

df, *F* and *p* from ANOVA; Three letter acronym represent sampling sites, see main text for more information; n: sample size

**Patterns of genetic variation and selection**

Polymorphism, divergence estimates, frequency spectrum test of neutrality and recombination estimates for mitochondrial DNA and nuclear DNA sequence data are summarized in Table II. Recombination was detected at Gapdh and Fib5 ($R > 0$). thus all subsequent analyses were performed with the largest independently segregating block ($R = 0$). I observed substantial levels of heterogeneity in sequence variation among loci: values of π varied from 0.002 - 0.02. Genetic differentiation at the mtDNA genes among localities, $K_{ST}$, was very high (average 89%, $p < 0.001$). A small, but significant differentiation was also recovered for the nuclear introns: $F_{ST}$ values varied from 2% to 14%.

Within sites, all microsatellites were polymorphic, in HW and linkage equilibrium. The number of alleles ranged from three to 13; Ho and He varied from 0.581 to 0.732 and 0.658 to 0.738, respectively. Genetic differentiation ranged from 6% to 11% in the eleven microsatellites (Table III).



**Table II.** Summary of the genetic variation in sequence data: mtDNA, autosomal loci and the Z-linked locus.

| Genome | | loci (n) | sites (bp) | *R* | S | H | Hd | π | θ# | *D* | Fu's Fs | *K*st |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| mtDNA | | ATP8 (164) | 168 | | 8 | 9 | 0.627 | 0.009 | 0.008 | 0.1205 | -0.728 | |
| | | ATP6 (164) | 578 | | 61 | 39 | 0.856 | 0.025 | 0.019 | 0.9212 | -0.732 | |
| | | All | 754 | | 70 | 46 | 0.870 | 0.021 | 0.017 | 0.8497 ($p = 0.865$) | -2.361 ($p = 0.370$) | 0.89** |
| nDNA | Autosomal | Gapdh (214) | 194 | 0 | 10 | 10 | 0.152 | 0.001 | 0.009 | -2.0854 ($p < 0.001$) | -15.7830 ($p = 0.001$) | |
| | | | 262 | 1 | 14 | 15 | 0.692 | 0.004 | 0.090 | -1.5126 ($p = 0.031$) | -9.205 ($p = 0.015$) | 0.05* |
| | | Fib5 (184) | 307 | 0 | 15 | 15 | 0.613 | 0.003 | 0.008 | -1.6842 ($p = 0.01$) | -9.9294 ($p = 0.005$) | |
| | | | 465 | 2 | 30 | 35 | 0.899 | 0.087 | 0.011 | -0.6040 ($p = 0.329$) | -13.6781 ($p = 0.005$) | 0.05* |
| | | 15691 (180) | 135 | 0 | 8 | 8 | 0.235 | 0.002 | 0.010 | -1.76876 ($p = 0.001$) | -7.272 ($p = 0.019$) | 0.02 |
| | Z-linked | BRM (193) | 257 | 0 | 11 | 13 | 0.511 | 0.002 | 0.008 | -1.7243 ($p = 0.003$) | -10.352 ($p < 0.001$) | 0.14** |

n (number of alleles), bp (base pairs), R (number recombination events), S (segregating sites), H (number of haplotypes), haplotype (Hd) and nucleotide (π) diversity, θ# theta per site, *D* (Tajima's D) and genetic divergence (Kst; Hudson et al. 1992). * $p < 0.01$, ** $p < 0.001$.

**Table III**. Genetic variation at 11 autosomal microsatellites genotyped for 285 individuals from 24 populations.

| Locus | Na | He (SE) | Ho (SE) | $F_{ST}$ |
|---|---|---|---|---|
| Eco2 | 3 | 0.476 (0.026) | 0.524 (0.039) | 0.091 |
| Eco12 | 9 | 0.848 (0.007) | 0.810 (0.029) | 0.073 |
| Eco13 | 9 | 0.836 (0.007) | 0.870 (0.018) | 0.067 |
| Eco14 | 7 | 0.722 (0.026) | 0.686 (0.036) | 0.091 |
| Eco16 | 5 | 0.641 (0.021) | 0.603 (0.039) | 0.102 |
| Eco17 | 5 | 0.654 (0.017) | 0.566 (0.040) | 0.084 |
| Eco30 | 5 | 0.699 (0.013) | 0.682 (0.035) | 0.106 |
| Eco32 | 7 | 0.773 (0.016) | 0.806 (0.026) | 0.095 |
| Eco42 | 3 | 0.501 (0.021) | 0.481 (0.035) | 0.111 |
| Eco56 | 7 | 0.805 (0.008) | 0.849 (0.018) | 0.076 |
| Eco66 | 4 | 0.628 (0.014) | 0.663 (0.032) | 0.066 |
| All | 6 | 0.689 (0.009) | 0.685 (0.012) | 0.088 |

Na: number of alleles; expected (He) and observed (Ho) heterozigosity; Standard error (SE)



The distribution of allele frequencies as measured by Tajima's *D* and Fu's Fs statistics varied among loci, ranging from -2.08 to 0.85 and -15.78 to -2.361, respectively. All four nuclear loci showed a significant skew in the allelic frequency spectrum, i.e. a high frequency of rare polymorphisms; *D* and Fu's Fs were significantly different from neutral expectations (Table II). This nucleus-wide skew towards rare alleles suggests that the Karoo Scrub-Robin has probably undergone a demographic expansion during the species' history. In contrast, Tajima's *D* for the mtDNA loci was positive and close to '1' (0.850), but not significant. A closer inspection of the allele frequency spectrum revealed an excess of both low frequency and high frequency polymorphism, disrupted by some intermediate frequency alleles. This is the pattern expected under positive selection: abundance of low-frequency polymorphisms and high-frequency variants (Fu 1997, Nielsen 2005).

The most striking result is the segregation of alternate amino acids within each of the two climatically well-defined areas (as depicted in Figure 4): two closely positioned replacement polymorphisms in the ATPase6 gene segregate in different climatic zones, whereas not a single amino acid substitution was recovered in the ATPase8 gene. The amino acid replacement I51V (Isoleucine for Valine at amino acid position 51) separated the western mesic sites from the interior arid populations, hereafter referred to as Western and Central groups. Isoleucine (I) was almost exclusively present in populations that experience relatively small variations in mean annual temperature due to the moderating

influence of the nearby Atlantic Ocean and receive most of their rainfall during the cold season. In contrast, Valine (V) at position 51 was recovered for all individuals sampled from the portion of the species range characterized by a summer rainfall regime. One site was exceptional: Bontebok (BTB, n = 4) where Valine was dominant (100%). Although not analysed in this paper given the small sample size (n = 2), it is worth mentioning that Helderberg was the single site where I found the co-occurrence of both forms of ATPase6 molecules (50% V, 50% I; depicted in Figure 4). An additional amino acid change at H63D was recovered exclusively within the most mesic area within the winter rainfall domain. Populations in this section of the environmental space have an Aspartic Acid (D) in position 63, as opposed to the remaining populations, which have a Histidine (H): the two forms of this replacement of ATPase6 are only sympatric at Beaverlac (62% D, 38% H, n = 8; Figure 4).

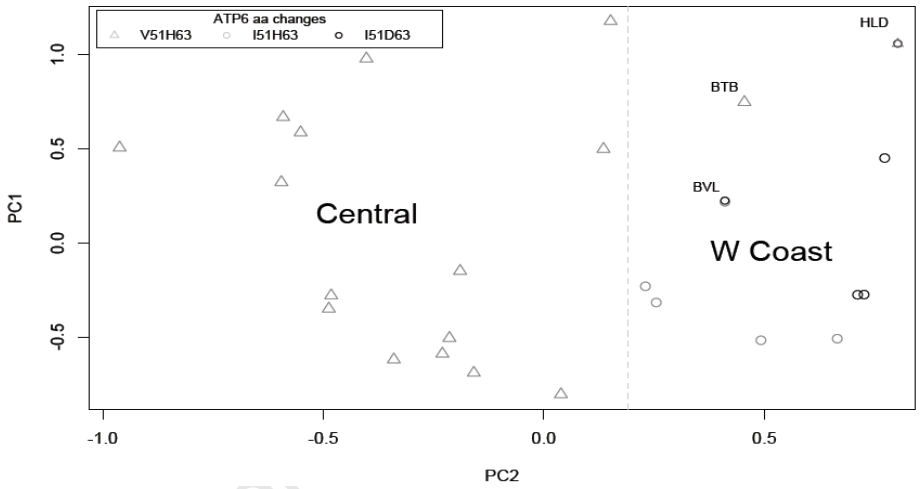


**Figure 4.** Environmental space of the two amino acid replacements, Isoleucine/Valine at amino acid 51 and Histidine/Aspartic Acid at amino acid 63, in the OXPHOS genes. Sites mention in the main text: Bontebock (BTB), Helderberg (HLD) and Beaverlac (BVL).

**Contemporary genetic structure and gene flow**

Overall, $F_{ST}$ values estimated for my microsatellite data were small (< 0.088) with the two most distant populations (aerial distance ≈ 1,200km) having a $F_{ST} = 0.056$. No signal of isolation by distance ($r = 0.103$, $p = 0.283$) was recovered. These results were further supported by the lack of genetic clustering found with the Bayesian clustering method implemented in STRUCTURE, which indicated high levels of gene flow between populations (Figure 5). The likelihood of the marginal posterior probability distribution was the highest when K = 1 (Ln P (D | K=1) = -1,1286). The low $F_{ST}$ values observed hinder the implementation of BayesAss to accurately estimate current gene flow.

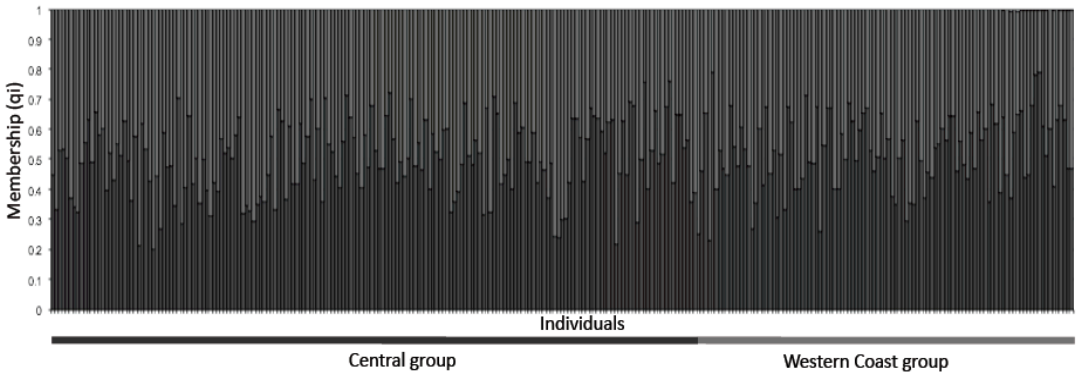


**Figure 5.** Graphical summary of STRUCTURE results based on 11 autosomal microsatellites using the maximum-likelihood estimate of K = 2. Each individual (total = 285) is represented by a vertical bar partitioned into two segments proportional to its membership to each population.



**Ecological divergence: correlation between phenotype, genotype and environment**

The AMOVA results confirmed that most of the mitochondrial variation was associated with environmental PC2-based groupings: 77% ($F_{CT}$ = 0.77, p < 0.001) between the two above mentioned groups (Western vs. Central; see Figure 4) when compared with 1.1% among localities within each group ($F_{SC}$ = 0.658, p < 0.001). It also revealed, nonetheless, that a significant but small portion of the variation was explained by differences among all localities (7.77%; $F_{ST}$ = 0.922, p < 0.001). In contrast, the intron data indicated that variance among groups was low and only significant at one locus: 0% for BRM (p = 0.954), 0.7% for Fib5 (p = 0.345), 0.1% for Gapdh (p = 0.157) and 4.6% for *15691* (p = 0.015). Most of the variation was rather significantly partitioned among all localities: BRM = 83% ($F_{ST}$ = 0.170, p < 0.001), Fib5 =95.6% ($F_{ST}$ = 0.04, p = 0.020), Gapdh = 93.6%, ($F_{ST}$ = 0.06, p = 0.010) and *15691* = 97.4% ($F_{ST}$ = 0.015, p < 0.015).

The analysis of variance of microsatellites alleles revealed a similar result, i.e. the hypothesis that population structure was associated with climate groupings was not supported (0.02%, $F_{CT}$ = 0.002, p = 0.38) and most of the variation, 90.3%, was distributed among all localities ($F_{ST}$ = 0.039, p < 0.001). Note that the AMOVA has an inherent spatial effect, which might partially affect the results. In order to remove the spatial proximity effect I used a Mantel test of matrix correspondence. A significant portion of mtDNA and morphological divergence could be explained by the spatial distribution of the environmental conditions (37% and 34%, respectively). However, as expected, spatially close sites were characterized by similar environmental conditions ($r$ = 0.607 p < 0.001). Once I removed the effect of spatial proximity,

PC2 still explained 39% (p = 0.008) of the mtDNA haplotypic distribution. In contrast, morphology was associated with geographical distance: populations far apart were phenotypically more distinct. This is consistent with the idea that the effect of migration decreases from the core towards the edge of the distribution where drift plays a greater role (Table IV).

Genetically (mtDNA genes) divergent populations were not generally morphologically divergent from each other ($r$ = -0.077, p = 0.752). There was a significant correspondence (27.8%, p = 0.011) between observed mtDNA variation and the hypothesis matrix (Figure 3) indicating that genetic variation among newly colonized populations is higher than between populations from the stable core. Thus, suggesting the possible role of drift during colonization, in creating the observed pattern in mtDNA. No correlation was, however, found between the hypothesis matrix and morphological divergence ($r$ = -0.068, p = 0.648). Overall, and after Bonferroni corrections, only mtDNA genetic variation correlated with environmental and range dynamics.

**Table IV.** Coefficients from matrix correlation tests showing the degree of congruence between environmental variation and mitochondrial divergence, as well as morphological patterns.

| | **Space** | **Environment** | | **Historical** |
|---|---|---|---|---|
| | GGD | PC1 | PC2 | Range shift |
| Mitochondrial variation | 0.100 ($p$ = 0.102) | -0.214 ($p$ = 0.998) | **0.396** (p < 0.001) | **0.278** ($p$ = 0.011) |
| | \| GGD | -0.288 ($p$ =1.000) | **0.388** ($p$ = 0.008)* | **0.278** ($p$ = 0.011)* |
| Morphology | **0.295** ($p$ = 0.002) | **0.263** ($p$ = 0.01) | **0.278** ($p$ = 0.004) | -0.097 ($p$ = 0.711) |
| | \| GGD | 0.154 ($p$ = 0.071) | **0.192** ($p$ = 0.038) | -0.068 ($p$ = 0.648) |

' | ' is used to delineate that the given variable was controlled in a partial matrix correlation test. GGD stands for geographical distance. Significant values in bold. * Significant results after Bonferroni correction for multiple tests ($\alpha$ = 0.05/3 = 0.017). Significant values in bold.

**The effect of environmental heterogeneity on gene flow**

I found opposite trends in gene flow according to the inheritance mode of the genetic marker (Figure 6). Migration rate ($m$) from Western to Central populations, for mtDNA, was 0.075 (0.015 – 0.2550, 90% high posterior density (HPD) interval; Figure 6C). In the opposite direction, the migration rate was 0.005 (90% HPD: 0.005 – 0.115, note that 0.005 was the first bin of the prior and therefore corresponds to zero; Figure 5D). Estimates were higher for the nuclear introns: from Western to Central $m$

= 17.05 (90% HPD: 0.55 – 86.25) and from Central to Western $m$ = 15.45 (90% HPD: 0.05 – 68.95). These estimates were 50 times higher than the values obtained for mtDNA genes, even after rescaling according to the effective population size (4-fold Ne). For microsatellites, $m$ estimates were even higher (> 100) than the values obtained with the intron data. The LLR tests clearly rejected a model of divergence in isolation ($m_1 = m_2 = 0$) whether analysing the four nuclear introns, the eleven microsatellites, or the mtDNA data (Table V). It is worth noting that the models that set gene flow from the inland Central to the Western group to zero ($m_2 = 0$) could not be rejected for the mtDNA genes and intron data (Table V). Together with the microsatellite data this asymmetrical gene flow indicates that the movement of alleles inland occurs at a greater frequency than the movement of alleles from inland to Western populations. These results suggest that mitochondrial divergence is most likely due to low levels of maternal gene flow (despite possible extensive migration of females) than to past population demography, which due to the more rapid coalescence of mtDNA may have resulted in faster sorting of alleles.

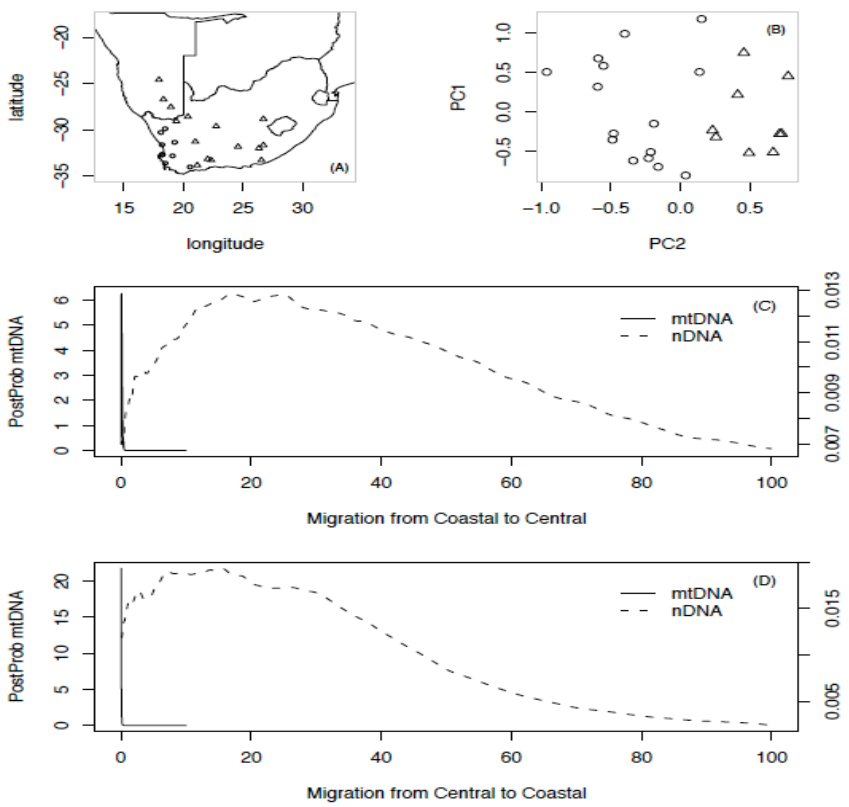


**Figure 6.** Coalescent estimates of migration for the two environmentally defined groups: geographical (A) and environmental (B) space for each group. (C) Eastwards migration $m_1$ (from Western Coast to Central) and (D) Westwards migration $m_2$ (from Central to Western Coast).

**Table V.** Likelihood ratio tests for a series of nested models applied to the ecologically defined groups (see Figure 6A-B). Divergence in isolation or directional migration was tested against the full (six parameter) model.

| Migration | Model | Data | LogL | df | 2LLR |
|---|---|---|---|---|---|
| Isolation | $\theta_1 = \theta_2 = \theta_a$ ; $m_1 = m_2 = 0$ | mtDNA | -13.569 | 4 | 17.8851* |
| | | nDNA | -460.517 | 4 | 916.669* |
| | $\theta_1, \theta_2, \theta_a$ ; $m_1 = m_2 = 0$ | mtDNA | -12.503 | 2 | 15.2833* |
| | | nDNA | -460.517 | 2 | 916.669* |
| | $\theta_1 = \theta_2, \theta_a$ ; $m_1 = m_2 = 0$ | mtDNA | -13.306 | 3 | 16.5423* |
| | | nDNA | -460.517 | 3 | 916.669* |
| | $\theta_1, \theta_2 = \theta_a$ ; $m_1 = m_2 = 0$ | mtDNA | -13.693 | 3 | 17.3574* |
| | | nDNA | -460.517 | 3 | 916.669* |
| | $\theta_2, \theta_1 = \theta_a$ ; $m_1 = m_2 = 0$ | mtDNA | -13.693 | 3 | 17.3574* |
| | | nDNA | -460.517 | 3 | 916.669* |
| Directional | $\theta_1, \theta_2, \theta_a$ ; $m_1, m_2 = 0$ | mtDNA | -5.0139 | 1 | -0.0002 |
| | | nDNA | -4.042 | 1 | 3.720 |
| | $\theta_1, \theta_2, \theta_a$ ; $m_1 = 0, m_2$ | mtDNA | -12.656 | 1 | 15.2838* |
| | | nDNA | -7.439 | 1 | 10.515* |

* significant after Bonferroni correction for multiple comparisons ($\alpha_{new} = 0.0065$)



## DISCUSSION

Ecological transitions have long been recognized as important for examining the influence of differential selection in space (e.g. Levene 1953, Endler 1986). Yet, empirical contributions (e.g. Smith et al. 1997/2001, Rosenblum 2006, Milá et al. 2009) are lagging, partially due to the way neutral theory (Kimura 1983) has inadvertently shaped the mind-set of population genetic empiricists over the past three decades. In this study I have combined different lines of evidence (as illustrated in Figure 7) to demonstrate that the range of the Karoo Scrub-Robin comprises several local evolutionary domains: i) the potential for local selection to occur due to geographically varying climatic conditions, ii) sharp eco-geographical transition of mitochondrial variants, which occupy a vast geographical area, iii) non-random distribution of amino acids with differing physicochemical proprieties (solubility and polarity) in distinct eco-geographically defined areas, iv) reduced effective migration at mtDNA among ecologically divergent groups, v) extensive gene flow as quantified with neutral genetic markers and, vi) lack of morphological divergence. The genetic variation observed in the mtDNA likely result from differential metabolic rates among individuals occupying different portions of the range, i.e. local adaptation.

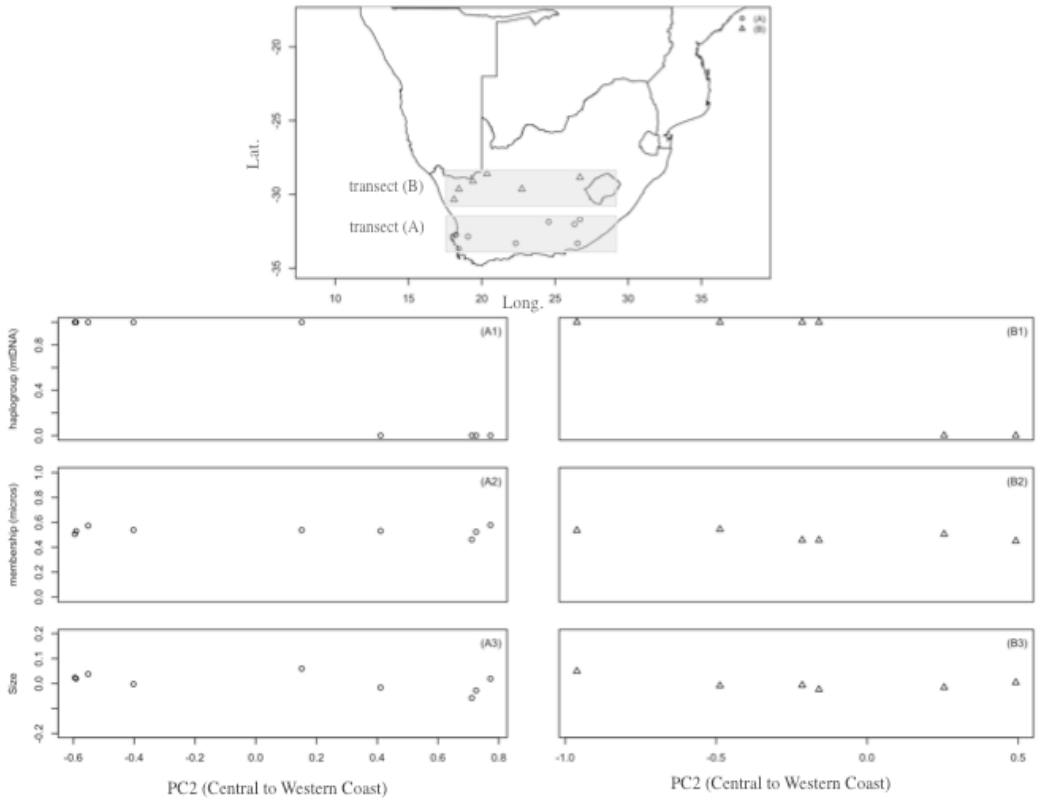


**Figure 7.** Mitochondrial haplogroup, population membership (as estimated using microsatellites in STRUCTURE) and morphological variation (size measured as regression of mass on tarsus-length) along two independent Central to Western Coast independent transects (A) and (B).



**Local adaptation: migration-selection balance**

The results presented here are consistent with the migration-selection model: strong natural selection maintaining alternative mtDNA haplotypes in populations occupying different ecological conditions, despite gene flow at neutral loci. The coalescent models of divergence with gene flow revealed a significantly better fit to the data compared to models with no gene flow (Table V). The 90% high posterior density interval of migration as estimated with the mtDNA and intron data included 'zero' from the Central to the Western group. Further, the analysis of nested models did not allow me to reject the model of asymmetrical gene flow. The possible absence of gene flow westwards might reflect differences in fitness of individuals in mismatched environments. Remarkably different estimates of gene flow between non-adaptive (nuclear) and adaptive (mtDNA) genes support my hypothesis. When scaled for the effective population size, mitochondrial gene flow was at least 50-fold smaller than expected under a

migration-drift model (e.g. from Western to Central group: nDNA $m = 4.26$, whereas mtDNA $m = 0.075$).

Although my intron data likely violate the assumptions of population demographic stability (negative Tajima's *D*) of the *isolation with migration* model, and thus the absolute estimate of non-adaptive gene flow may not be accurate, it was recently demonstrated that this model is quite robust to such violation (Strasburg and Rieseberg 2010). Furthermore, under a scenario of population growth, parameter estimates other than population size were comparable to the stable demographic scenario (Strasburg and Rieseberg 2010). Since I was primarily interested in gene flow estimates, I consider this violation very unlikely to significantly bias the results. The fact that only four eco-geographical peripheral populations were significantly different in morphological traits and hence that sites from throughout most of the species distribution were similar in body mass and tarsus-length lends further support to the homogenizing effect of recurrent gene flow in shaping the current patterns of variation among non-adaptive traits.

The hypothesis that changes in mtDNA OXPHOS genes might be a common mechanism for facilitating rapid adaptation to particular environmental conditions has received increasing support from intra-specific studies (birds: Cheviron and Brumfield 2009, humans: Balloux et al. 2010). However, identifying the actual gene(s) underlying any adaptive response is a challenging task. Changes in the physicochemical features of one or a few amino acids can affect the function/structure of a given enzyme/protein (e.g. MC1R: Theron et al. 2001, Violet-sensitive Opsin: Shi and Yokoyama 2003; *Agouti*: Linnen et al. 2009) and thus have dramatic implications in the physiological pathway in which they are involved. The population genetics approach implemented here (i.e. allele frequency spectrum) revealed the increase of mutations at high frequencies, which might be a signal of positive selection or alternatively, a possible mimicking effect of past demographic dynamics.

**Divergence: selection *versus* demography**

The observed discrepancy between the nuclear and mitochondrial genomes is consistent with the expectations of a model where selection is strong and acts on a single trait (Nosil et al. 2009a). An appealing argument with which to explain the maintenance of such strong spatial structure in the mtDNA haplotypes, in the absence of a current geographical barrier to gene flow, would be male biased dispersal. However, this is unlikely for this species given the results from direct and indirect measures of gene flow, which have demonstrated that in the Karoo Scrub-Robin, females (as in birds generally; Greenwood 1980) are the long-distance dispersers, whereas male dispersal distance follows a leptokurtic distribution (see chapter three).

The type of genetic clines observed here (see Figure 7A1, B1) can also be formed by the diffusion of alleles after a period of divergence in isolation (Endler 1977). However, this seems an unlikely explanation, for the following reasons. First, the geological history of the area offers no plausible scenario for the model of divergence-in-isolation because the most recent geological event, establishment of the present day shape of the Great Escarpment dates to the late Miocene (Partridge 1997). Second, the hypothesis of allopatric divergence due to habitat unsuitability via climatic cycling during the Plio-Pleistocene (e.g. Matthee and Flemming 2002, Herron et al. 2005, Tolley et al. 2006, Outlaw et al. 2007, Swart et al. 2009) was not supported by the historical geographical range models, which only predicted a single refugium.

Such a sharp change in allele frequencies can also be a consequence of a range expansion, a phenomenon known as 'surfing' (Excoffier and Ray 2008). The comparison between the predicted historical and current ranges suggests a range expansion westward, after the Holocene maximum. During the expansion, rare mtDNA variants 'surfing' on the front of the colonization 'wave' might have increased in frequency due to strong genetic drift. The genetic signature observed in the nuclear genome (i.e. significant negative Tajima's D and Fu's Fs) as well as the location of morphologically different populations at the trailing (GTN and PRK) and leading (KBG and DWK) edges of the moving range, gives further support to the hypothesis of a range expansion out of a single refugium. Nonetheless, secondary contact or 'surfing' of alleles during expansion fails to explain the maintenance of the spatial segregation of mitochondrial haplotypes despite extensive gene flow (introns and microsatellites). In the absence of any spatial difference in the survival of genotypes or asymmetrical competition for some kind of resource (Goldberg and Lande 2007), theory predicts that two groups will merge into each other's range as a function of time since contact *t,* scaled by dispersal distances per generation $\sigma$ (neutral diffusion model; Endler 1977). Therefore, the spatial difference in gene frequencies will disappear as a consequence of the swamping effect of gene flow. In this particular system where females are the long-distance disperser and the mean velocity of the spread of maternal alleles is quite high ($\sigma$ females = 892m/gen; chapter three) the merger between the previously diverged populations should be well underway. Given all of the above, different evolutionary forces may be responsible for creating and maintaining the observed variation: mutation creating new haplotypes, drift perhaps helping to increase their frequency in the population, and selection 'trapping' the haplotypes that are locally adapted.

**Extreme environments**

The southern African arid- and semi-arid biomes have a unique combination of environmental parameters, particularly highly seasonal rainfall, which affects resource availability, such as water and

food. Nevertheless, and despite their high mass-specific metabolism and high body temperature, some passerine birds are permanent residents in such harsh environments. Empirical data has demonstrated that selection favours reduction of the amount of daily energy required to sustain vital metabolic pathways in birds living in such areas (Tieleman et al. 2003) and also that the mitochondrial genome is partially involved in the control of the metabolic rate in birds (Tieleman et al. 2009). In this context, the results suggest that populations might have adapted physiologically, with minimal morphological change. Evaporative water loss in the Karoo Scrub-Robin can be reduced by mechanisms other than surface area reduction. For instance, by reducing the basal metabolic rate birds are also reducing water loss (Williams and Tieleman 2005). The Western group is distributed over an area where food is available almost all-year round with its abundance peaking during winter. In contrast, in the central area, food and water are scarce during the winter (when temperatures can be below freezing). This poses an extra challenge to the interior birds. Any mitochondrial variant that would allow individuals to meet the daily energetic requirements to regulate body temperature and reduce water loss would be advantageous in such extremely cold and less productive areas. Because the mitochondrial genome is transmitted as a unique linkage block, the amino acid replacements observed in the F0 domain of the ATP synthase can have an adaptive function or alternatively be hitchhiking with variation in other OXPHOS gene(s) upon which selective pressures might be acting.

Together with physiology, behavioural plasticity in seasonal and daily selection of microhabitats that provide protection from the heat and cold might also play an important role in allowing birds to survive the drastic temperature amplitude and potential starvation in these harsh arid environments. Further investigation that would allow putting the current findings in the context of the causal effects on individual fitness seems promising, especially if we are to understand the genetic, physiological and behavioural basis of adaptation to xeric environments. Understanding adaptation to aridity has become of critical importance as desertification and drought are expected to increase and hence affect the distribution of many species worldwide, and in particularly the terrestrial animal species of western and central southern Africa (Erasmus et al. 2002).

## ANNEXE A

### Species Distribution

To assess the extent to which past climatic oscillations are likely to have influenced the Karoo Scrub-Robin geographic range I modeled its paleodistribution at two-time periods. The hypothesized historical distributions under the past climates at 6 000 and 21 000 years before present were generated using the maximum entropy algorithm implemented in MaxEnt (Phillips et al. 2006) with bioclimatic variables produced by Hijmans et al. (2005) from the Paleoclimatic Modelling Intercomparison Project ECHAM3 atmospheric general circulation model and downscaled at the 30 s (1 $km^2$). I used ten independent bioclimatic variables (Pearson's correlation coefficient $r < 0.70$) extracted from 181 point-localities (records from museum specimens compiled by Richard Dean and my field expeditions). MaxEnt was first trained with 75% of the data and the remaining 25% used to test each model. Model performance was assessed with commonly used 'area under the receiver operating curve' – AUC measure (Elith et al. 2006). For the current model AUC was 0.892. The models predict the Karoo Scrub-Robin to have experienced range changes due to climatic oscillations, with a likely absence from the western coast and southern Namibia during the late Holocene. Yet, at this particular time the species is predicted to have had a continuous distribution that very much resembles the present day distribution.

CHAPTER FIVE

# Dynamics of a range expansion: placing adaptation in the context of the colonization of a new habitat



## INTRODUCTION

Biogeography has a long history of considering climate as the primary factor shaping the distributional ranges of terrestrial species (Grinnell 1914, 1917, Gaston 2003, Sexton et al. 2009). Under this paradigm, species are expected to adjust or shift their ranges in order to track climatic conditions (e.g. Moritz et al. 2008, Tingley et al.2009). Further, multiple studies have demonstrated that past climatic oscillations affected species distributions causing them to retract and expand (e.g. Lessa et al. 2003, Hewitt 2004). However, the implicit assumption of this paradigm, namely that species ecological requirements are static through time, influences the interpretation of species ranges and their limits.

When new habitat becomes available, the establishment of new populations in the novel environment depends not only on the ability of dispersing individuals to reach those patches (e.g. van Bocxlaer et al. 2010, Duckworth 2008), but also on their ability to adapt to new conditions (e.g. Parmesan et al. 1999, Rosenblum 2006, Myles et al. 2007), either from existing variation or through the evolution of a new trait(s). An increase in fitness in the new habitat accelerates colonization and facilitates range expansion. The demographic growth and selective pressures associated with the expansion into novel habitat is expected to leave a signature in genome-wide patterns of variability in contemporary populations. Population genetics theory posits that while newly advantageous variants can increase in frequency and erase previous diversity (Maynard-Smith and Haigh 1974, Storz et al. 2004, Pool and Nielsen 2007), polymorphisms that do not influence fitness are affected only by changes in the demographic trajectory of the species/population (Kimura 1956, Tajima 1989, Hudson 1990).

Compelling examples have shown that different processes can cause heterogeneous genomic divergence (e.g. Crispo et al. 2006, Nosil et al. 2008, 2009a). Natural selection results in the divergence of those portions of the genome that mediate local adaptation, whereas gene flow tends to homogenize other portions. This observation is particularly relevant in understanding the evolution of the ranges of species, because peripheral populations tend to behave as sinks, unless natural selection results in local adaptation (Kirkpatrick and Barton 1997) and thus they became source populations. Adaptation to conditions at the margin of the range can promote a range and niche expansion (Kawecki 2008).

Since the late Miocene, the southern Africa subcontinent has experienced a general trend toward increasing aridity and seasonality, with an associated shift from closed subtropical woodland to sparse and shrubby vegetation (Tyson 1986, Cowling et al. 1999, Chase and Meadows 2007). The regionally endemic Karoo Scrub-Robin, *Cercotrichas coryphaeus,* is a bird species whose current range spans the arid and semi-arid zones of the subcontinent, characterized by a west to east precipitation seasonality and aridity

gradient. Plumage colour differences have been used as diagnostic taxonomic traits to define two subspecies: *C. c. cinerea* and *C. c. coryphaeus* (Collar 2005, Oatley 2005), which have a contiguous distribution (Figure 1). *Cercotrichas c. cinerea*, the form with greyish upper parts, is found in the winter rainfall regime area where the habitat is characterized by semi-arid shrubs and dense thickets. *Cercotrichas c. coryphaeus,* the brownish colour morph, occurs in the central arid area of the subcontinent, which receives most of the rain during the summer and where sparse and low shrubs dominate the vegetation. There is a steep cline in *Cercotrichas coryphaeus* mitochondrial genome coincident with the climatic gradient, particularly the west to east shift in precipitation seasonality (from winter to summer rain fall regimes), despite extensive neutral gene flow across the whole range of the species, as indicated by microsatellite data analysed in chapter four. This suggests a role for the mitochondrial genome in allowing populations to adapt to local conditions. Interestingly, there is a concordance between the ranges of the two subspecies, the environmental gradient and the mtDNA cline (Figure 1), making this a compelling system with which to explore the nexus between ecology and evolution: i.e. to understand how factors affecting population demographic dynamics determine the evolutionary history of a species.



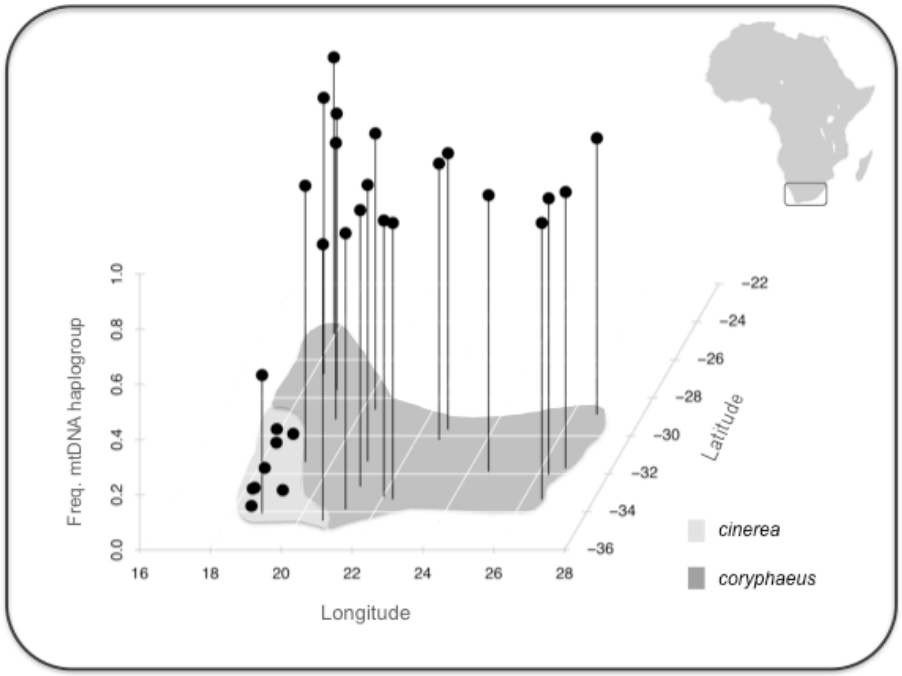


**Figure 1**. Current geographical distribution of the two subspecies *Cercotrichas coryphaeus cinerea* and *C. coryphaeus coryphaeus* underlying the spatial representation of the steep cline found in the mitochondrial DNA reported in chapter four.

I examined patterns of nuclear genetic variation in *Cercotrichas coryphaeus* and compared them to population genetics predictions of the degree and spatial structure of sequence variation expected under a demographic (Tajima 1989; Watterson 1986, Rogers and Harpending 1992) and spatial expansion (Ray et al. 2003, Excoffier 2004, Excoffier 2009) to test three competing demographic scenarios (Figure 2a). In the first scenario, the appearance of a new adaptive trait in a peripheral population facilitates the species' establishment in the new habitat. The population subsequently expands its range into the new habitat, but without exchanging migrants with the ancestral population until recent post-range expansion contact with extensive introgression (Figure 2a, model A). I predict a strong signature of demographic growth, reflected as low genetic diversity and an excess of rare alleles in the colonising population. Further, the majority of coalescent events are expected to have occurred recently. In the second scenario, colonisation of the new habitat is facilitated by the novel variant of a trait. The new population expands and continues to exchange migrants, but only for a limited time post colonization of the new habitat (Figure 2a, model B). Under this model, I expect to find an excess of rare alleles in the population occupying the newly colonised habitat, moderate estimates of genetic diversity, and migration events concentrated at a given time in the past. In the third scenario, the appearance of a newly adaptive trait in a peripheral population facilitates the invasion of the new habitat and a range expansion, but the two populations continued to (and still do) exchange migrants. Under this scenario I predict homogeneous neutral genetic variation across the entire range as a consequence of migration and gene flow, no clear spatial pattern in the location of new and old alleles, and migration events continuing through time (Figure 2a, model C).

In all scenarios, migration is expected to have a homogenizing effect across the whole genome, with the exception of at the adaptive trait. Regardless of the amount and timing of neutral gene flow, if the expansion was facilitated by the increased fitness associated with the phenotype of a particular mitochondrial haplotype (see chapter four), I expect the newly colonized population to exhibit reduced variation at the adaptive trait. Given the evidence of recent gene flow presented on chapter 4, I have only considered scenarios where recent gene flow was possible.

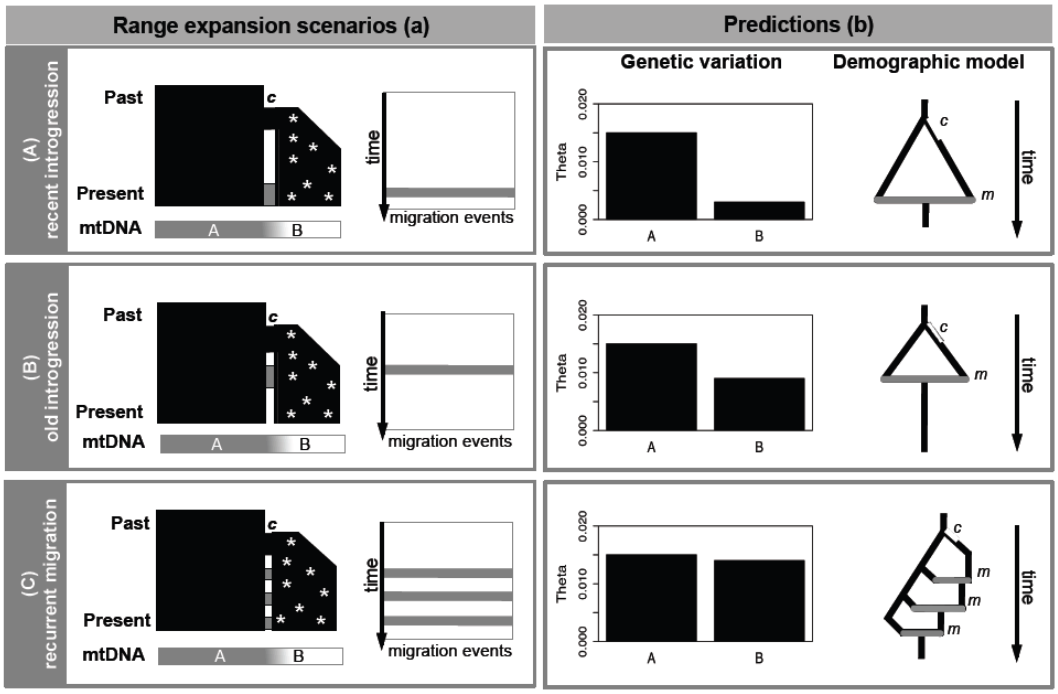


**Figure 2**. Three competing demographic scenarios of invasion of a new habitat from a peripheral population: (a) graphical representation of the underlying demographic dynamics and current pattern of variation at mtDNA; (b) predictions of genetic variation and timing of migration events. *c:* establishment of a small population in the new habitat (arid interior); ∗: new adaptive trait; *m*: migration event. Note: no



## MATERIAL AND METHODS

### Characterization of current climatic niche and range dynamics

The geographical distributions of the subspecies as defined by Oatley (2005) and Collar (2005) were used to group point locality data into subspecies *cinerea* and *coryphaeus*. I used a Principal Component Analysis to reduce the ten independent bioclimatic variables (Hijmans et al. 2005) into two orthogonal axes. The PC scores were then used as dependent variables in an ANCOVA to test whether subspecies (fixed factor) occupy different climatic niches after removing the effect of latitude and longitude (covariates). All statistical analyses were performed using R (R Development Core Team 2011).

In order to assess whether the climatic niche of the *coryphaeus* and *cinerea* have changed spatially during the Plio-Pleistocene climatic oscillations I used the maximum entropy models, as implemented in MaxEnt v. 3.3.3. (Phillips et al. 2006; Phillips and Dudik 2008). I combined presence-only data (46 points for *coryphaeus* and 17 points for *cinerea*) with climatic layers to first model the current distribution of the Karoo Scrub-Robin and the two subspecies at a resolution of 1 $km^2$ using ten independent bioclimatic variables (Pearson's correlation coefficient $r < 0.70$) produced by Hijmans et al. (2005). The accuracy of the models was assessed by partitioning the occurrence data into training (80%) and testing (20%) subsets, as well as by estimating the 'area under the receiver operating curve' – AUC. Point-locality data used in these models were obtained primarily from my field expeditions and supplemented with a few records obtained from Peter Nupen, Dawie de Swardt and Gordon Schultz. Although museum specimens with geographical reference were available, the geographic coordinates were not detailed and thus, I chose to have fewer, but more accurate points. The potential paleo-geographical range of the species and subspecies were inferred by projecting the current model into two time periods, representing the extreme climatic conditions during the Quaternary: the Last Glacial Maximum (21 000 years before present - ybp) and Holocene Optimum (6 000 ybp). Data from the Paleoclimatic Modelling Intercomparison Project ECHAM3, downscaled at the 30 s (1 $km^2$) resolution were used to the reconstruct of the past climate surfaces.



**Genetic data surveyed**

*Nuclear introns*

To test among the three demographic scenarios outlines (Figure 2a) I used sequence data from six nuclear introns, an extension of the dataset of four introns I used in chapter four. Sequences for Gapdh-intron11 (Friesen et al. 1997, chromosome 1), βFib-intron5 (Fuchs et al. 2004, Kimball et al. 2009, chromosome 4), *15691* (Backström et al. 2008, chromosome 5) and BRM-intron15 (Goodwin 1997, Z chromosome) were supplemented with two nuclear loci: TGFb2-intron 5 (*Gallus gallus* chromosome 3; Burt and Paton 1991, Primmer et al. 2002) and *26438* (*Gallus gallus* chromosome 3; Backström et al. 2008). PCR amplifications were performed in a total volume of 10 µl with 10 – 25 ng of genomic DNA, GeneAmp 10× PCR Gold Buffer, 2.0 - 2.5 mM $MgCl_2$, 0.3 mM of each dNTP, primer concentrations of 0.15 µM and 0.5 U of Taq polymerase (Roche). The thermocycling profile comprised of an initial denaturizing step at 95 °C for 3 min followed by 35 cycles at 95°C for 30 seconds, and a locus-specific annealing temperature of 55°C-60°C for 30 seconds, and 72°C for 30 seconds, with a final 7 min extension step at 72 °C. PCR products were cycle sequenced in both forward and reverse directions using

the ABI BigDye Terminator Kit v3.1 (Applied Biosystems, Foster City, CA, USA) and then analyzed on an ABI 3730 automated sequencer. Sequences were edited and aligned using CodonCode Aligner v3.5.2 (CodonCode Corporation 2009) and Geneious Pro v5.0 (Biomatters Ltd. 2010). Multiple length-polymorphisms were found at the locus *26438*; I thus truncated the sequences at the first indel. The gametic phase was determined using PHASE 2.1 (Stephens et al. 2001, Stephens and Donnelly 2003), which was run twice for each locus ($10^4$ main iterations and $10^3$ burn-in, -x100 option) and the highest-probability haplotypes used for subsequent analyses (Harrigan et al. 2008, Garrick et al. 2010).

*Microsatellites and mtDNA*

A subset (100 individuals) of the dataset of eleven microsatellites reported in chapter four (developed as described in chapter two) was used to complement the demographic modelling implemented with nuclear introns, as explained under the *Demographic models section,* with the rationale that the inherent features (mutation rate) of these genetic markers would improve the characterisation of the demographic trajectory of the species. The mitochondrial DNA dataset used in chapter four was used to assess genetic variation for both the species as a whole and the subspecies.

**Intragenic recombination**

Under the infinite sites model (ISM; Kimura 1969) the occurrence of recombination between two polymorphic sites at a given locus will originate four different haplotypes: two recombinants and two parental (Hudson and Kaplan 1985). This expectation is used in the four-gamete test (Hudson and Kaplan 1985) to estimate the minimum number of recombinant events observed within each locus. However, during population growth the increasing number of mitotic events, and hence mutations, is likely to create new haplotypes that mimic recombination. Under such a demographic scenario the ISM is violated and recombination events can be overestimated. As such, besides using the four-gamete test as implemented in the program DnaSP v5.10.1 (Librado and Rozas 2009) to detect intragenic recombination events, I also used the Φw statistic (Bruen et al. 2006) as estimated with SplitsTree v. 4.10 (Huson and Bryant 2006). The latter test (Φw) measures the genealogical similarity between closely linked sites and then discerns whether it is due to recurrent mutation or recombination. The data sets were further analysed using the single breakpoint (SPB) and genetic algorithms for recombination detection (GARD) methods (Pond et al. 2006) as performed through the web interface of HyPhy (Datamonkey; Pond et al. 2005). Comparing these methods will also provides insights with respect to the underlying demographic processes. If a population growth signal is present in the data, the four-gamete test applied to the nuclear loci should

recover multiple recombination events, whereas the Φw and SPB and GARD tests should find no recombination.

**Population genetics analyses and tests of demographic changes**

The multilocus sequence data were used to conduct multiple analyses to determine whether I could detect a signal of demographic and spatial expansion. I implemented statistical tests developed to detect departures from the Wright-Fisher model (Fisher 1930, Wright 1931), which assumes a constant population size. I considered three tests that use distinct information contained in the sequence data: Tajima's *D* (Tajima 1989), $R_2$ (Ramos-Onsis and Rozas 2002) and Fu's *Fs* (Fu 1997). The first two tests are based on the allele frequency spectrum of mutations and the latter test is based on the haplotype distribution. For all estimates I examined the total number of mutations rather than the number of segregating sites, because in a few instances I observed four different nucleotides segregating. The significance of the demographic statistics was determined by comparing the empirical values with 10 000 coalescent simulations conditioned on the observed sample size and nucleotide diversity, assuming a standard neutral model with no recombination, as performed in DnaSP v5.10.1 (Librado and Rozas 2009). All test were implemented for the entire species range and also separately for the two subspecies. Significantly negative Fu's *Fs*, positive $R_2$ and negative Tajima's *D* are often considered evidence of a population expansion.

To infer the possible direction of the expansion I searched for the geographical axis that maximized multilocus differentiation, estimated as the principal component scores of a covariance matrix of genetic distances between among 16 sites ($N_{average}$ = 5 individuals). This was conducted by fitting a linear model where latitude or longitude was the explanatory variables, and PC1 scores the dependent variable.

**Demographic models**

Using the *isolation with migration* model (Wakeley and Hey 1998, Hey and Nielsen 2007), I estimated migration rates between two populations/ecotypes (see chapter four for details). However, this model assumes that populations have been exchanging migrants at a constant rate since divergence. In order to reconstruct the history of colonization of the arid interior I built three demographic models in which the main difference was the timing of migration events: i) recent, ii) old or iii) continuous. I used Approximate Bayesian Computation (ABC) performed in the DIYABC package v1.0.4.41 (Cornuet et al. 2008) to infer whether the observed genetic variation is compatible with any of the three candidate demographic models. Each scenario is described as a sequence of events (Figure 2b): the establishment of

a small marginal population, period(s) of isolation and subsequent introgression event(s). The demographic parameters used in each model and mutation rates were drawn from a range of priors reported in Table BI – Annexe B. The analyses were implemented for the intron and microsatellite data separately, as an independent test of the consistency of the model selected as the most likely scenario. Genetic variation was summarized into six (intron) and three (microsatellite) estimators that contain information about the level of genetic polymorphism and allele frequency spectrum. I simulated $3x10^5$ data sets per scenario, and used both direct comparison of the summary statistics, as well as logistic regression for the most similar simulated datasets and observed data to derive a posterior probability for each scenario.

**Demographic parameter estimates and divergence time**

I used the coalescent-based *isolation with migration* model (Wakeley and Hey 1998, Hey and Nielsen 2007) to infer changes in the effective population sizes of the ancestral source population ($\theta_a$) and the two descendent populations (subspecies) currently occupying distinct climatic niches ($\theta_{cinerea}$, $\theta_{coryphaeus}$), as well as to estimate migration rate ($m_{cinerea \text{ to } coryphaeus}$ and $m_{coryphaeus \text{ to } cinerea}$), as implemented in IMa (Hey and Nielsen 2007). After multiple preliminary runs to determine the appropriate prior ranges for each of the three parameters and confirming proper MCMC mixing, the final runs were performed with 30 chains with a geometric heating scheme ($g_1 = 0.3$ and $g_2 = 0.9$). The first $1x10^6$ steps were discarded as the burn-in and the MCMC allowed to continue until effective sample size > 500. I used an *inheritance scalar* to adjust the parameters in the model: 0.75 - Z-linked and 1.0 – autosomal. The *Likelihood-mode* was then used to test formally if the full model of distinct population size ($\theta_a$, $\theta_1$, $\theta_2$) fitted the data significantly better than the models of constant population size ($\theta_a = \theta_1 = \theta_2$). The significance of the test statistics (2 x log-Likelihood ratio) was assessed using a chi-squared test (Hey and Nielsen 2007). I used the point estimates of the effective size for the ancestral and descendent populations to calculate a growth ratio $\alpha_i$ whereby the growth ratio of *cinerea* $\alpha_1 = \theta_{cinerea} / \theta_a$, and *coryphaeus* $\alpha_2 = \theta_{coryphaeus} / \theta_a$. In order to estimate divergence time, in years, I used a species-specific generation time of one year (see chapter three) and assumed a neutral mutation rate of $1.35 \times 10^{-9}$ substitutions/site/year and $1.45 \times 10^{-9}$ substitutions/site/year for the autosomal and the z-linked loci, respectively (Ellegren 2007). The geometric mean of the mutation rate for the six loci was $1.92 \times 10^{-6}$ per substitution/locus/year.

## RESULTS

### Range dynamics

After removing the possible effects of geography (co-variates: latitude and longitude), the difference in the current climatic niches of *cinerea* and *coryphaeus* was highly significant ($PC1_{climate}$: $F_{4, 53} = 24.77$, $P < 0.001$; $PC2_{climate}$: $F_{4, 53} = 25.94$, $P < 0.001$; Figure 3). The distribution models, based on the current climatic niche, revealed that the Pleistocene climatic fluctuations had little effect on the paleodistribution of the subspecies (Figure 4). The paleodistribution models further indicates that the climatic conditions currently occupied by the two subspecies have been in place since 21 000 ybp.

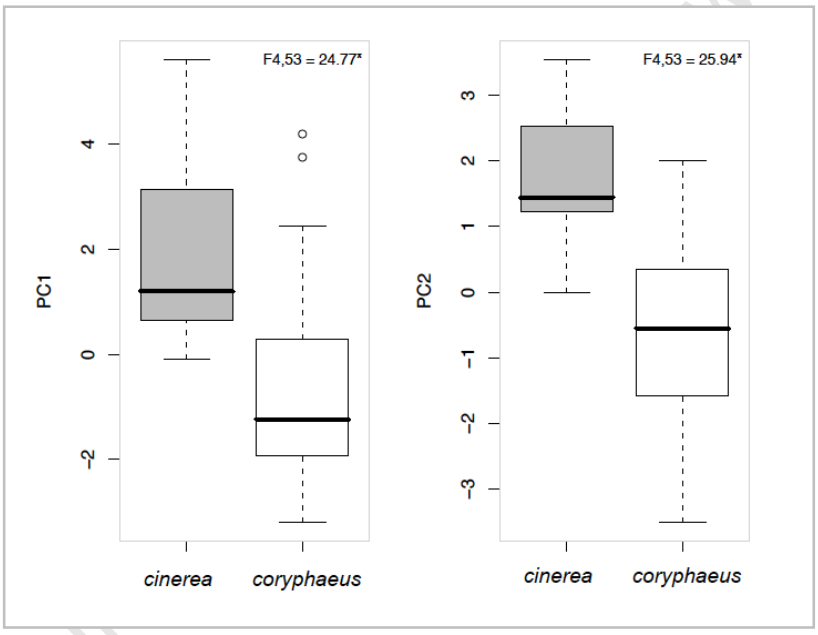


**Figure 3**. Climatic niche divergence along the first two principal component axes (PC1 and PC2) between *C. c. cinerea* and *C. c. coryphaeus.*

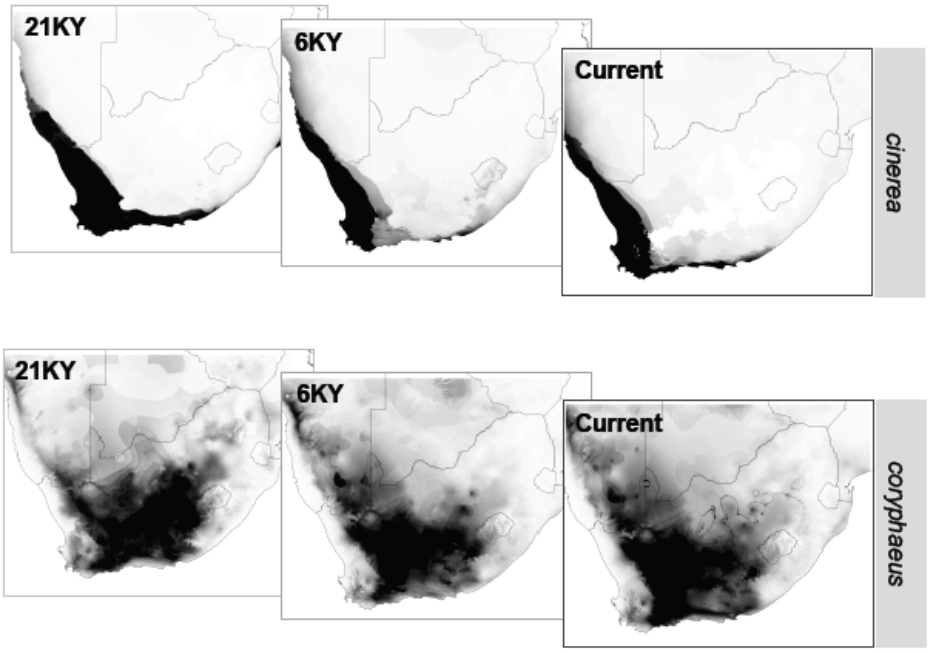


**Figure 4**. Geographical projection of the current climatic niche and predicted past range of the two subspecies *cinerea* and *coryphaeus*. 6KY: Holocene, 6 000 years before present; 21KY: Last Glacial Maximum, 21 000 years before present).



**Historical demography**

*Genetic variation and demographic expansion*

Using the four-gamete test I detected a few recombination events in four (Fib5, Gapdh, TGFb2 and *26438*) out of the six loci analysed. Nevertheless, the Φw test and the SPB and GARD methods did not detect any recombination events, and thus all the subsequent analyses were performed using the full sequences.

Of the 1 825 autosomal base pairs analysed, 108 were polymorphic. The multilocus Waterson's theta for the total pooled samples was $\theta w$ (SD) = 0.011 (0.004). The levels of nucleotide diversity were similar between the two subspecies (Figure 5), with the standard deviations of $\theta w$ for *cinerea* and *coryphaeus* overlapping (0.0090 ± 0.003; 0.0089 ± 0.003). In contrast, *coryphaeus*, the subspecies that occurs in the arid interior, exhibited a significantly reducted mtDNA nucleotide diversity when compared with *cinerea*: $\theta w$ *coryphaeus* (SD) = 0.0062 (0.002) vs. $\theta w$ *cinerea* (SD) = 0.0164 (0.002).

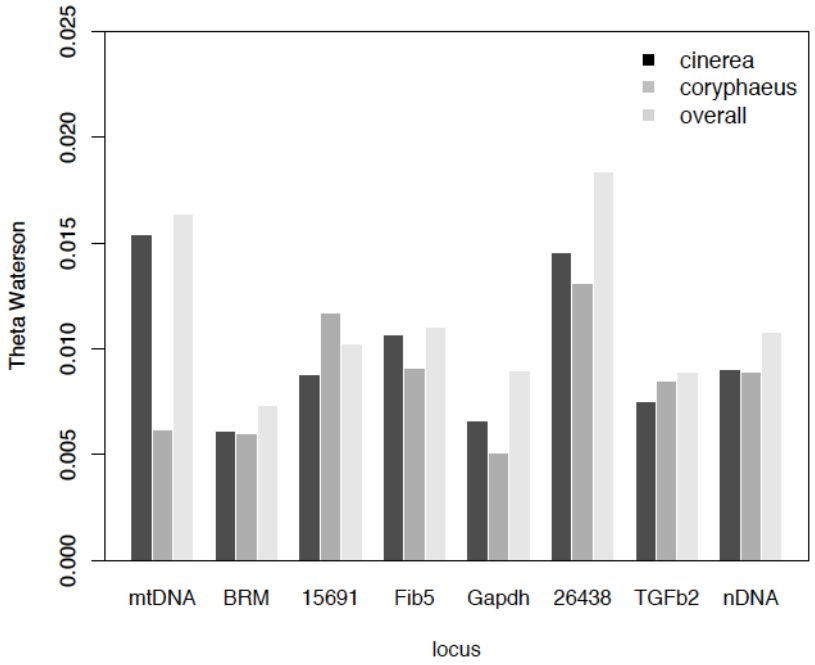


**Figure 5.** Genetic variation, *θw,* measured for mtDNA, five autosommal loci (15691, Fib5, Gapdh, 26438, TGFb2), one Z-linked locus (BRM) and overall nuclear introns for the species and within each subspecies.

Only two loci in *cinerea* and *coryphaeus* showed a significant skew towards an excess of rare alleles, as measured by Tajima's *D* statistics. Tajimas's *D* values ranged from -1.578 to 0.329 in *cinerea*, and from -2.154 to -0.3319 in *coryphaeus*. Although negative in the majority of loci, coalescent simulations revealed that the observed mean values within each subspecies were not significantly negative (Table I). The $R_2$ value was significant in only one locus for *cinerea*. In contrast, Fu's *Fs* differed significantly from expected values for every locus in *coryphaeus* and in all but one locus in *cinerea* (Figure 6). The shapes of the allele frequency spectra in both subspecies revealed multiple modes for the majority of the loci. There was an excess of low-frequency alleles as well as a high proportion of intermediate frequency alleles (Annexe B: Figure B1). I found no geographical trend in genetic differentiation across the whole range of the species ($PC1_{introns}$ ~ latitude: $r^2 = 0.064$; $PC1_{introns}$ ~ longitude: $r^2 = 0.024$) or within the range of each subspecies.

**Table I**. Estimates of genetic diversity ($\theta w$) and tests statistics of demographic change (T*D* – Tajimas's *D*, $R_2$, *Fs* – Fu's *Fs*) across six nuclear introns.

| | *cinerea* | | | | *coryphaeus* | | | | Overall | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Locus** | ***θw*** | **T*D*** | **$R_2$** | ***Fs*** | ***θw*** | **T*D*** | **$R_2$** | ***Fs*** | ***θw*** | **T*D*** | **$R_2$** | ***Fs*** |
| **15691** | 0.008 (0.005) | **-1.578** | 0.048 | **-3.760** | 0.012 (0.004) | **-2.154** | 0.038 | **-10.494** | 0.010 (0.004) | **-1.769** | 0.021 | **-7.272** |
| **BRM** | 0.006 (0.003) | **-1.530** | 0.044 | **-4.430** | 0.006 (0.003) | **-1.447** | 0.043 | **-6.798** | 0.007 (0.003) | **-1.627** | 0.029 | **-10.352** |
| **Fib5** | 0.011 (0.003) | -0.329 | 0.088 | -4.793 | 0.009 (0.003) | -0.332 | 0.008 3 | **-8.344** | 0.011 (0.003) | -0.604 | 0.067 | **-13.670** |
| **Gapdh** | 0.007 (0.003) | -1.152 | **0.052** | **-5.544** | 0.005 (0.002) | -0.770 | 0.064 | **-2.434** | 0.009 (0.003) | **-1.512** | **0.034** | **-19.112** |
| **26438** | 0.015 (0.006) | -1.413 | 0.052 | **-9.409** | 0.013 (0.005) | -1.336 | 0.051 | **-6.937** | 0.018 (0.006) | **-1.667** | 0.034 3 | **-20.187** |
| **TGFb2** | 0.008 (0.003) | -0.679 | 0.079 | **-16.230** | 0.009 (0.003) | -1.141 | 0.058 | **-23.546** | 0.009 (0.003) | -1.075 | 0.053 | **-34.404** |

Significant values in bold. Standard deviation in parenthesis.

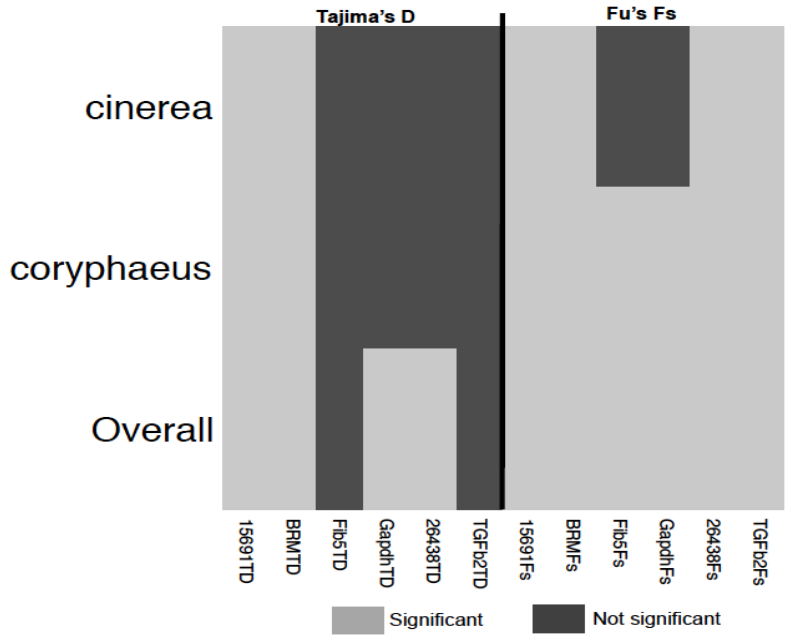


**Figure 6**. Summary of the results obtained with the two most commonly used tests statistics implemented to detect departures from the constant population size model: Tajima's *D* and Fu's *Fs*. Suffix 'TD' and 'Fs' after locus name denote the test statistics, Tajima's D and Fu's Fs, respectively.

*Reconstructing the history of colonization and timing of migration*

The approximate Bayesian framework used here revealed that the expansion into the arid interior (new habitat) likely involved the establishment and demographic growth of a small marginal population that exchanged migrants at multiple times (Figure 7). Based on the intron data, the model with highest posterior probability (0.9822; 95% high posterior density: 0.9786 – 0.9859) was model B (old introgression). By contrast, the microsatellite data supported the scenario of a recent introgression (model A) with posterior probability 0.7204 (95% high posterior density: 0.7122 – 0.7285). Both the direct and logistic regression methods consistently recovered the same demographic model.

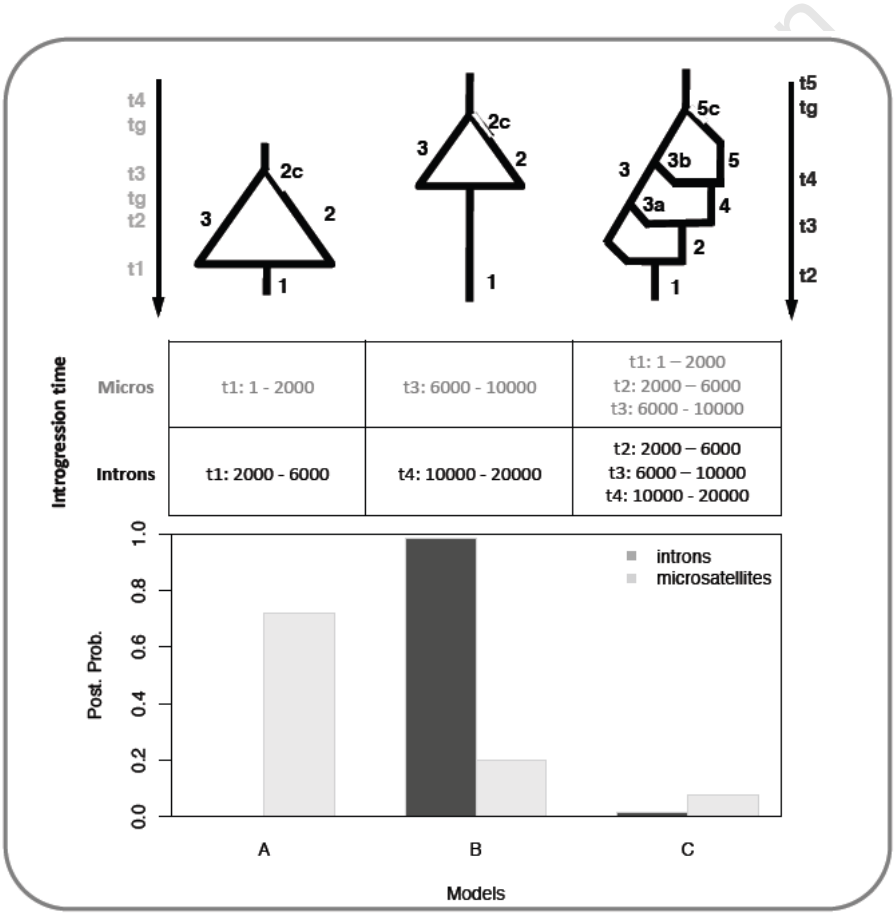


**Figure 7.** Microsatellites and intron data support (posterior probability) for the three demographic models tested under the ABC framework. Details for the priors selected for the size of population 1 – *i* and timing of events t1 - t2 are reported in Table BI (see Annexe B).

*Migration directionality, population size and time of divergence*

The coalescent-based MCMC simulations implemented in IMa allowed me not only to estimate population sizes, but also to compare models of divergence with constant *vs* distinct effective population sizes. There was a clear signal of directional migration, with migration rates from the west (*cinerea*) towards the arid interior (*coryphaeus*) being nine times greater than in the reverse direction (Table II). The size of the ancestral population was significantly smaller than the estimates obtained for *coryphaeus* (Table II), and the effective population size of *coryphaeus* was larger than the value estimated for *cinerea,* although the 90% HPD overlapped. The growth ratio of *coryphaeus* ($\alpha_2$: 6.600) was six times larger than of *cinerea* ($\alpha_1$: 1.311). Using log-Likelihood ratio tests to compared nested models with the full model, I could reject all models of stable populations sizes (i.e. $\theta_1 = \theta_2 = \theta_a$) and models where the ancestral population and *coryphaeus* had the same effective size (i.e. $\theta_1, \theta_2 = \theta_a$; Table III). All the models of divergence in isolation ($m_1 = m_2 = 0$) were rejected (see Annexe B, Table BII) confirming the results reported in chapter four. The scaled divergence time between *cinerea* and *coryphaeus* was 0.253 (90% HPD: 0.147 – 0.485), which given the mutation rates assumed (see methods) corresponds to 131 144 years before present (76 205 - 251 425 ybp) and corroborated the most likely demographic scenario of an old colonization of the arid interior as suggested by the Approximate Bayesian approach.



**Table II:** Demographic parameters estimated under a model of *isolation with migration* using IMa software, using six nuclear introns. $\theta$: population size, *m*: migration, *t*: time of divergence

| Model Parameters | Highest Posterior Probability | 90% HDP |
|---|---|---|
| *θ ancestral* | 1.3652 | 0.637 - 2.397 |
| *θ cinerea* | 1.7289 | 0.576 - 5.613 |
| *θ coryphaeus* | 9.071 | 3.549 - 39.652 |
| *m cinerea - coryphaeus* | 14.050 | 0.050 - 74.550 |
| *m coryphaeus - cinerea* | 4.350 | 0.050 - 39.050 |
| *t* | 0.253 | 0.147 - 0.485 |

HPD: high posterior density.

**Table III**: Likelihood ratio test of nested models with equal and differential population sizes against the full model ($\theta_1$ , $\theta_2$ , $\theta_a$). $\theta_1$ = population size for *cinerea* ; $\theta_2$ = population size for *coryphaeus.*

| Population size | Model | logLikelihood (Model \| data) | 2LLR (df) |
|---|---|---|---|
| Equal size | $\theta_1 = \theta_2 = \theta_a$ , $m_1$, $m_2$ | -15.8766 | **20.6294** (2) |
| | $\theta_1 = \theta_2 = \theta_a$, , $m_1 = m_2$ | -23.413 | **35.7023** (3) |
| Different for *coryphaeus* | $\theta_1 = \theta_a$ , $\theta_2$ , $m_1$, $m_2$ | -5.9422 | 0.7607 (1) |
| | $\theta_1 = \theta_a$ , $\theta_2$ , $m_1 = m_2$ | -10.063 | 9.0024 (2) |
| Stable for *coryphaeus* | $\theta_1$ , $\theta_2 = \theta_a$ , $m_1$ , $m_2$ | -11.2608 | **11.3978** (1) |
| | $\theta_1$ , $\theta_2 = \theta_a$ , $m_1 = m_2$ | -15.4411 | **19.7584** (2) |

df: degrees of freedom. Bold indicate tested that are significant for $\alpha = 0.01$.
2LLR = 2 (logLikelihood full-model / logLikelihood alternative model).



## DISCUSSION

The analyses of nucleotide variation in the two subspecies of the Karoo Scrub-Robin *cinerea* and *coryphaeus* provided insights into the demographic processes underlying an adaptive range expansion. Present occurrence data revealed that the two subspecies are largely confined to different climatic niches (Figure 3) and that the climatic requirements have been present for at least the past 21 000 years. The multilocus intron data revealed: similar patterns of genetic variation between subspecies; significantly negative Fu's *Fs* values; allele frequency spectra with an excess of low-frequency alleles but with multiple modes; effective population size for *coryphaeus* nine times higher than *cinerea*; the colonization of the arid interior (new habitat) was estimated to have occurred during the Pleistocene (between 251 000 years – 76 000 years ago). Furthermore, by analysing microsatellite and sequence data independently, I could demonstrate that there has been effective migration between the two subspecies since the colonization of the arid zone. All these findings, together with the lack of correlation between the population size (large; Table II) and mitochondrial variation observed (small; Figure 5) in the population inhabiting the arid interior are consistent with a model of trait-dependent niche expansion (Figure 2, model C), i.e. the range and demographic expansion was facilitated by the appearance of a newly adaptive mitochondrial haplotype and accompanied by neutral gene flow (see also chapter four).

**Demographic and range expansion**

The predictive models of the past geographical range of the subspecies did not support the hypothesis of a range expansion caused by the climatic oscillations that characterized the Quaternary and are thought to have affected Southern Africa (Figure 4). The data presented here rather suggested that suitable climates persisted over a large area, closely concordant to the present-day distribution of the species and subspecies. It is, however, important to consider that these models are not an absolute prediction of the fundamental or realized niche of an organism. Moreover, projecting current climatic requirements onto paleoclimatic surfaces assumes temporal stability of the climatic niches of the species/populations, thus overlooking the possibility that populations can adapt to new environmental conditions (e.g. Parmesan 2006 and references there in; see also Gavrilets and Losos 2009 for classical examples of adaptive evolution).

Classic population genetic summary statistics did not revealed any particular trend in genetic variation between subspecies, as predicted under the scenario of range expansion by colonization of new habitat followed by recurrent gene flow. Implementing statistical tests that use different information contained in the genetic data, and hence have different sensitivity to the effects of expansion, provided interesting insights into the demographic trend of the Karoo Scrub-Robin and the two recognized subspecies.

Spatially explicit simulations shown that when a range expansion is accompanied by demographic growth and populations exchange large numbers of migrants ($N_em$ = 100 individuals) the signature left is similar to one resulting from demographic growth alone, causing Tajima's *D* and Fu's *Fs* to be significantly negative (Ray et al. 2003). In contrast, when the number of migrants entering the population is smaller ($N_em$ = 10 migrants), genealogies are characterized by coalescent events occurring within the population but also events dating back to the time of expansion. Low migration rates reduce the effectiveness of Tajima's *D* to detect demographic expansions, whereas Fu's *Fs* does not suffer from the same limitation (Ray et al. 2003). These results obtained by Ray et al. (2003) and Excoffier (2004) allowed me to reconcile the discrepancy between Tajima's *D* and $R_2$ with Fu's *Fs* under the scenario of range expansion with lower migration rates ($N_em$ = 10). However, it is important to note that the value modeled as a *low* rate of gene flow rate is larger than '1', the theoretical number of migrants among demes required for them to evolve as a panmitic unit (Wright 1943, Kimura and Maruyama 1971). The lack of a longitudinal/latitudinal trend in genetic dissimilarity across the whole species' range and within subspecies provides further support to the scenario of range expansion followed by gene flow from the source population. Effective migration not only evens out the genetic diversity, it also spreads old alleles across the range. Interestingly, the coalescent-based estimates of migration using intron data supported the

scenario of expansion into the arid zone. Gene flow was asymmetrical, with migration rates from *cinerea* into *coryphaeus* 10 times larger than the estimates in the opposite direction.

Demographic dynamics during range changes strongly influence genetic patterns (Excoffier and Ray 2008, McInerny et al. 2009). The similarity in allele frequencies found at microsatellites across the range of the Karoo Scrub-Robin (chapter four) is indicative of a panmitic population. But, is it a consequence of range expansion with recurrent gene flow or the result of a colonization of new habitat followed by isolation and recent contact? Although central to understanding the demographic dynamics and the evolutionary history of populations/species, answering this question requires the estimation of both rate and timing of migration: a task that is technically challenging, as revealed in a recent simulation study by Strasburg and Rieseberg 2011.

My approach to the before mentioned challenge involved the simulation of three scenarios (Figure 2) under an approximate Bayesian computation framework (Beaumont et al. 2002). These analyses revealed that the Karoo Scrub-Robin likely expanded its range from a small peripheral population with subsequent exchange of migrants with the source population. Using microsatellite and intron data independently revealed to be complementary and allowed me to better understand the demographic trend of the species. By comparing the estimates of ancestral population size with the current effective population size of both subspecies revealed dramatic demographic expansion in the arid zone *coryphaeus* ($\alpha_2 >> \alpha_1$). The split between *cinerea* and *coryphaeus,* estimated from the nuclear genetic variation to date back to the mid-Pleistocene is consistent with ABC model B (old introgression), the one that best fitted the intron data. The population size estimated with IMa could be biased, because this model assumes that populations maintain a constant size after they diverged, an assumption violated by the data presented here. However, Strasburg and Rieseberg (2009) demonstrated by simulations that the trend, if any, is to slightly underestimate current population sizes, with no effect on other parameters estimates. Therefore, I contend that these results convincingly show a demographic expansion.



**Adaptive evolution during range expansion into novel habitat**

The geographic range of a species is determined by the interplay of ecological and evolutionary processes. Understanding which and how ecological factors maintain or disrupt the demographic dynamics in populations at the margins of their range has been an area of intensive research (reviewed by Sexton et al. 2009). From an ecological perspective, and without arguing for any adaptive change, pushing the range limits and colonizing newly available habitat depends on the niche breadth and behavioural plasticity of peripheral populations. But a species can also expand its range by adapting to marginal habitats (Kawecki 2008): when a new trait appears and improves fitness in the peripheral sink habitat that

usually characterizes the leading edge of an expansion. Some studies, have demonstrated that adaptive evolution might occur during or after the establishment of population in a new environment. These include the selective sweep at pigmentation genes after the expansion of human populations out of Africa (Myles et al. 2007), the change in leg size (longer) during range expansion of cane Toads *Bufo marinus* in Australia (Phillips et al. 2006) or the shift in skin colour in eastern fence Lizards *Sceloporus undulatus*, during colonization of a newly available habitat characterized by distinct substratum colour (Rosenblum et al. 2007).

Since the mid-Miocene, southern Africa has experienced a trend towards increasing aridification and seasonality (Cowling et al. 1997, 1999, Richardson et al. 2001). This climatic trend has caused the progression from closed subtropical woodland to more open drought-tolerant biomes dominated by shrubs and succulents, and an accompanying decrease in primary productivity. These new habitats were thus characterized the new selective regimes, with respect to climate, resources availability, competitors and predators. In the Karoo Scrub-Robin, the fact that the subspecies *cinerea* and *coryphaeus* have divergent climatic niches, together with the large effective population size and low genetic variation found in the mtDNA of *coryphaeus*, the subspecies inhabiting the most arid portion of the species range, suggests that expansion into a newly available habitat was mediated by a mitochondrial variant, which enhanced individual fitness in the new climatic and nutritional niche. The congruent geographical variation in plumage colouration and mtDNA variants could simply be a by-product of a physiological response to altered climate or food resources through mitochondrial function. But, it could also have an adaptive function such as background matching, given the time spent by these birds on the ground while foraging, and thereby reducing predation risks. Background matching is found in other ground-feeding birds inhabiting the same region, e.g. *Certhilauda* Larks (Dean et al. 1991, Ryan et al. 1997).

In the traditional perspective, the process of matching between organismal traits and its environment (adaptive evolution) results from spatially varying selection maintaining adaptive variants. Under this viewpoint, selective factors are extrinsic to the organisms and thus the directionality of causation is from the environment to the individual. However, several researchers have proposed a feedback loop where organisms can modify their relationship with the environment; this idea has been termed niche construction (Odling-Smee 1988, Odling-Smee et al. 2003). One of the most well-know examples of such phenomena is the fixation of a lactase variant in humans associated with pastoralism (Gerbault 2011). If we define niche as a feature inherent to individuals, the niche construction theory could offer an alternative perspective about the process of adaptive range expansion. If some individuals start to explore a portion of the habitat that became available and where selective factors are different from the ancestral pressures, they likely change some axes if their *n*-dimensional niche. The offspring will thus

inherit the niche and the trait(s) associated with the niche construction. Subsequently those traits may increase in frequency or perhaps do to fixation (Day et al. 2003). In the Karoo Scrub-Robin, the invasion of the arid habitat could have been possible due to the opportunity to explore new food resources and/or the ability to reduce water loss by altering activity patterns (behaviour). The role of mitochondria in the energetic metabolism and the evidence that calorie restriction affects respiratory efficiency (Das 2006) suggests that the mtDNA variant might be a niche construction associated trait. Whether the increase in frequency of a particular mitochondrial variant was in the first place driven by directional selection (extrinsic to the individuals) or a result of niche construction when moving into a new habitat is still unclear, but warrants further investigation.

The unprecedented availability of refined analytical methods allowed me to model the underlying historical demography of the Karoo Scrub-Robin and reveal strong support for a scenario of an adaptive range expansion with recurrent gene flow. The results obtained here give further support to the arguments articulated in chapter four that natural selection is currently maintaining the populations of this bird adapted to local conditions. Furthermore, it illustrates the importance of considering the role of population demographic dynamics (ecological factors) in evolution and also vice-versa (Kokko and Lopez-Sepulcre 2006, Schoener 2011).

**ANNEXE B**

**Table BI**. Parameters and priors used for each of the scenarios simulated under the Approximate Bayesian Computation framework. Mutation rates: autosomal introns [$10^{-9}$ - $10^{-7}$], Z-linked loci [$10^{-9}$ - $10^{-8}$] and microsatellites [$10^{-4}$ - $10^{-3}$].

| Scenario | Parameter | Description | Priors Range |
|---|---|---|---|
| Colonization with recent introgression | N1, N2, N3 | population size | [100 - 100 000] |
| | N2c | marginal population | [10 – 500] |
| | t1 / t2 | time introgression | [1 - 6 000] / ]2 000 – 6 000] |
| | tc | time of establishment | [1- 500] |
| | t2 / t4 | time of divergence | ]2 000 – 6 000] / ]10 000 – 20 000] |
| | ra1 | introgression rate | [0.001 – 0.999] |
| Colonization with old introgression | N1, N2, N3 | population size | [10 - 100 000] |
| | N2c | marginal population | [10 – 500] |
| | t3 / t4 | time introgression | [1 - 6 000] |
| | tc | time of establishment | [1- 500] |
| | t4/ t5 | time of divergence | ]10 000 – 20 000] / ]20 000 – 300 000] |
| | ra1 | introgression rate | [0.001 – 0.999] |
| Colonization with recurrent migration | N1, N2, N3, N3a, N4, N3b, N5 | population size | [10 - 100 000] |
| | N5c | marginal population | [10 – 500] |
| | tc | time of establishment | [1- 500] |
| | t1 / t2 | time introgression | [1 - 2 000] / ]2 000 – 6 000] |
| | t2 / t3 | time introgression | ]2 000 – 6 000] / ]6 000 – 10 000] |
| | t3 / t4 | time introgression | ]6 000 – 10 000] / [10 000 – 20 000] |
| | t4 / t5 | time of divergence | ]10 000 – 20 000] / ]20 000 – 300 000] |
| | ra1, ra2, ra3 | introgression rate | [0.001 – 0.999] |

**Table BII**: Likelihood ratio test of nested models of divergence in isolation ($m_1 = m_2 = 0$). $m_1$ = migration from *cinerea* into *coryphaeus;* $m_2$ = migration from *coryphaeus* into *cinerea.*

| | Model | logLikelihood (Model \| data) | 2LLR (df) |
|---|---|---|---|
| Isolation | $\theta_1 = \theta_2 = \theta_a$ , $m_1 = m_2 = 0$ | -460.517 | **909.9103 (4)** |
| | $\theta_1$ , $\theta_2 = \theta_a$ , $m_1 = m_2 = 0$ | -460.517 | **909.9103 (3)** |
| | $\theta_1 = \theta_2$ , $\theta_a$ , $m_1 = m_2 = 0$ | -460.517 | **909.9103 (3)** |
| | $\theta_1 = \theta_a$ , $\theta_2$ , $m_1 = m_2 = 0$ | -460.517 | **909.9103 (3)** |
| | $\theta_1$ , $\theta_2$ , $\theta_a$ , $m_1 = m_2 = 0$ | -460.517 | **909.9103 (2)** |

df: degrees of freedom. Bold indicate tested that are significant for $\alpha = 0.01$.

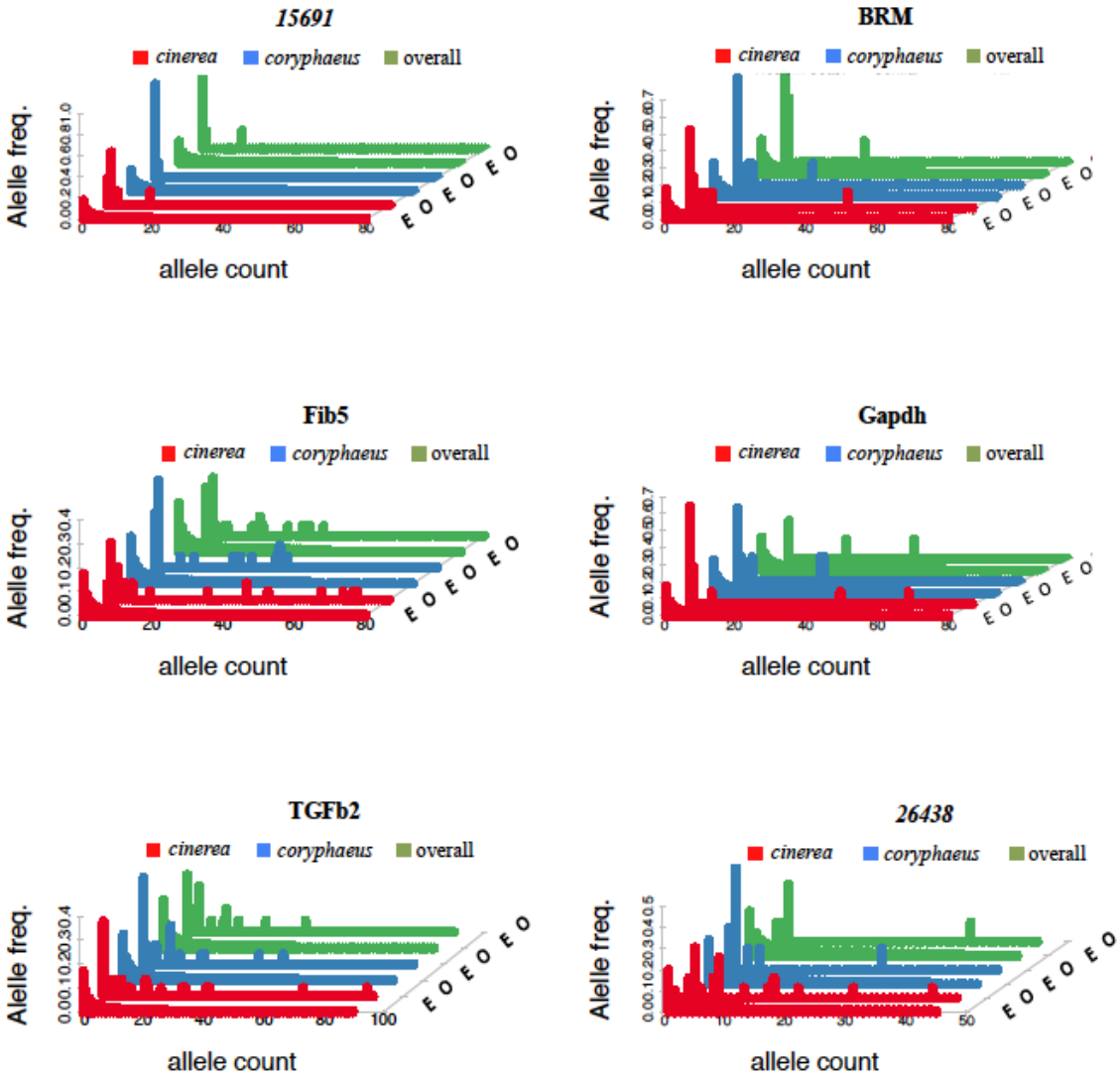


**Figure B1**. Allele frequency spectra for *cinerea* (red), *coryphaeus* (blue) and overall (green), for each nuclear locus analysed. 'O' and 'E' denote the variation observed and the variation expected under a model of stable population size, respectively.

CHAPTER SIX

# Eco-geographical context of morphological and genetic divergence in a forest-dwelling Scrub-Robin of southeast Africa



## INTRODUCTION

The study of the origin of species, i.e. speciation process, has been the career of enthusiastic debate among evolutionary biologists. Two distinct philosophical approaches have emerged. The earliest, proposed by Mayr (1942, 1947) and Dobzhansky (1951) promotes a retrodictive view of the process, inferring the causes of reproductive isolation from a snapshot of the late or completed stage of the process. The second and more recent approach is a population-level perspective (Schemske 2000, Schluter 2001, Turelli et al. 2001, Kirkpatrick and Ravigné 2002, Nosil and Rundle 2004), which brings population genetics concepts into the framework. Proponents of this latter approach posit that speciation is a continuous process of divergence and therefore best understood by studying how ecological and genetic components interact to restrict gene flow among populations. Several empirical studies have demonstrated the importance of considering the continuity of the speciation process (e.g. Ogden and Thorpe 2002, Rosenblum 2006, Butlin et al. 2008, Nosil et al. 2008, Thorpe et al. 2008, Pereira and Wake 2009).

The population-centered perspective has led to a renewed emphasis on the role of ecological factors in the process of speciation. In fact, ecology, i.e. factors that affect population demographic dynamics and distribution, was the underlying concept of the majority of the theories of speciation proposed in the early days of the *modern synthesis* (e.g. Mayr 1942, 1947; Dobzhansky 1951). Yet, its relevance in the process of speciation has remained unappreciated mainly due to the emphasis given to the geographical pattern (geographical theory of speciation; Mayr 1942) over the mechanisms that underlie it. Mayr's formulation about the role of geographical isolation can be elaborated in terms of ecological and micro-evolutionary mechanisms. Fundamentally, geographical isolation results from the failure to adapt to the environmental conditions that disrupted the initially continuous range. While natural selection 'pushes' individuals far from the area where fitness is low, migration is reduced and consequently gene flow diminished. This affects not only the demographic dynamics of local populations, but also the dynamics of the overall range. Depending on how much local conditions affect the demography of particular phenotypes (i.e. natural selection), the populations can either diverge adaptively and possibly attain reproductive isolation, or the process of divergence can be reversed if migration and gene flow resumes

The geographical pattern of intra-specific organization of populations in distinct but contiguous populations (conjunct distributions, *sensu* Mayr 1963) is common in nature. This geographical pattern is thought to be a response of organisms to local environmental conditions rather than a response to physiographical features (Endler 1977), such that the contiguous populations are often designated as ecotypes, i.e. populations within a species that occupy distinct environmental conditions and exhibit

phenotypic differences. Species where ecotypes remain contiguous thus form interesting systems with which to study the underlying mechanisms that may promote divergence.

The forest-dwelling Brown Scrub-Robin, *Cercotrichas signata* exemplifies the concept of a conjunct distribution. The bird's range spans a narrow and scattered ribbon of forest in the southern and eastern parts of southern Africa, from -33S to -23S; across this range there is considerable variation in climatic conditions. Two subspecies are currently recognized: the dark olive-brown *C. s. signata* and the paler, smaller and shorter-billed *C. s. tonguensis* (Hockey et al. 2005). Because *signata* spans the temperate and subtropical forest of Pondoland and highlands of Swaziland, Mpumalanga and Limpopo (reaching 2000 m above sea level) and *tonguensis* is restricted to the lowland tropical forests of the Maputaland region, the two subspecies are also considered ecotypes (term used hereafter). The boundary between the ranges of the two ecotypes is centered on the coastal plains between St. Lucia Estuary and Mtunzini (Hockey et al. 2005, Figure 1). This area is considered a biogeographical transition between the temperate and sub-tropical Pondoland and the tropical Maputaland region (Poynton 1961, Van Wyk 1996). It also coincides with the region where the 18°C mid-winter isotherm crosses southern Africa, a climatic feature widely used to divide tropical from non-tropical climates (Koppen Climatic System; Koppen 1936). The current distribution of the forests occupied by this Scrub-Robin differs from its historical range. The dry and cold climate that dominated southern Africa during the Last Glacial Maximum is though to have caused a major contraction of forest into small isolates from which it expanded when the climate ameliorated following the glacial retreat (Eeley et al. 1999, Lawes et al 2007).

In this study I explore the ecological and evolutionary processes underlying the conjunct distribution in the Brown Scrub-Robin in an attempt to provide insights about the continuous process of the formation of species in the southeast coast of Africa, an understudied part of the world. To achieve this goal I used a framework of three linked components. In the first step, I characterized the eco-geographical context of the contemporary, conjunct distribution of the Brown Scrub-Robin, with the respect to the nature of the environmental changes that historically affected the habitat, i.e. climate. If during isolation the dimensions of the niche that are affected by climatic factors were conserved, Brown Scrub-Robins should be absent from any areas that conform to the set of climatic conditions hypothesized to have disrupted the species distribution in the past (predicted as unsuitable). Also, the climatic requirements of populations should reciprocally predict the distribution of each other in the geographical context. In contrast, if disruption of the range resulted in divergence along the climatic axis of the niche, or any other related axis of the *n*-dimensional niche, ecotypes should be mutually exclusive in the geographical context. In the second component, I quantified morphological divergence with the rationale that body mass and body size have functional consequences for multiple life-history features, and are therefore

subject to selective pressures (Gaston et al. 2007). Particularly in birds, the geographical variation in body size, with larger-body size associated with higher latitudes and cooler climates, is thought to result from the interaction of species-specific physiology with the environment and resource availability (Olson et al. 2009). In the final step, I examined patterns of genetic structure and gene flow along the geographical axis that crosses the boundary between the two ecotypes and three predicted forest remnants. If divergence was initiated by the failure of individuals to adapt to changing conditions that disrupted the range for a given period of time, and thus impeded gene flow, drift could promote divergence. As ecological factors promoted population demographic growth, migration among the isolated populations was eventually resumed. However, because niches remained similar, complete reproductive isolation is unlikely and gene flow should proceed. Alternatively, natural selection could have played an important role during the reduction and eventual cessation of migration associated with the disruption of the range and lead to adaptive divergence. Populations can be found at different stages of divergence: 1) locally adapted if natural selection maintains the adaptive traits at higher frequencies regardless of neutral gene flow or, 2) at the final step if genes/traits contributing to the adaptive divergence have a pleiotropic effect on or are linked to other traits that in turn affect reproductive isolation and impede gene flow.

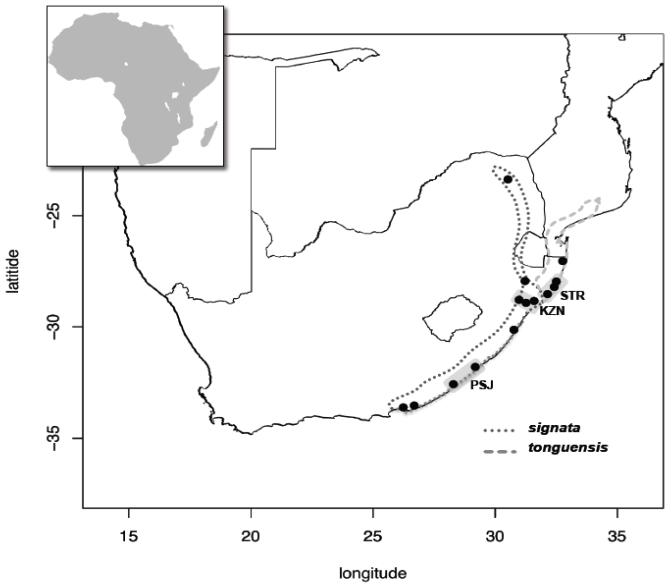


**Figure 1.** Geographical distribution of the Brown Scrub-Robin, *Cercotrichas signata,* along the southeastern coast of Africa, with the range of the two ecotypes depicted. The points mapped correspond to sites from where ecological, morphological and genetic data were gathered. PSJ: Port St. Johns coastal scarp forest in the Eastern Cape; KZN: scarp forest in KwaZulu-Natal, STR: Indian Ocean coastal forest in the vicinity of St. Lucia Estuary in northern KwaZulu-Natal.

## METHODS

### Eco-geographical context

Climatic data were obtained from the WorldClim database (Hijmans *et al.* 2005). DIVA-GIS v5.4 (Hijmans *et al.* 2004) was used to extract data for nineteen climatic variables at a resolution of 30 arc-seconds (1 $km^2$ grid) for each of the geo-referenced localities. In order to access the spatial context of the climatic niches of *signata* and *tonguensis* I used the maximum entropy models, as performed in MaxEnt v. 3.3.3. (Phillips et al. 2006, Phillips and Dudik 2008). Presence-only point locality data used in these models was primarily obtained from my field expeditions and supplemented with records of adult birds kindly provided by Richard Dean, Craig Symes and Joseph Heymans. I modelled the distribution of climatic variables supporting the occurrence of the species (36 localities) and also the two ecotypes independently (28 sites for *signata* and 8 sites for *tonguensis*; supplementary material Table I) at a resolution of 1 $km^2$ using eight independent bioclimatic variables (Pearson's correlation coefficient $r < 0.80$) produced by Hijmans et al. (2005). The accuracy of the models was assessed by estimating the 'area under the receiver operating curve' – AUC. The predicted probability of occurrence obtained was spatially projected using ArcMap 9.3.

In order to estimate the geographical extent of overlap of the climatic niche, and thus predict the possibility of co-occurrence, I first converted the cumulative prediction obtained with MaxEnt into a binary prediction, i.e. presence vs. absence, using the minimum probability of occurrence as a threshold. Deciding which threshold is the most appropriate is subjective. However, since the Brown Scrub-Robin is a territorial year-round, I made the simplifying assumption that occurrence data reflects a resident bird. All the routines involving GIS were performed in ArcMap 9.3 (ESRI).

### Morphological variation

Birds were captured using mist-nets and walk-in traps. Morphological data was collected from 81 individuals from 14 localities across the range of the species: 55 *C. c. signata* from ten sites and, 26 *C. c. tonguensis* from four localities. However, samples sizes from four locations within *C. c. signata* were small and thus I decided to use geographical proximity and prior knowledge about forest type occupied to combine those into two sites. Therefore, morphological analyses were performed with individuals grouped in 12 sites ($n_{per\ site} > 3$). Tarsus-length was measured with a digital caliper to an accuracy of 0.1 mm, wing-length (from the carpal joint to tip of the longest primary feather) was measured with a wing rule to an accuracy of 0.1 mm, and body mass was measured with a 50 g Pesola spring balance to an accuracy of 0.5 g.

Body mass (g), wing-length (mm) and tarsus-length (mm) of adult birds (n = 81: 29 females and 52 males) were used to quantify morphological variation across the entire range of the species. The sex of the birds was determined using a PCR-based assay (Fridolfsson and Ellegren 1999). Prior to all the analysis, morphological data were $\log_{10}$ transformed to meet assumptions of normality. Also, I used the residual values from the regression of log-transformed wing on tarsus as a proxy of structural body size. Exploratory scatterplots indicated morphological variation within the range of the species and between sexes (Annexe C, Figure C1). Therefore, I designed a factorial analysis (uni-variate ANOVA) with geography and sex as fixed-effect factors to test three null hypotheses regarding morphological variation: i) no geographical effect, ii) no sexual dimorphism and iii) no geographical effect on sexual dimorphism, represented by the interaction geography*sex. All statistical analyses were implemented in R 2.10.1 (R Development Core Team 2011).

**Selective regime**

I used principal components analysis (PCA) on a correlation matrix of nineteen climatic variables (see eco-geographical context section above for details) to derive orthogonal PC axes. The first two PC axes accounted for 78.42% of the variation in the climate data, and hence only PC1 and PC2 were retained for further analysis. These axes were used as the explanatory variables to test for the effect of climate on the amount of morphological variation using a regression approach. In addition to the previous linear models, I also tested for a non-linear relationship between morphology and climate, by including a quadratic term in the model. The rationale behind this was that a subtle change in the climatic conditions should have a dramatic effect on phenotype. R 2.10.1 (R Development Core Team) was used to perform the statistical analyses.



**Genetic data**

*DNA amplification*

The blood or muscle samples collected in the field, from 14 sites, were used to quantify genetic variation. Genomic DNA was extracted from blood/tissue samples using DNeasy kits (Qiagen, Valencia, CA). I sequenced two mitochondrial protein-coding genes (ATP6 and ATP8; E. Bermimgham Lab) and six nuclear introns: Gapdh-intron11 (*Gallus gallus* chromosome 1, Friesen et al. 1997), βFib-intron5 (*Gallus gallus* chromosome 4, Fuchs et al. 2004, Kimball et al. 2009), *15691* (*Gallus gallus* chromosome 5, Backström et al. 2008) and BRM-intron15 (*Gallus gallus* Z chromosome, Goodwin 1997), TGFb2-intron 5 (*Gallus gallus* chromosome 3, Burt and Paton 1991, Primer et al. 2002) and *26438* (*Gallus gallus* chromosome 3, Backström et al. 2008). PCR amplifications were performed in a total volume of 10 µl

with 10 – 20 ng of genomic DNA, GeneAmp 10x PCR Gold Buffer, 0.5 U of Taq polymerase (Roche), 2.0-2.5 mM $MgCl_2$, 0.3 mM of each dNTP and primer concentrations of 0.15 µM. The thermocycling profile consisted of an initial denaturation step at 95°C for 3 min followed by 35 cycles at 95°C for 30 seconds, and a locus-specific annealing temperature of 55°C - 60°C for 30 seconds, and 72°C for 30 seconds, with a final extension step at 72°C for 7 min. PCR products were cycle sequenced in both forward and reverse directions using the ABI BigDye Terminator Kit v3.1 (Applied Biosystems, Foster City, CA, USA) and then analyzed on an ABI 3730 automated sequencer. Sequences were edited and aligned using CodonCode Aligner v3.5.2 (CodonCode Corporation 2009) and Geneious Pro v5.4 (Biomatters Ltd 2011). Length polymorphisms were found in two nuclear loci: TGFb2 and *26438*. For TGFb2, I chose to remove indels from the alignments prior to further analysis. For *26438*, because length polymorphism was located at the very end of the fragment I decided to truncate the sequence data at the indel in the forward direction.

*Population Structure and Gene flow*

Prior to analyses the gametic phase was determined using PHASE 2.1 (Stephens et al. 2001, Stephens and Donnelly 2003). The PHASE algorithm was run twice for each locus ($10^4$ main iterations and $10^3$ burn-in, -x100 option) and the highest-probability alleles used for subsequent analyses (Harrigan et al. 2008, Garrick et al. 2010). Recombination was tested by screening the sequences for possible breaking points using the GARD module (Pond et al. 2007) as implemented in DataMonkey, the Web interface for HyPhy (Pond et al. 2006).

I used STRUCTURE v2.3 (Pritchard et al. 2000, Hubisz et al. 2009) to identify groups of randomly mating individuals with different allele frequencies, by minimizing deviations from Hardy-Weinberg expectations and linkage disequilibrium. The analysis was restricted to individuals that were missing data at no more than one locus (n = 63) using the *admixture* model with *correlated allele frequencies*. I ran five pseudo-replicates with $10^6$ Markov-Chain-Monte-Carlo iterations following a burn-in of $10^5$. Because mtDNA is inherited as a linkage group I contend that it cannot be incorporated in the STRUCTURE analysis. Therefore, I used the analysis of molecular variance (AMOVA, Excoffier et al. 1992) to test whether the mtDNA variation was consistent with the nuclear genetic clusters defined by STRUCTURE. The AMOVA was implemented in ARLEQUIN v3.1 (Excoffier et al. 2005) and the statistical significance of the *F* statistics was assessed using $10^5$ permutations.

Given the contemporary genetic structure (see results) I used the coalescent-based *isolation with migration* model, as implemented in software IMa (Wakeley and Hey 1998, Hey and Nielsen 2007), to

test whether the boundary between the two ecotypes has been *impermeable* since divergence, or whether gene flow has occurred. In order to answer this question the IMa analysis was restricted to populations closer to the current geographic boundary: KwaZulu-Natal scarp (KZN, n = 12) in the sub-tropical Pondoland and St. Lucia region (SLR, n = 10) in tropical Maputaland. I also compared migration between two forest blocks within the Pondoland region: KZN and the Port St. Johns coastal scarp (PSJ, n = 16). These areas are thought to have supported forest remnants during the LGM (Eeley et al. 1999; Hughes et al. 2005). Two separate analyses were performed including: i) six nuclear and mtDNA loci and, ii) only six nuclear loci. After preliminary runs to determine the appropriate range of priors, as well as confirming proper MCMC mixing, the final runs were performed with 30 coupled chains with a geometric heating scheme ($g_1 = 0.3$ and $g_2 = 0.9$) and allowed to continue for at least $3x10^6$ steps. The first $8x10^5$ steps were discarded as the burn-in. Because the three types of marker are characterized by different effective population size, I included an *inheritance scalar* to adjust the parameters in the model: 0.25 for mitochondrial, 0.75 for Z-linked and 1.0 for autosomal loci

## RESULTS

### Eco-geographical axis of morphological divergence

Together, PC1 and PC2 of the climate data accounted for 78.42% of the variation that characterized the range of the Brown Scrub-Robin. Factor loadings indicated that PC1 correlates positively with minimum temperature of the coldest month (0.3010) and precipitation of the coldest quarter (0.2953), whereas PC2 correlates positively with mean temperature of the wettest quarter (0.3651) and maximum temperature of the warmest month (0.3271). PC1 reflects a major difference in precipitation seasonality between the cooler highlands of Limpopo and KwaZulu-Natal versus the warmer eastern coastal lowlands, regardless of latitude. PC2 describes a tropical-temperate climate gradient with PC2 scores decreasing towards the southern portion of range (highest latitude; Figure 2).

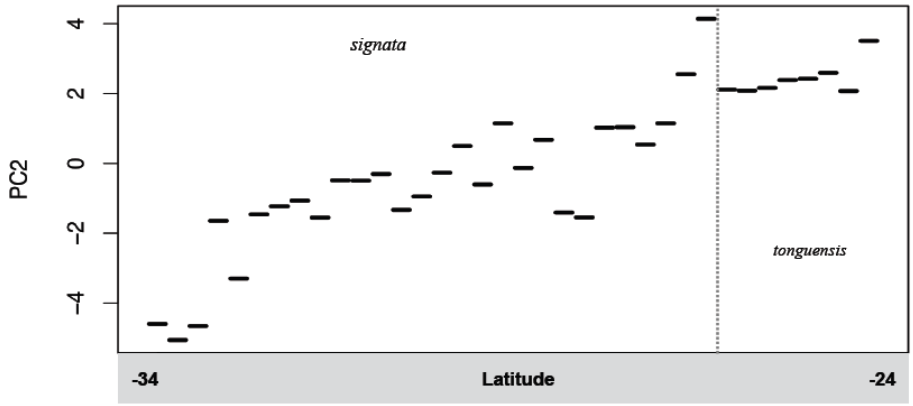


**Figure 2.** Multidimensional summary of climatic variation across the geographical range of the Brown Scrub-Robin as quantified in the second principal component of nineteen climatic variables. PC2 represents the temperate-tropical gradient that runs along the south to north axis of the Brown Scrub-Robin range.

There was a significant effect of sex, i.e. sexual dimorphism, for all morphological variables quantified (mass: $F = 8.025$, $p = 0.006$; wing: $F = 49.072$, $p < 0.001$; tarsus: $F = 17.690$, $p < 0.001$; size: $F = 16.3728$, $p < 0.001$). Furthermore, the ANOVA revealed significant differences in all traits between birds from different geographical areas (Table I). The interaction sex*geography was non-significant for all traits quantified, indicating that females were smaller than males regardless of geographical location. Explicitly testing for differences between ecotypes, I found that *signata* and *tonguensis* differ in mass ($F = 124.167$, $p < 0.001$), wing-length ($F = 34.707$, $p < 0.001$) and structural size ($F = 35.889$, $p < 0.001$), but not in tarsus-length ($F = 1.422$, $p = 0.236$).

**Table I.** Effect of geography and sex on morphological variation of Brown Scrub-Robins as revealed by two-way ANOVA. $N_{females} = 29$, $N_{males} = 52$.

| Variable | Effect | df | *F* |
|---|---|---|---|
| Mass | geography | 11 | $14.175^{***}$ |
| | sex | 1 | $8.026^{**}$ |
| | geography*sex | 10 | 0.593 |
| Wing | geography | 11 | $6.669^{***}$ |
| | sex | 1 | $49.072^{***}$ |
| | geography*sex | 10 | 2.033 |
| Tarsus | geography | 11 | $2.033^{*}$ |
| | sex | 1 | $17.690^{***}$ |
| | geography*sex | 10 | 0.448 |
| Size | geography | 11 | $5.057^{***}$ |
| | sex | 1 | $16.373^{***}$ |
| | geography*sex | 10 | 1.694 |

All morphological variables except tarsus-length were significantly and positively correlated with climate, in a linear relationship (Table II). Despite this significance the correlation with PC1 (precipitation seasonality) was relatively weak and explained little of the variation in the data. The quantitative traits changed concomitantly and were strongly correlated with the temperate-tropical climate gradient quantified by PC2. Including a quadratic factor in the model testing the effect of the temperate-tropical gradient (PC2) explained more of the variability (mass: $R^2 = 0.57$, $p < 0.001$; wing: $R^2 = 0.288$, $p < 0.001$; size: $R^2 = 0.269$, $p < 0.001$) than did the linear model (Table II). This indicates that changes in mass, wing and structural body size rather than changing gradually along the gradient, are affected to a greater extent by extreme conditions and thus best explained by a quadratic function. Birds in the tropical localities of Maputaland (*tonguensis*) were smaller than birds from the subtropical and temperate portion of the range (*signata;* see Annexe Figure C1).

**Table II.** The effect of climate on quantitative traits (morphology). The non-linear relationship between morphology and climate was modelled by including a quadratic term.

| **Models** | **$R^2$** | ***p*** |
|---|---|---|
| Mass ~ PC1 | 0.037 | 0.08 |
| Mass ~ PC1 + $PC1^2$ | 0.149 | <0.01 |
| Mass ~ PC2 | 0.202 | <0.01 |
| Mass ~ PC2 + $PC2^2$ | 0.572 | <0.001 |
| Tarsus ~ PC1 | 0.008 | 0.413 |
| Tarsus ~ PC1 + $PC1^2$ | 0.066 | 0.067 |
| Tarsus ~ PC2 | 0.037 | 0.587 |
| Tarsus ~ PC2 + $PC2^2$ | 0.032 | <0.01 |
| Wing ~ PC1 | 0.065 | 0.02 |
| Wing ~ PC1 + $PC1^2$ | 0.131 | <0.001 |
| Wing ~ PC2 | 0.091 | <0.01 |
| Wing ~ PC2 + $PC2^2$ | 0.288 | <0.001 |
| Size ~ PC1 | 0.058 | 0.03 |
| Size ~ PC1 + $PC1^2$ | 0.083 | 0.03 |
| Size ~ PC2 | 0.097 | <0.01 |
| Size ~ PC2 + $PC2^2$ | 0.269 | <0.001 |

**Eco-geographical axis of genetic divergence**

The geographical projection of climatic niches revealed areas of reciprocal suitability. A portion of the geographical range of *signata* was predicted as climatically suitable for *tonguensis* and vice-versa (Figure 3). The models also predicted a small area (approximate linear distance of 50 km), south of St. Lucia region, where co-occurrence of both ecotypes is unlikely (Figure 3 – co-occurrence). Thus, both ecotypes could co-occur and potentially introgress if able disperse across the 50 km gap. However, the Bayesian clustering analysis of nuclear genotypes revealed two distinct nuclear genomes (LnP (K = 1) = -1043.84 ± 30.78 *vs.* LnP (K = 2) = -663.5 ± 53.72) with strong geographical affinities. Members of cluster I were all from the tropical Maputaland and members of cluster II all from subtropical and temperate areas, with no signal of recent introgression (Figure 4). All individuals with the exception of one were assigned to one of the clusters with a probability of membership higher than 0.980 across the five STRUCTURE runs. Only one bird sampled in the Maputaland region had a mean probability of membership of 0.842 ± 0.023 across five runs.

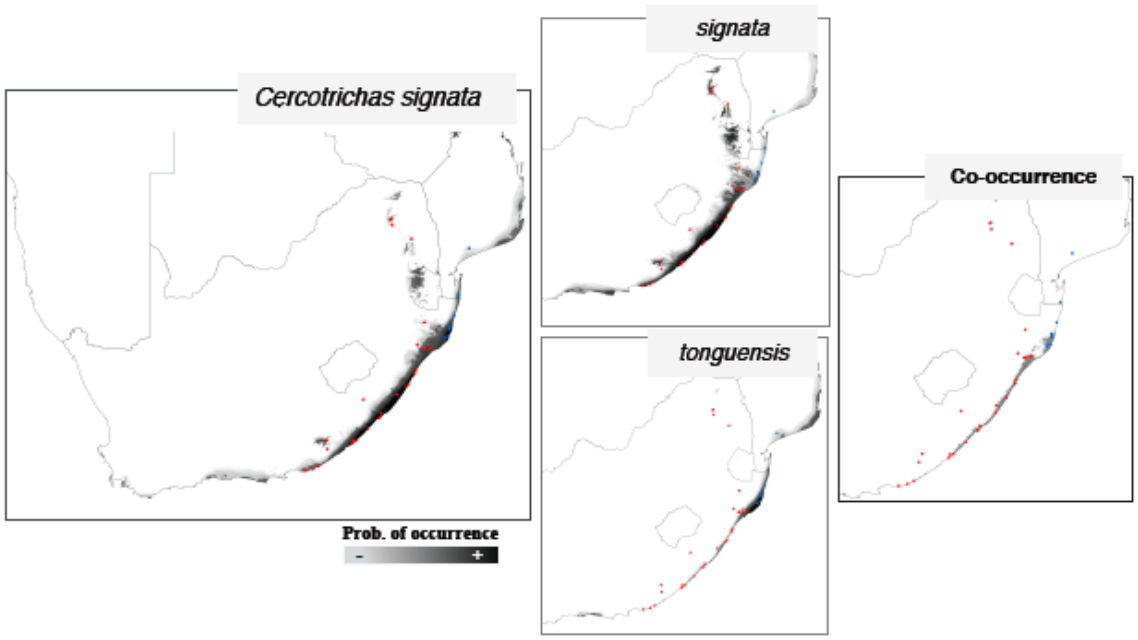


**Figure 3.** Geographical projection of the climatic niche of *Cercotrichas signata* and the two ecotypes: *signata* and *tonguensis*. The niches of both *signata* and *tonguensis* were reciprocally predicted into a small portion of each other ranges – region of co-occurrence. The portion depicted in white represents an area where the climatic niches do not overlap. Red and blue dots indicate point locality data, for the occurrence of ecotypes *signata* and *tonguensis*, respectively.

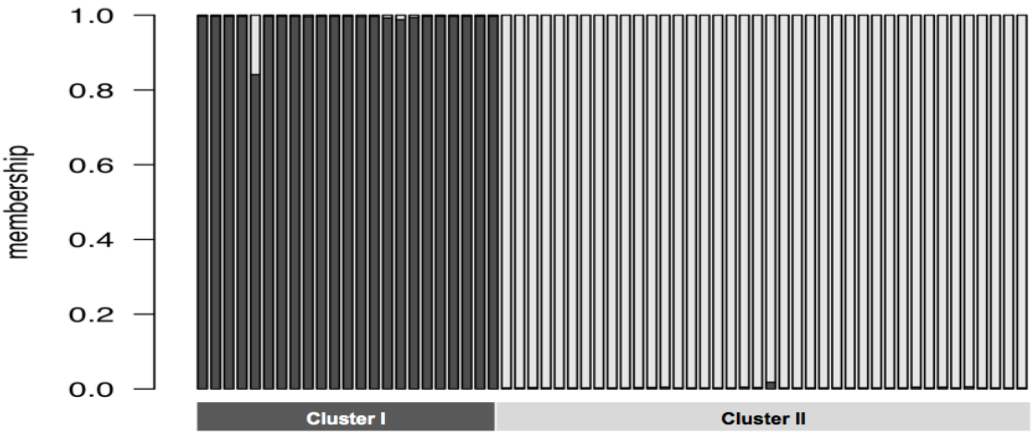


**Figure 4.** Results from Bayesian clustering analysis of genotypic data (six nuclear introns) showing two distinct gene pools (LnP (K=2) = -663.5) corresponding to the two ecotypes: *tonguensis* (cluster I) and *signata* (cluster II).



Using STRUCTURE membership to define two major genetic groups, the AMOVA implemented with the mitochondrial data revealed the same strong geographical structure: the majority of the mitochondrial genetic variation found was between ecotypes (96%; $F_{CT}$ = 0.962, $p < 0.001$) and only a smaller portion among populations within the ecotypes (3.27%; $F_{SC}$ = 0.117, $p$ = 0.008). Because of this strikingly high differentiation between groups living in tropical vs. sub-tropical-temperate forests (44 segregating sites; 4.8% sequence divergence) and the fact that mitochondrial loci examined here underlie a key enzyme from the oxidative phosphorylation pathway (OXPHOS) I translated the DNA sequences to the amino acids. This revealed that of the seven non-synonymous mutations found at the ATP synthase subunit 6, five amino acids segregate between the two nuclear genetic cluster and coincide with the currently defined ecotypes/subspecies (Figure 5).

The coalescent model of *isolation with migration* revealed different trends in the amount of gene flow detected between the two groups of populations compared ($SLR_{tongensis}$ vs. $KZN_{signata}$ ; $KZN_{signata}$ vs. $PSJ_{signata}$), and also that the estimates of gene flow vary in response to whether or not the mtDNA data was included in the analyses. Using the full data set, the migration rate between $SLR_{tongensis}$ and $KZN_{signata}$ was estimated to be zero: $m_{KZN\ -\ SLR}$ = 0.0015 (90% High Posterior Density: 0.0015 – 1.4535), $m_{SLR\ -\ KZN}$ = 0.0015 (90% HPD: 0.0015 – 0.4275). Note that 0.0015 was the first bin of the prior and therefore

corresponds to zero in both directions. When analysing only the nuclear data, estimates of migration were slightly different. Introgression from KZN$_{signata}$ into SLR$_{tongensis}$ was higher than in the opposite direction: $m_{KZN\text{-}SLR}$ = 0.319 (90% HPD: 0.001 - 0.977), $m_{SLR-KZN}$ = 0.001 (90% HPD: 0.001 – 0.1670). However, the 90% high posterior densities included zero (0.001 was the first bin).

In contrast, migration between populations between PSJ$_{signata}$ and KZN$_{signata}$, estimated with full data only, was nearly 13 times larger than between KZN and SLR populations (Table III): $m_{PSJ\text{-}KZN}$ = 13.525 (90% HPD: 5.474 – 30.525), $m_{KZN-PSJ}$ = 0.0075 (90% HPD: 0.0075 – 12.5475). Excluding mtDNA from the data set had a major effect on the directionality of gene flow. Migration southward (from KZN to PSJ) became significantly different from zero: $m_{PSJ\text{-}KZN}$ = 20.475 (90% HPD: 9.275 - 41.475), $m_{KZN-PSJ}$ = 2.783 (90% HPD: 0.203 – 13.403), indicating that nuclear gene flow between these two populations occurs in both directions.



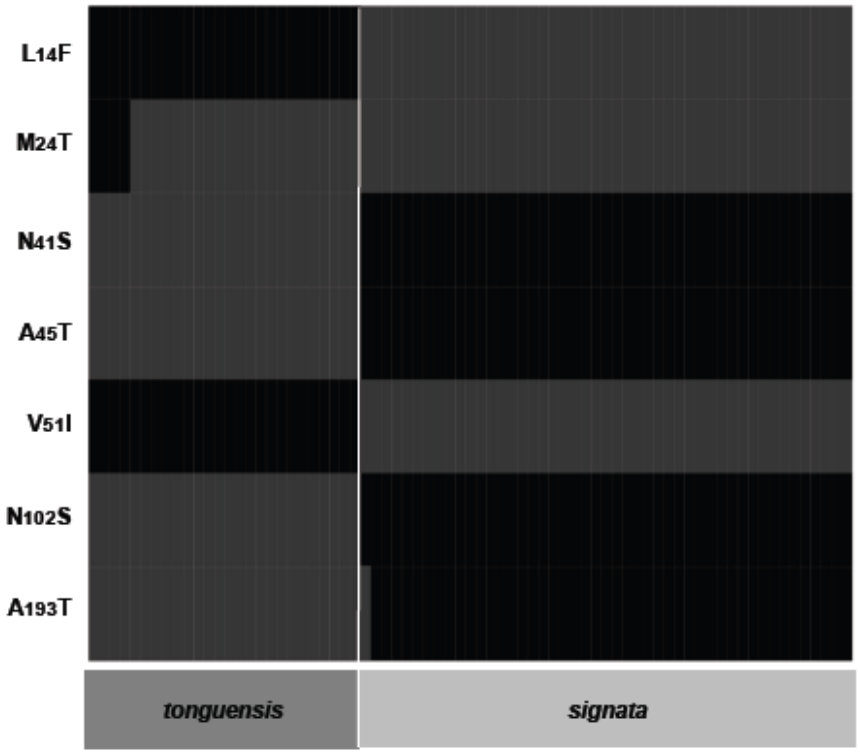


**Figure 5.** Amino acid changes in the ATP synthase subunit 6, which is encoded in the mitochondrial genome. Five, out of the seven, changes segregate in concert with morphology, climate and nuclear genotypes. L: Leucine, F: Phenylalanine, M: Methionine, T: Threonine, N: Aspargine, S: Serine, A: Alanine, V: Valine, I: Isoleucine.

**Table III.** Coalescent estimates of migration (*m*) between: (1) populations in the coastal forests of St. Lucia Region - SLR, and KwaZulu-Natal scarp forest - KZN; and (2) populations in the KZN scarp forests - KZN and Port St Johns coastal scarp forests - PSJ.

| Populations | Analysis | *m* (HiPt) | 90% HPD |
|---|---|---|---|
| **(1)** SLR vs. KZN | **Seven loci** (mtDNA + nuclear DNA) | $m_{SLR - KZN} = 0.0015$ | 0.0015 – 0.4275 |
| | | $m_{KZN - SLR} = 0.0015$ | 0.0015 – 1.4535 |
| | **Six loci** (nuclear DNA) | $m_{SLR - KZN} = 0.0010$ | 0.0010 – 0.1690 |
| | | $m_{KZN - SLR} = 0.3210$ | 0.0010 – 0.9730 |
| **(2)** KZN vs. PSJ | **Seven loci** (mtDNA + nuclear DNA) | $m_{KZN - PSJ} = 0.0075$ | 0.0075 – 12.5475 |
| | | $m_{PSJ - KZN} = 13.525$ | 5.4750 – 30.5250 |
| | **Six loci** (nuclear DNA) | $m_{KZN - PSJ} = 2.7825$ | 0.2025 – 13.4025 |
| | | $m_{PSJ - KZN} = 20.4750$ | 9.2750 – 41.4750 |

HiPt: high posterior density, HPD: high posterior probability



## DISCUSSION

When considering speciation as a continuous process, populations can be found at distinct stages of divergence depending on the interaction between ecological and evolutionary processes. Combining eco-geographical, morphological and genetic data provided insights into the process of divergence in the southern African endemic Brown Scrub-Robin. Two distinct stages of divergence are exemplified in this bird, a species that has had its demographic dynamics altered and its range disrupted, as a consequence of the effect of the Pleistocene climatic oscillations on the habitat.

The concerted sharp transition in morphological traits, nuclear genotypes and mtDNA haplotypes (Figure 6 A-E) in the area of transition between temperate and tropical climates (south of St. Lucia); the unlikely co-occurrence of ecotypes found in the same area; the reciprocal prediction of climatic niches into each other's geographical range (Figure 3); and the lack of gene flow between adjoining populations from the two ecotypes that had occupied remnant forest patches suggests that this area is where natural selection currently changes direction, and thus impedes gene flow between populations. These results indicate that these populations are in a latter stage of the divergence process. However, whether divergence was initiated in different environments with natural selection promoting adaptive divergence and thereby impeding gene flow; or whether divergence started via the reduction of migration and

consequently gene flow between populations in similar environments, after which natural selection increased the frequency of better adapted phenotypes, still remains unanswered.

In contrast, the lack of morphological and genetic divergence, and the extensive gene flow between populations currently occupying forests that were historically isolated (KZN and PSJ) and which occur in distinct portions of the climatic gradient (Figure 6 A-E) revealed that changes in selective pressures altered the population dynamics and promote the reestablishment of migration and hence the reversal of the divergence process.

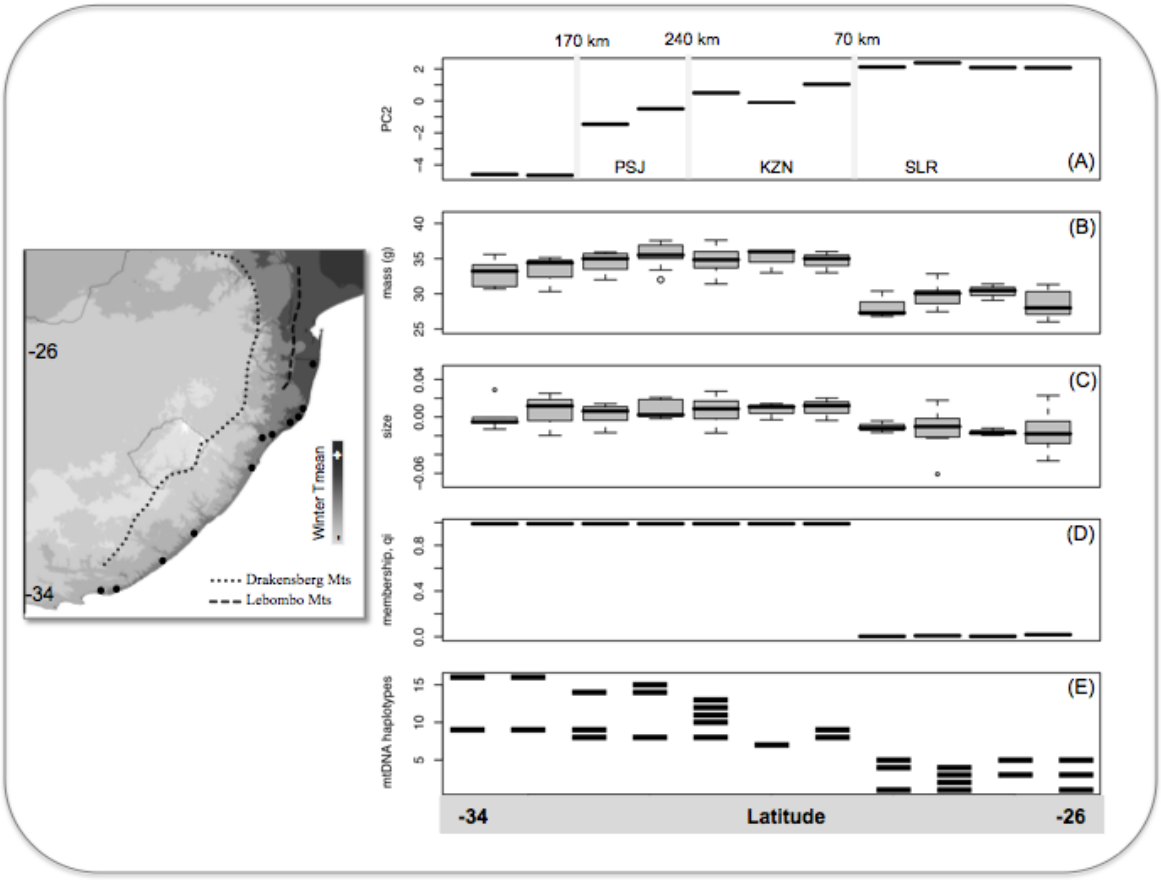


**Figure 6.** Concerted change in climate (A), mass (B), size (residuals of wing-length on tarsus length) (C), nuclear genotype (D) and, mtDNA haplotypes (E) along the temperate to tropical transect (S to N) that crosses the range of the Brown Scrub-Robin in southeastern Africa. Side map depicting the change in mean temperature during winter (WinterTmean = Bio11: mean temperature of the coldest quarter; light grey represents lower temperatures and dark grey higher temperatures) and the location of the two main topographical features: the Drakensberg and Lebombo Mountains.

**Ecological determinants of the divergence process**

Since the early 20th century, climate has been regarded as the primary factor that affects density and the geographical distribution of populations (Grinnell et al. 1914, Gaston 2003). In 1967, Janzen proposed a mechanism to explain how thermal adaptation could lead to speciation. In his conceptual model, populations living in areas where temperature ranges are narrow are expected to evolve greater physiological specialization and thus have reduced ability to disperse across climatically unsuitable regions. The role of selection on thermoregulatory capabilities is also one of the main adaptive mechanisms proposed to explain variation in body size: larger-bodied individuals loose heat more slowly, and thus have higher fitness in cooler climates (Smith et al. 1995). In contrast, smaller body size reduces heat loading which is advantageous in warmer climates (McNab 1979). Also, after body mass, avian basal metabolism is most strongly (and negatively) related to ambient temperature (e.g. White et al. 2007, Wiersma 2007, Jetz et al. 2008). The significant difference in structural body size (wing/tarsus residuals) and body mass found in the Brown Scrub-Robin populations along the climatic axis representing the temperate-tropical gradient, with smaller birds occupying warmer climates, coupled with the prediction of an area of less suitable climatic conditions for *signata,* suggests the role of natural divergent selection on traits involved in physiological adaptation.

Alternatively, morphological divergence can be explained by factors other than abiotic variables. For instance, population demographic dynamics and thus the fate of a phenotype can be determined by inter-individual relationships (Sutherland 1996, e.g. Duckworth and Badyaev 2007). In the social context the process of divergence can be caused by an individual condition and how it affects that individual's relationship with others, particularly with respect of competition for resources, including mates. Under this model, divergence can occur between populations that differ slightly in their non-social environment (abiotic niche) if traits involved in social interactions increase mating success of some individuals (West-Eberhard 1983). The two ecotypes of Brown Scrub-Robin differ in plumage colour and white patterning on the throat, with *signata* being larger and darker with narrower white markings than *tonguensis*. It is therefore possible that a trait (e.g. plumage colouration and pattern, size) involved in social interactions, such as mate attraction or competition with same-sex conspecifics, increased in frequency as gene flow progressively diminished. This sub-component of natural selection can promote assortative mating which in turn may lead to the loss of polymorphism. Song is another trait that could have taken the lead in the process of divergence (Slabbekoorn and Smith 2002, Toews and Irwin 2008) and may be co-varying with morphology. Although the difference between the song structure of the two forms is considered to be subtle (Oatley 1998), the actual quantification of the differences and its experimental manipulation (playback) would allow testing whether song leads the divergence process or rather, if any variation

observed in that trait is a consequence of the process.

**Speciation in the forest of southeast Africa**

In the resumé of the historical symposium *The biogeography of south-east Africa*, Poynton (1961) described a gradual north-to-south impoverishment of tropical flora and fauna along the southeastern coast of Africa (the 'subtropical subtraction effect'), with a major loss of tropical elements occurring in the region of St. Lucia Estuary. The causes of this abrupt loss of biodiversity around St Lucia still puzzle biogeographers because no abrupt geophysical or climatic changes occur in this area. Regarding forest faunal assemblages, Lawes et al. (2007) demonstrated that the forest archipelago in the scarps of northern KwaZulu-Natal, but south of St. Lucia, harbour elements from both temperate and tropical faunal assemblages. The major cause behind this is hypothesized to be the climatic oscillations during the Quaternary. Modelling the distribution of forests in KwaZulu-Natal, Eeley et al. (1999) concluded that forest habitat was much reduced and fragmented as a consequence of lowered temperatures and precipitation during the Last Glacial Maximum. Small remnants of forest might have persisted in the costal scarp of the southern parts of the Eastern Cape near Port St Johns (Hughes et al. 2005), and also along the scarps of northern KwaZulu-Natal (Eeley et al. 1999). With subsequent climatic amelioration, Afromontane forest expanded into the highlands, scarp forest re-established in Maputaland and coastal tropical forest with East African affinities expanded south into Maputaland (Lawes et al. 2007).

The Quaternary climatic oscillations might have disrupted the range of the forest-dwelling Brown Scrub-Robin, causing populations to be eco-geographically segregated. By comparing the coalescent estimates of migration between two pairs of populations ($KZN_{signata}$ vs. $STR_{tongensis}$ and $KZN_{signata}$ vs. $PSJ_{signata}$), that had occupied isolated forest remnants allowed to understand the relative roles of migration, drift and selection in the divergence process. Constrained by their initial genetic background and the lack of gene flow, populations followed distinct evolutionary trajectories. At a given point, this divergence resulted in reproductive isolation between populations inhabiting the scarp forests of KwaZulu-Natal (*signata*) and populations in the coastal Indian Ocean forest of the St. Lucia region (*tonguensis*). Regardless of the data set analysed (morphology, mtDNA, nuclear loci), populations in the scarp forests of KwaZulu-Natal seem to have diverged from populations in the Santa Lucia region without exchanging effective migrants, through selection and drift. This suggests that the initial fragmentation of these populations was driven by the appearance of a set of conditions that disrupted the range and progressively diminished the movement of individuals. In contrast, whilst selection and drift might have occurred during the putative isolation of populations occupying the coastal scarps forest of the Eastern Cape (*signata*) and the scarp forests of KwaZulu-Natal, the effect was temporary. When selective pressures changed, post

LGM climatic amelioration, the process of divergence was reversed and nuclear gene flow resumed in both directions between these two populations. However, introgression of he mitochondrial genome seems to be asymmetrical, as suggested by the results obtained when including mtDNA data in the analysis (i.e., zero migration southward). This last result may underscore the role of the mitochondrial genome in facilitating adaptation to local conditions, either climatic or food availability (reviewed by Das 2006, Gershoni et al. 2009). Brown Scrub-Robins from the temperate forest might have a broader climatic niche and thus tolerate subtropical conditions, whereas the birds from sub-tropical scarp forest of KZN might have adapted to a narrow set of conditions and thus despite neutral gene flow, natural selection would counter-balance migration and reduce gene flow along the tropical-temperate gradient.

Whether the cessation of gene flow between populations in the transition between the subtropical (KZN scarp forest) and tropical (Santa Lucia coastal forest) climate was the cause or the consequence of environmental-driven and/or social selection is arguable. The fact that morphological differences exist between ecotypes does not prove that these were the traits involved in the divergence process. However, it certainly suggests these are currently co-varying with the causal traits to maintain ecotypes reproductively isolated. Only by concentrating future work in the archipelago of small forest patches at the boundary between the two ecotypes will we be fully able to understand if reproductive isolation is a consequence of divergent selection on physiological tolerances and/or sexual selection on plumage, size or song.

## ANNEXE C

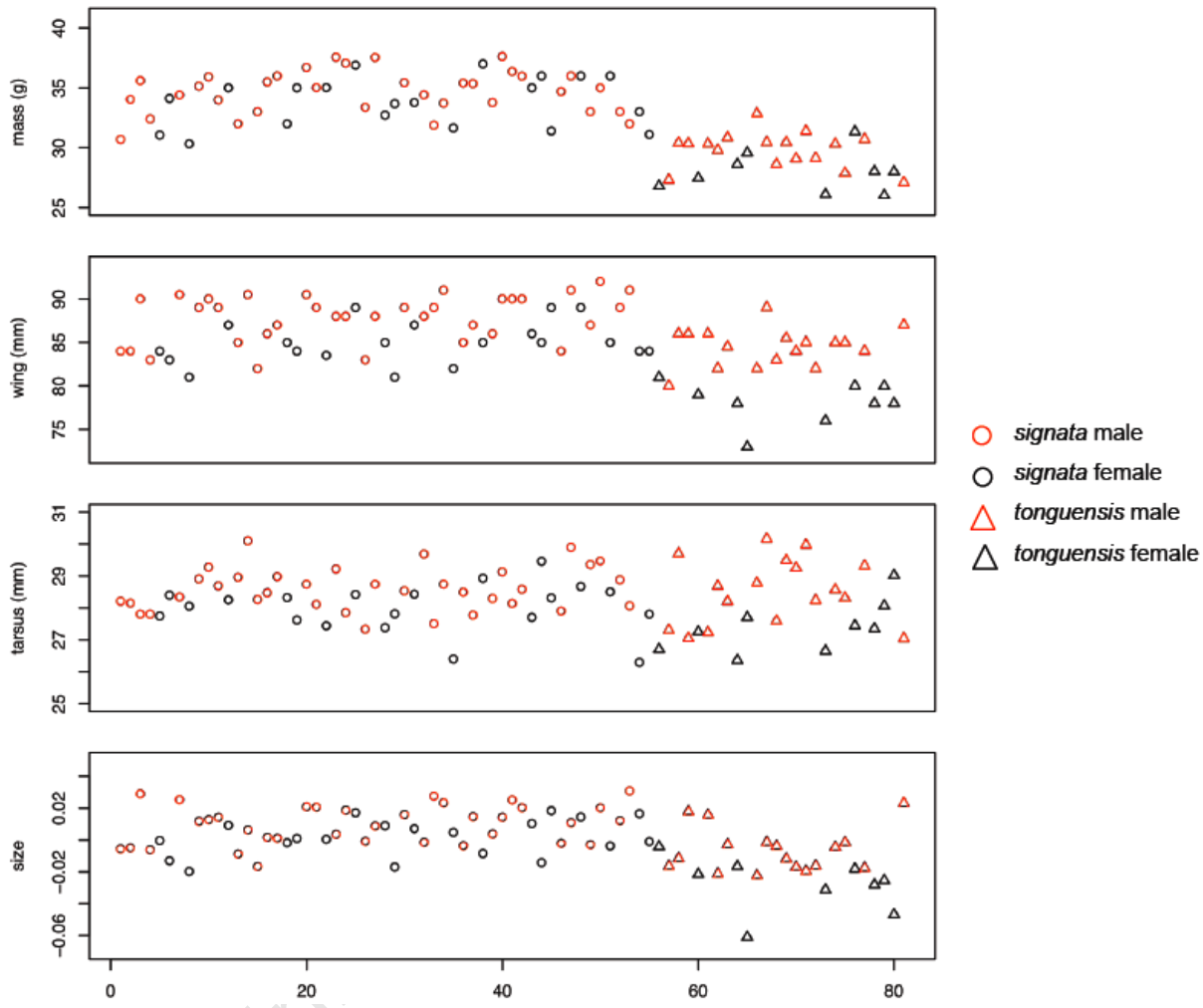


**Figure C1.** Dimorphism in quantitative traits measured across in the Brown Scrub-Robin range. Triangles represent *tonguensis* and circles *signata*. Sex is colour coded.

# Conclusion

Similarly to other areas of science, the intellectual contributions in evolutionary biology tend to become dichotomized. This is particularly true with respect to perspectives on the role of ecology in the process of divergence and formation of species (see Schluter 2009). It is quite difficult, however, to conceive of a natural scenario where factors that constrain the demographic dynamics of populations are absent. Even in the traditional models of geographical speciation (Mayr 1942, 1963), the disruption of a species range may be interpreted as a result of changes in population dynamics due to the inability of individuals to keep up with changes in the environment.

In the last decade, several studies have once again highlighted the importance of considering the ecological context in the continuous process of population divergence and speciation (Rundel and Nosil 2005, Nosil 2008, 2009b). Studies of several organisms are currently considered textbook examples of segregation into alternative ecological niches at different stages of the speciation process (e.g. *Geospiza* Darwin's Finches, *Pyrenestes* African Seedcrackers, *Gasterosteus* sticklebacks, *Anolis* lizards, *Tinema* walking-sticks, *Mimulus* Monkeyflowers). Inspired by these and other integrative research programs, I have tried to tread new ground with respect to mechanisms underlying divergence and speciation in the understudied southern African subcontinent, with a particular focus on the arid-zone.



## CONCEPTUAL CONTRIBUTIONS FROM THE SOUTHERN AFRICAN AVIFAUNA

### Dispersal and gene flow

Dispersal has multiple ecological and evolutionary consequences (Clobert et al. 2001, Kokko and López-Sepulcre, 2006). In the Karoo Scrub-Robin, the integration of behavioural ecology and genetic data at a micro-geographical scale revealed that the social context has consequences on the arrangement of kin, determines the identity of dispersers and the distance moved. These results provided insights into the role of ecological and evolutionary causes of dispersal. However, studying the evolution of dispersal is not only complicated by the multiplicity of benefits and costs, but also by the fact that they vary with population dynamics and spatial scale of the study (Ronce et al. 2007; Nathan et al. 2008). The population studied here had a high density and was located in the most mesic portion of the range (highest productivity and predictability) at the current limit of the distribution. Understanding how sociality and spatial variation in habitat features influence dispersal and its evolution would be an interesting line of research. An experimental design that would replicate the current study along the aridity gradient that crosses the range of the Karoo Scrub-Robin, as well as including populations from different locations

along the edge of the range would provide crucial knowledge about the plasticity of dispersal behaviour and how it affects the sociality and meta-population dynamics.

Individuals might respond differently to ecological constrains if a given variant of a trait facilitates dispersal. Empirical studies have found that a pattern of spatial variation in selection on dispersal-associated traits is common: individuals with higher dispersal capabilities tend to be at greater frequency near the edge of the expanding range (e.g. Cane Toad *Bufo marinus* Phillips et al. 2006; Western Bluebird *Sialia mexicana* Duckworth 2008). The geographical variation in population dynamics can feedback and alter the frequency of adaptive traits that in turn will indirectly result in selection of dispersal behaviour and facilitate a range expansion, as illustrated in the Karoo scrub-robin.

**Adaptation and speciation**

The two *Cercotrichas* species studied are members of the same clade (revealed by molecular phylogeny) that occupy opposite ends of the habitat spectrum that can be found in the southern Africa subcontinent: the Karoo Scrub-Robin *C. coryphaeus* occurs in the semi-arid Karoo and Fynbos and, the Brown Scrub-Robin *C. signata* in endemic to the evergreen forest. The ecotype/subspecies pairs within each species are at different stages of divergence. The taxa *C. s. signata* and *C. s. tonguensis* exhibited increased genotypic clustering, lower neutral gene flow, greater phenotypic divergence and higher levels of reproductive isolation than the *C. c. coryphaeus* and *C. c. cinerea.* This indicates of a later stage in the speciation process within the Brown Scrub-Robin and only a partial progress towards speciation, i.e. adaptation to local conditions, in the Karoo Scrub-Robin. The historical changes in eco-geographical context that characterized the evolution of the species seem to be relevant and indicate that selection on multiple dimensions of the niche results in greater divergence.

The finding that the mitochondrial genome might have evolved adaptively in response to selective pressures imposed by new climatic conditions, coupled with the role of mitochondrial genes in the regulation of energetic metabolism suggests that birds living in harsh habitats might evolve local physiological mechanisms to allow them to meet their daily energetic requirements. In the case of the Karoo Scrub-Robin, testing the relationship between adaptive genetic variation and phenotypic function of OXPHOS genes in a physiological mechanistic framework would be an exciting line of future research.

Recent compelling experimental studies have demonstrated mito-nuclear incompatibilities in an intertidal Copepode (Ellison and Burton 2006), the Stonechat (Tielleman et al. 2009) and the Seed Beetle (Arnqvist et al. 2010), thus providing evidence for the relevance of metabolic pathways in the process of adaptive divergence. Natural selection counteracts migration and drift to impede the breakdown of the cyto-nuclear genetic network involved in cellular energetics, maintaining populations adapted to different

environments and leading the divergence process.

The *genomics era* puts us in an unprecedented position to unravel the genetic basis of adaptation. Moreover, exploring the phenotype-genotype nexus by coupling experimental ecological data with genomic data would provide insights on how passerine birds, characterized by their diurnal habits, high body temperature and metabolic rates, adapted to live in the harsh and unpredictable xeric habitats of southern Africa.

Going even further, the functional role of the mitochondrial genome and the co-evolution with the nuclear genome deserves further investigation as a possible physiological mechanism that can allow individuals to survive local thermal conditions and/or food availability and hence may initiate the process of divergence and lead to speciation in *Cercotrichas* Scrub-Robins.

I close by re-calling that natural selection, the process that results in the change of frequency of a variant in a polymorphic trait in a population, is an empty concept regarding the function of the trait and the identity of the individuals that have higher reproductive success. As such, adding more information to the premises used to define the process, can lead us to sub-aspects of natural selection. In this dissertation, I focused primarily on traits that are potentially associated with environmental constrains, not because I consider other factors less relevant, but because those constrains are likely to be the initiators of the divergence process. Other aspects of natural selection are equally important. For instance, sexual selection, when the trait expressed in males leads to non-random mating due to female preferences, or social selection if the trait is associated with competition with conspecifics to gain access to resources (other than mates). Therefore, considering the complex dynamics among the different aspects of natural selection and gene flow with empirical manipulations of dispersal (i.e. translocations), although challenging, would be a significant contribution to gain an understanding of the mechanisms by which gene flow can constrain or promote adaptive divergence owing to traits important in habitat matching, social competition and sexual interactions.

# REFERENCES


## A

Abzhanov A., Protas M., Grant B.R., Grant P.R., Tabin C.J. 2004. Bmp4 and morphological variation of beaks in Darwin's finches. *Science*, 305: 1462 – 1465.

Abzhanov A., Kuo W. P., Hartmann C., Grant B. R., Grant P. R., Tabin C. J. 2006. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature,* 442: 563 – 567.

Allender C.J., Seehausen O., Knight M.E., Turner G.F., Maclean N. 2003. Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. *Proceeding of the National Academy of Sciences USA*, 100: 14074 – 14079.

Angert A.L., Bradshaw H.D., Schemske D.W. 2008. Using experimental evolution to investigate geographic range limits in monkey-lowers. *Evolution,* 62: 2660 – 75.

Arnqvist G., Dowling D.K., Eady P., Gay L., Tregenza T., Tuda M., Hosken D.J. 2010. Genetic architecture of metabolic rate: environment specific epistasis between mitochondrial and nuclear genes in an insect. *Evolution*, 64: 3354 – 3363.

## B

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

Ballard J., and Melvin R. 2010. Linking mitochondrial genotype with organismal phenotype. *Molecular Ecology*, 119: 1523 – 1539.

Balloux F., Handley L.J.L., Jombart T., Liu H., Manica A. 2009. Climate shaped the worldwide distribution of human mitochondrial DNA sequence variation. *Proceedings of the Royal. Society B*, 276: 3447 - 3455.

Beaumont M.A., Zhang W., Balding D.J. 2002. Approximate Bayesian computation in population genetics. *Genetics*, 162: 2025 – 2035.

Bradshaw H.D., and Schemske D.W. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkey-flowers. *Nature*, 426: 176–178.

Beaumont M., and Nichols R. 1996. Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B,* 263: 1619 - 1626.

Bengtsson B.O. 1978. Avoiding inbreeding: at what cost? *Journal of Theoretical Biology*, 73: 439 – 444.

Bicudo J.E.P.W., Buttemer W.A. Chappell M.A., Pearson J.T., Bech C. 2010. Ecological and Environmental Physiology of Birds. Oxford University Press, Oxford, UK.

Blomqvist D., Andersson M., Kupper C., Cuthill I.C., Kis J., Lanctot R.B., Sandercock B.K., Szekely T., Wallander J., Kempenaers B. 2002. Genetic similarity between mates and extra-pair parentage in three species of shorebirds. *Nature*, 419: 613 – 615.

Bottema C.D.K., Sarkar G., Cassady J.D., Ii S., Duttonand C.M., Sommer S.S. 1993. Polymerase chain-reaction amplification of specific alleles: a general method of detection of mutations, polymorphisms, and haplotypes. *Methods in Enzymology*, 218: 388 – 402.

Bowler D.E, and Benton T.G. 2005. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, 80: 205 – 225.

Bruen, T.C., Philippe H., Bryant D. 2006. A simple and robust statistical test detecting the presence of recombination. *Genetics*, 172: 2665 – 2681.

Bulmer M.G. 1972. Multiple niche polymorphism. *American Naturalist*, 106: 254 – 257.

Burt D.W., and Paton I.R. 1991. Molecular cloning and primary structure of the chicken transforming

growth factor-beta2 gene. *DNA and Cell Biology*, 10: 723 – 734.

Butlin R.K., Galindo J., Grahame J.W. 2008. Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philosophical Transactions of the Royal Society B*, 363: 2997 – 3007.

**C**

Chase B.M., and Meadows M.E. 2007. Late Quaternary dynamics of southern Africa's winter rainfall zone. *Earth Science Reviews*, 84: 103 – 138.

Cheviron Z.A., and Brumfield R.T. 2009. Migration-selection balance and local adaptation of mitochondrial haplotypes in Rufous-collared Sparrows (*Zonotrichia capensis*) along an elevational gradient. *Evolution*, 63: 1593 – 1605.

Clobert J., Danchin E., Dhondt A.A., Nichols JD. 2001. Dispersal. Oxford University Press, New York, USA.

Clutton-Brock T.H. 1989. Female transfer and inbreeding avoidance in mammals. *Nature*, 337: 70 – 71.

Cockburn A., Osmond H.L., Mulder R.A., Double M.C., Green D.J. 2008. Demography of male reproductive queues in cooperatively breeding superb fairy-wrens (*Malurus cyaneus)*. *Journal of Animal Ecology*, 77: 297 – 304.

Collar N. J. 2005. Family Turdidae (Thrushes). *In* Handbook of the Birds of the World. Vol. 10. Cuckoo-shrikes to Thrushes. J. del Hoyo, A. Elliot and D. A. Christie., eds., pp. 514 – 807, Lynx Edicions, Barcelona, Spain.

Cornuet J.M., Santos F., Beaumont M.A., Robert C.P., Marin J-M., Balding D.J., Guillemaud T., Estoup A. 2008. Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, 24: 2718–2719.

Costantini C., Ayala D., Guelbeogo W., Pombi M., Some, C., Bassole I., Ose K., Fotsing J-M., Sagnon N.F., Fontenille, D., Besansky N., Simard F. 2009. Living at the edge: biogeographic patterns of habitat

segregation conform to speciation by niche expansion in *Anopheles gambiae*. *BMC Ecology,* 9: 16.

Covas R., and Griesser M. 2007. Life history and the evolution of family living in birds. *Proceeding of the Royal Society B,* 274: 1349 – 1357.

Cowling RM, Richardson DM, Mustart PJ. 1997. Fynbos. *In* Vegetation of Southern Africa, R.M. Cowling, D.M. Richardson and S.M. Pierce, eds., Cambridge University Press, Cambridge, UK.

Cowling R.M., and C. Hilton-Taylor. 1997. Phytogeography, flora and endemism. *In* Vegetation of Southern Africa, R.M. Cowling, D.M., Richardson and S.M. Pierce, eds., Cambridge University Press, Cambridge, UK.

Cowling R.M., Cartwright C.R., Parkington J.E., Allsopp J.C. 1999. Fossil wood charcoal assemblages from Elands Bay Cave, South Africa: implications for Late Quaternary vegetation and climates in the winter-rainfall fynbos biome. *Journal of Biogeography*, 26: 367–378.

Coyne J.A., and Orr H.A. 2004. Speciation. Sinauer Associates Inc., Suderland, USA.

Crispo E., Bentzen P., Reznick D.N., Kinnison M.T., Hendry A.P. 2006. The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecol*ogy, 15:49 – 62

**D**

Darwin C. 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London, UK.

Das J. 2006. The role of mitochondrial respiration in physiological and evolutionary adaptation. BioEssays, 28: 890 – 901.

Day R.L., Laland K.N., and Odling-Smee J. 2003. Rethinking Adaptation: The Niche Construction Perspective. *Perspectives in Biology and Medicine,* 46: 80 – 95.

Dean W.R.J., Milton S.J., Watkeys M.K., Hockey P.A.R. 1991. Distribution, habitat preference and

conservation of the Red Lark *Certhilauda burra*. *Biological Conservation,* 58: 257-274.

Dean W.R.J., and Milton S.J. 2004. The Karoo: Ecological Patterns and Processes. Cambridge University Press, Cambridge, UK.

Desmet P.G., and Cowling R.M. 1999. Biodiversity, habitat and range-size aspects of a flora from a winter-rainfall desert in north-western Namaqualand, South Africa. *Plant Ecology*, 142:23 - 33.

Dieckmann U., Doebeli M., Metz J.A.J., Tautz D., eds. 2004. Adaptive speciation. Cambridge University Press, Cambridge, UK.

Dobson F.S. 1982. Competition for mates and predominant juvenile male dispersal in mammals. *Animal Behaviour,* 30: 1183 – 1192.

Dobzhansky T. 1951. Genetics and the Origin of Species. 3rd edition. Columbia University Press, New York, USA.

Doebeli M., and Dieckmann U. 2003. Speciation along environmental gradients. *Nature*, 421: 254 –264.

Doerr E.D., Doerr V.A.J. 2006. Comparative demography of treecreepers: evaluating hypotheses for the evolution and maintenance of cooperative breeding. *Animal Behavior*, 72: 147 – 159.

Double M.C., Peakall R., Beck N.R., Cockburn A. 2005. Dispersal, philopatry and infidelity: dissecting local genetic structure in superb fairy-wrens (*Malurus cyaneus*). *Evolution*, 59: 625 – 635.

Duckworth R.A., and Badyaev A.V. 2007. Coupling of dispersal and aggression facilitates the rapid range expansion of a passerine bird. *Proceedings of the National Academy of Sciences USA*, 104: 15017 – 15022.

Duckworth R.A. 2008. Adaptive dispersal strategies and the dynamics of a range expansion. *American Naturalist,* 172: S4 – S17.

**E**

Eckert C.G., Samis K.E., Lougheed S.C. 2008. Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology*, 17: 1170 – 88.

Eeley H.A.C., Lawes M.J., Piper S.E. 1999. The influence of climate change on the distribution of indigenous forest in KwaZulu-Natal, South Africa. *Journal of Biogeography,* 26: 595 – 617.

Ehrlich P.R., and Raven P.H. 1969. Differentiation of populations. *Science*, 165: 1228 – 1232.

Ekman J., Dickinson J.L., Hatchwell B.J., Griesser M. 2004. Delayed dispersal. *In* Cooperative breeding in birds. W.D. Koenig and J.L. Dickinson, eds., Cambridge University Press. Cambridge, UK.

Ellegren H. 2007. Molecular evolutionary genomics of birds. *Cytogenetic and Genome Research*, 117: 120 – 130.

Ellison C.K., and Burton R.S. 2006. Disruption of mitochondrial function in interpopulation hybrids of *Tigriopus californicus*. *Evolution*, 60: 1382 – 1391.

Endler J. 1977. Geographic variation, speciation and clines. Princeton University Press, Princeton, USA.

Endler J. 1986. Natural selection in the wild. Princeton University Press, Princeton, USA.

Epperson B.K. 2003. Geographical genetics. Princeton University Press. Princeton, USA.

Erasmus B.F.N., Van Jaarsveld A.S., Chown S.L., Kshatriya M.,Wessels K.J., 2002. Vulnerability of South African animal taxa to climate change. *Global Change Biology*, 8: 679 – 693.

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

Excoffier L. 2004. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology*, 13: 853–64.

Excoffier L., Laval G., Schneider S. 2005. Arlequin ver. 3.1: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1: 47 - 50.

Excoffier L., and Ray N. 2008. Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology and Evolution*, 23: 347 - 351.

Excoffier L., Foll M., Petit R.J. 2009. Genetic consequences of range expansions. *Annual Review of Ecology, Evolution and Systematics*, 40: 481–501.

**F**

Feder, J.L. 1998. The apple maggot fly, *Rhagoletis pomonella*: flies in the face of conventional wisdom about speciation? *In* Endless Forms: Species and Speciation. D. J. Howard and S.H. Berlocher, eds., Oxford University Press, New York, USA.

Feder, J.L., Berlocher, S.H., Roethele, J.B., Dambroski, H., Smith, J.J., Perry, W.L., Gavrilovic V., Filchak K.E., Rull J., Aluja M. 2003 Allopatric genetic origins for sympatric host plant shifts and race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences USA*, 100: 10314 – 10319.

Fisher R.A. 1930. The genetical theory of natural selection. Oxford University Press, Oxford, UK.

Fitzpatrick B.M., Fordyce J.A., Gavrilets S. 2009 Patterns, processes and geographic modes of speciation. *Journal of Evolutionary Biology,* 22: 2342 – 2347.

Fridolfsson A-K., and Ellegren H. 1999. A simple and universal method for molecular sexing of non-ratites birds. *Journal of Avian Biology*, 30: 116 - 121.

Friesen, V.L., Congdon B.C., Walsh H.E., Birt T.P. 1997. Intron variation in Marbled Murrelets detected using analyses of single-stranded conformation polymorphisms. *Molecular Ecology* 6: 1047 – 1058.

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

Fu Y.X., and W.H. Li. 1993. Statistical tests of neutrality of mutations. *Genetics*, 133: 693 – 709.

**G**

Garant D., Forde S.E., Hendry A.P. 2007. The multifarious effects of dispersal and gene flow. *Functional Ecology,* 21: 434 – 443.

Garrick R.C., Sunnucks P., Dyer R.J. 2010. Nuclear gene phylogeography using PHASE: dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation. *BMC Evolutionary Biology*, 10: 1 – 17.

Gaston K. 2003. The Structure and Dynamics of Geographic Ranges. Oxford UniversityPress, Oxford, UK.

Gaston K., Chown S., Evans K. 2007. Ecogeographical rules: elements of a synthesis. *Journal of Biogeography*, 35: 483 – 500.

Gavrilets S., and Losos J. B. 2009. Adaptive radiation: contrasting theory with data. *Science,* 323: 732 – 737.

Gerbault P., Liebert A., Itan Y., Powell A., Currat M., Burger J., Swallow D.M., Thomas M.G. 2011 Evolution of lactase persistence: an example of human niche construction. *Philosophical Transactions of the Royal Society B*, 366: 863 – 877.

Gershoni M., Templeton A.R., Mishmar D. 2009. Mitochondrial bioenergetics as a major motive force of speciation. *BioEssays*, 31: 642 – 650.

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

Goldberg E.E., and Lande R. 2007. Species' borders and dispersal barriers. *American Naturalist*, 170: 297 – 304.

Goodwin G.H., 1997. Isolation of cDNAs encoding chicken homologues of the yeast SNF2 and Drosophila Brahma proteins. *Gene,* 184: 27 - 32.

Goudet J. 2002. FSTAT, a program to estimate and test gene diversities and fixation indices (v2.9.3.2). Institute of Ecology, Biology Building, UNIL, Lausanne, Switzerland.

Grant P., and Grant.R. 2008. How and Why Species Multiply: The Radiation of Darwin's Finches, Princeton University Press, New Jersey, USA.

Grahame J., Wilding C.S., Butlin R.K. 2006. Adaptation to a steep environmental gradient and an associated barrier to gene exchange in *Littorina saxatilis*. *Evolution*, 60: 268 – 278.

Greenwood P.J., and Harvey, P.H. 1982. The natal and breeding dispersal of birds. *Annual Review of Ecology and Systematics,* 13: 1 - 21.

Greenwood P.J. 1980. Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour,* 28: 1140 – 1162.

Grinnell J. 1914. Barriers to distribution as regards birds and mammals. *American Naturalist,* 48: 248 – 245.

Grinnell J .1917. The niche relationship of the California thrasher. *The Auk,* 34: 427 – 433

Grinnell J. 1922 The role of the 'Accidental'. *The Auk*, 39: 373-380.

Grinnell J. 1924. Geography and Evolution. *Ecology*, 5: 225 – 229.

Guo S.W., and Thompson E.A. 1992. Performing the exact test of Hardy–Weinberg proportions for multiple alleles. *Biometrics*, 48: 361 – 372.

**H**

Hamilton W.D., and May R.M. 1977. Dispersal in stable habitats. *Nature,* 269: 578 – 81

Handley L.J., and Perrin N. 2007. Advances in our understanding of mammalian sex-biased dispersal. *Molecular Ecology*, 16: 1559 – 1578.

Hanski I., and Gilpin M.E. 1997. Metapopulation biology: ecology, genetics and evolution. Academic Press, San Diego, USA.

Harrigan R.J., Mazza M.E., Sorenson M.D. 2008. Computation vs. cloning: evaluation of two methods for haplotype determination. *Molecular Ecology Resources,* 8: 1239–1248.

Harris T.R., Caillaud D., Chapman C.A., Vigilant L. 2009. Neither genetic nor observational data alone are sufficient for understanding sex-biased dispersal in a social-group-living species. *Molecular Ecology*, 18: 1777-90.

Herron M.D., Waterman J.M., Parkinson C.L. 2005. Phylogeny and historical biogeography of African ground squirrels: the role of climate change in the evolution of *Xerus*. *Molecular Ecology*, 14: 2773 – 2788.

Hewitt G.M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B,* 1442: 183–195.

Hijmans R.J., Guarino J., Bussink C., Mathur P. , Cruz M., Barrentes I., Rojas E. 2004. DIVA-GIS ver. 4.0. A geographic information system for the analysis of species distribution data. *Bioinformatics*, 19: 2496 – 2497.

Hijmans, R.J., Cameron S.E., Parra J.L., Jones P.G., Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25: 1965–1978.

Hochachka P.W., and Somero G.N. 2002. Biochemical adaptation: mechanism and process in physiological evolution. Oxford University Press, Oxford, UK.

Hoekstra H.E., Hirschmann R.J., Bundey R.A., Insel P.A., Crossland J.P. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science*, 313: 101 – 104.

Hoffman J.I., and Amos W. 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology*, 14: 599 – 612.

Hubisz M.J., Falush D., Stephens M., Pritchard J.K. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9: 1322 - 1332.

Hudson R.R., Slatkin M., Maddison W.P. 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics*, 132: 583 – 589.

Hudson R.R., and Kaplan N.L. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, 111: 147 – 164.

Huson D.H., and Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23: 254 – 267.

**J**

Janzen D.H. 1967. Why mountain passes are higher in the tropics. *American Naturalist,* 101: 233 – 249.

Jetz W., Freckleton R.P., McKechnie A.E. 2008. Environment, Migratory Tendency, Phylogeny and Basal Metabolic Rate in Birds. PLoS ONE 3(9): e3261

**K**

Kalinowski S.T., Taper M.L., Marshall T.C. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 16: 1099 – 1106.

Kawecki, T.J., and Ebert, D. 2004. Conceptual issues in local adaptation. *Ecology Letters*, 7: 1225 – 1241.

Kawecki TJ. 2008. Adaptation to marginal habitats. *Annual Review of Ecology, Evolution and Systematics*, 39: 321–424.

Kim Y., and Gulisija D. 2010. Signatures of recent directional selection under different models of population expansion during colonization of new selective environments. *Genetics*, 184: 571 - 85.

Kimball R.T., Braun E.L., Barker F.K., Bowie R.C.K., Braun M.J., Chojnowski J.L., Hackett S.J., Han K.-L., Harshman J., Heimer-Torres V., Holznagel W., Huddleston C.J., Marks B.D., Miglia K.J., Moore W.S., Reddy S., Sheldon F.H., Smith J.V., Witt C.C., Yuri T. 2009. A well-tested set of primers to amplify regions spread across the avian genome. *Molecular Phylogenetics and Evolution*, 50: 654 – 660.

Kimura M. 1956. A model of genetic system which leads to closer linkage by natural selection. *Evolution*, 10: 278 – 287.

Kimura M. 1969. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, 61: 893–903.

Kimura M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, UK.

Kirkpatrick M., and Ravigné V. 2002. Speciation by natural and sexual selection: models and experiments. *American Naturalist*, 158: S22 - S35.

Kirkpatrick M., and Barton N.H. 1997. Evolution of a species' range. *American Naturalist* 150: 1 – 23.

Koenig W.D., and Dickinson J.L. 2004. Introduction. *In* Ecology and evolution of cooperative breeding in birds, W.D. Koenig, J.L. Dickinson, eds., Cambridge University Press, Cambridge, UK.

Koenig W.D., and Haydock J. 2004. Incest and incest avoidance. *In* Cooperative breeding in birds, W.D. Koenig and J.L. Dickinson, eds., Cambridge University Press, Cambridge, UK.

Kokko H., and López-Sepulcre A. 2006. From individual dispersal to species ranges: Perspectives for a changing world. *Science*, 313: 789 – 79.

Kokko H., and López-Sepulcre A. 2007. The ecogenetic link between demography and evolution: can we bridge the gap between theory and data? *Ecology letters*, 10: 773 – 782.

Komdeur J., and Hatchwell B.J. 1999. Kin recognition: function and mechanism in avian societies. *Trends in Ecology and Evolution,* 14: 237 – 241

Koppen W. 1936. Das geographisca System der Klimate, *In* Handbuch der Klimatologie, Koppen W. and Geiger G., eds., Borntraeger, Berlin, Germany.

Kryazhimskiy S., and Plotkin J.B. 2008. The Population Genetics of dN/dS. *PLoS Genetics*, 4: e1000304.

**L**

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

Lee J-W., Lee Y-K., Hatchwell B.J. 2010. Natal dispersal and philopatry in a group-living but non-cooperative passerine bird, the vinous-throated parrotbill. *Animal Behaviour*, 79: 1017 - 1023.

Lenormand T. 2002. Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, 117: 183 – 189.

Lessa E.P., Cook J.A., Patton J.L. 2003. Genetic footprints of demographic expansion in North America, but not Amazonia, during the late Quaternary. *Proceeding of the National Academy of Sciences USA,* 100: 10331 – 10334.

Levene H. 1953. Genetic equilibrium when more than one ecological niche is available. *American Naturalist*, 87: 331 – 333.

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

Lillywhite H.B., and Navas C.A. 2006. Animals, energy and water in extreme environments: Perspectives from Ithala 2004. *Physiological and Biochemical Zoology,* 79: 265 - 273.

Linnen C.R., Kingsley E.P., Jensen J.D., Hoekstra H.E. 2009. On the origin and spread of an adaptive allele in deer mice. *Science*, 325: 1095 – 1098.

Lloyd P., Taylor W.A., du Plessis M.A., Martin T.E. 2009. Females increase reproductive investment in response to helper-mediated improvements in allo-feeding, nest survival, nestling provisioning and post-fledging survival in the Karoo scrub-robin *Cercotrichas coryphaeus*. *Journal of Avian Biology*, 40: 400 - 411.

**M**

McNab B.K. 1979. The influence of body size on the energetics and distribution of fossorial burrowing mammals. *Ecology*, 60: 1010 – 1021.

Malécot G. 1967. Identical loci and relationship. *In* Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, L.M. Lecam and J. Neyman, eds., California University Press, Berkeley, USA.

Mallet J. 1995. A species definition for the modern synthesis. *Trends in Ecology and Evolution,* 10: 294 – 299.

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

Mantel N.A. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27: 209 - 220.

Matthee C.A., and Flemming A.F. 2002. Population fragmentation in the southern rock agama, *Agama atra*: more evidence for vicariance in southern Africa. *Molecular Ecology*, 11: 465 - 471.

Maynard-Smith J., and Haigh N.D.J. 1974. The hitch-hiking effect of a favourable gene. *Genetic Research*, 23: 23–35.

Mayr E. 1942. Systematics and the origin of species. Columbia University Press, New York, USA.

Mayr E. 1947. Ecological factors in speciation. *Evolution,* 1: 263 – 288.

Mayr E. 1963. Animal species and evolution. Harvard University Press, Cambridge, USA.

McCracken K.G., Bulgarella M., Johnson K.P., Kuhner M.K., Trucco J., Valqui T.H., Wilson R.E., Peters J.L. 2009. Gene flow in the face of countervailing selection: adaptation to high altitude hypoxia in the βA hemoglobin subunit of yellow-billed pintails in the Andes. *Molecular Biology and Evolution*, 26: 815 - 827.

McInerny G.J., Turner J.R.G., Wong H.Y., Travis J.M.J., Benton T.G. 2009. How range shifts induced by climate change affect neutral evolution. *Proceeding of the Royal Society B.* 276: 1527 – 1534.

Metcalf C.J.E., and Pavard, S. 2007. Why evolutionary biologists should be demographers *Trends in Ecology and Evolution*, 22: 205 - 212.

Midgley J.J., Cowling R.M., Seydack A.H.W., van Wyk G.F. 1997. Forest. *In* Vegetation of Southern Africa, R.M. Cowling, D.M., Richardson and S.M. Pierce, eds., pp. 167 - 188. Cambridge University Press, Cambridge, UK.

Milá B., Wayne R.K., Fitze P., Smith T.B. 2009. Divergence with gene flow and fine-scale phylogeographic structure in the wedge-billed woodcreeper *Glyphorynchus spirurus*, a Neotropical rainforest bird. *Molecular Ecology*, 18: 2879 - 2995.

Milton S.J., Yeaton R.I., Dean W.R.J., Vlok J.H.J. 1997. Succulent Karoo. *In* Vegetation of Southern Africa, R.M. Cowling, D.M., Richardson and S.M. Pierce, eds., pp. 131 - 166. Cambridge University Press, Cambridge, UK.

Mishmar D., Ruiz-Pesini E., Golik P., Macaulay V., Clark A.G., Hoseini S., Brandon M., Easely K., Chen E., Brown M. D., Sukerniki R.I., Olckers A., Wallace D.C. 2003. Natural selection shaped regional mtDNA variation in humans. *Proceedings of the National Academy of Sciences USA*, 100: 171 - 176.

Moore J., and Ali R. 1984. Are dispersal and inbreeding avoidance related? *Animal Behaviour*, 32: 94 – 112.

Moreau R.E. 1948. Ecological isolation in a rich tropical avifauna. *Journal of Animal Ecology* 17: 113 - 126.

Moritz C., Patton J.L., Conroy C.J., Parra J.L., White G.C., Beissinger S.R. 2008. Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science,* 322:261–264.

Mucina L., and Rutherford M.C. 2006. The Vegetation of South Africa, Lesotho and Swaziland, South African National Biodiversity Institute, Pretoria, South Africa.

Mullen L.M., and Hoekstra H.E. 2008. Natural selection along an environmental gradient: A classic cline in mouse pigmentation. *Evolution*, 62: 1555 - 1569.

Myles S., Somel M., Tang K., Kelso J., Stoneking M. 2007. Identifying genes underlying skin pigmentation differences among human populations. *Human Genetics*, 120: 613 - 621.

**N**

Nachman M. 2006. Detecting selection at the molecular level, pp. 103 - 118, *In* Evolutionary Genetics, concepts and case studies, C.W. Fox and J.B. Wolf, eds., Oxford University Press, Oxford, UK.

Nathan R., Getz W. M., Revilla E., Holyoak M., Kadmon R., Saltz D., Smouse, P. E. 2008 A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences USA,* 105: 19 052 – 19 05

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

Nielsen R. 2005. Molecular signatures of natural selection. *Annual Review of Genetics*, 39: 197 – 218.

Nielsen R., and Wakeley J. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, 158: 885 - 896.

Nix H.A. 1986. A biogeographical analysis of Australian elapid snakes. pp. 4-10 *In* Snakes: Atlas of Elapid Snakes of Australia. Australian Bureau of Flora and Fauna, R. Longmore, ed., Canberra.

Nosil P. 2004. Reproductive isolation caused by visual predation against migrants between divergent environments. *Proceeding of the Royal Society B*, 271: 1521 – 1528.

Nosil P., Egan S.P., Funk D.J. 2008. Heterogeneous genomic differentiation between walking-stick ecotypes: 'isolation by adaptation' and multiple roles for divergent selection. *Evolution*, 62: 316 – 336

Nosil P., Funk D.J., Ortiz-Barrientos D. 2009a. Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, 18: 375 - 402.

Nosil P., Harmon L.J., Seehausen O. 2009b. Ecological explanations for (incomplete) speciation. *Trends Ecology and Evolution*, 24: 145–156.

**O**

Oatley T., and Arnott G. 1988. Robins of Africa. Acorn Books, Randburg and Russel Friedman Books:

Halfway House, South Africa.

Oatley T.B. 2005. Karoo Scrub-Robin *Cercotrichas coryphaeus*. *In* Roberts Birds of Southern Africa, Hockey PAR, Dean WRJ and Ryan PG, eds. VII edition, pp. 942 - 943. The Trustees of the John Voelcker Bird Book Fund, Cape Town, South Africa.

Ogden R., and Thorpe R.S. 2002. Molecular evidence for ecological speciation in tropical habitats. *Proceedings of the National Academy of Sciences USA*, 99: 13612 - 13615.

Oksanen J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2008. VEGAN: Community Ecology Package. R Package Version 1.15-1.

Odling-Smee F.J. 1988. Niche constructing phenotypes. *In* The role of behavior in evolution. H.C. Plotkin,ed., MIT Press, USA

Odling-Smee F.J., Laland K.N., Feldman M.W. 2003. Niche Construction: The Neglected Process in Evolution. Monographs in Population Biology. Princeton University Press, New Jersey, USA.

Outlaw R.K., Voelker G., Outlaw D.C. 2007. Molecular systematics and historical biogeography of the rock-thrushes (Muscicapidae: *Monticola*). *Auk*, 124: 561 – 577.

**P**

Palmer A.R., and Hoffman M.T. 1997. Nama Karoo. *In* Vegetation of Southern Africa R.M. Cowling, D.M., Richardson and S.M. Pierce, eds., pp. 167 - 188. Cambridge University Press, Cambridge, UK.

Parmesan C., Ryrholm N., Stefanescu C., Hill J.K., Thomas C.D., Descimon H., Huntleyk B., Kaila L., Kullberg J., Tammaru T., Tennent W.J., Thomas J.A., Warren M. 1999. Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, 399: 579 – 583

Parmesan C. 2006. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology and Evolutionary Systematics*, 37: 637 – 669.

Partridge T.C. 1997. Evolution of landscapes. pp. 5-20 *In* Vegetation of Southern Africa. R.M. Cowling, D.M. Richardson and S.M. Pierce, eds., Cambridge University Press. Cambridge, UK.

Peakall R., and Smouse P.E. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6: 288 – 295.

Pereira R.J., and Wake D.B. 2009. Genetic leakage after adaptive and non-adaptive divergence in the *Ensatina eschscholtzii* ring species. *Evolution,* 63: 2288 - 2301.

Perrin N., and Mazalov V. 2000. Local competition, inbreeding, and the evolution of sex-biased dispersal. *American Naturalist,* 155: 116 – 27.

Perrin N., and Goudet, J. 2001. Inbreeding, kinship, and the evolution of natal dispersal. *In* Dispersal, J. Clobert, E. Danchin, A. Dhondt and J.D. Nichols, eds., Oxford University Press, Oxford, UK.

Petrie M., and Kempenaers B. 1998. Extra-pair paternity in birds: explaining variation between species and populations. *Trends in Ecology and Evolution*, 13: 52 – 58.

Pedersen P.L., Ko Y.H., Hong S. 2000. ATP Synthesis in the year 200: Defining the different levels of mechanism and getting a grip on each. *Journal of Bioenergetics and Biomembranes*. 32: 423 - 432.

Phillips B.L., Brown G.P., Webb J.K., Shine R. 2006. Invasion and the evolution of speed in toads. *Nature*, 439: 803

Phillips S.J., and Dudik M. 2008. Modeling of species distributions with MaxEnt: new extensions and a comprehensive evaluation. *Ecography,* 31: 161 – 175.

Pond S.L.K., Frost S.D.W, Muse S.V. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics,*

21: 676 – 679.

Pond S.L.K., Posada D., Gravenor M.B., Woelk C.H., Frost S.D.W. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Bioogy and Evol*ution, 23: 1891 – 1901.

Pool J.E., Nielsen R. 2007. Population size changes reshape genomic patterns of diversity. *Evolution,* 61: 3001 – 3006.

Poynton, J.C., 1961. Biogeography of south-east Africa. *Nature,* 189: 801 – 803.

Price T. 2008. Speciation in birds. Roberts and Company, Greenwood Village, USA.

Primmer C.R., Borge T., Lindell J., Saetre, 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., Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155: 945 – 959.

du Plessis M.A., Siegfried W.R., Armstrong A.J. 1995. Ecological and life-history correlates of cooperative breeding in South-African birds. *Oecologia*, 102: 180 – 188.

**Q**

Queller D.C., and Goodnight K.F. 1989. Estimating relatedness using genetic markers. *Evolution,* 43: 258 – 275.

**R**

R Development Core Team. 2011. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ramos-Onsins S.E., and Rozas J. 2002. Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, 19: 2092 – 2100.

Ray N., Currat M., Excoffier L. 2003. Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology and Evolution*, 20:76–86.

Raymond M., and Rousset F. 1995. GENEPOP (version 3.4): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86: 248 - 249.

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

Richardson J.E., Weitz F.M., Fay M.F., Cronk Q.C.B., Linder H.P., Reeves G., Chase M.W. 2001. Rapid and recent origin of species richness in the Cape Flora of South Africa. *Nature,* 412: 181 – 183.

Ricklefs R.E., and Wikelski M. 2002. The physiology/life history nexus. *Trends in Ecology and Evolution*, 17: 462 - 468.

Rogers A.R., and Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, 9: 552–692.

Rolán-Alvarez E., Johannesson K., Erlandsson J. 1997. The maintenance of a cline in the marine snail *Littorina saxatilis*: the role of home site advantage and hybrid fitness. *Evolution*, 51: 1836 – 1847.

Ronce O. 2007. How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *The Annual Review in Ecology, Evolution, and Systematics*, 38: 231 – 53.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, 145: 1219 - 1228.

Rosenblum E.B. 2006. Convergent evolution and divergent selection, lizards at the White Sands ecotone. *American Naturalist*, 167: 1 – 15.

Rosenblum E.B., Hickerson M.J., Moritz C. 2007. A multilocus perspective on colonization accompanied by selection and gene flow. *Evolution*, 61: 2971 – 2985.

Rosenblum E.B., Römpler H., Schöneberg T., Hoekstra H.E. 2010 Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proceedings of the National Academy of Sciences USA,* 107: 2113 - 2117

Rundle H.D., and Nosil P. 2005. Ecological speciation. *Ecology Letters*, 8: 336 – 352.

Ryan P., Hood I., Bloomer P., Komen J., Crowe T. 1997. Barlow's Lark: a new species in the Karoo Lark *Certhilauda albescens* complex of southwest Africa. *Ibis*, 140: 605 - 619.

**S**

Saccheri I., and Hanski, I. 2006. Natural selection and population dynamics. *Trends in Ecology and Evolution*, 21: 341 – 347.

Schemske D.W. 2000. Understanding the origin of species. *Evolution,* 54: 1069 – 1073.

Schluter D. 2000. The ecology of adaptive radiation. Oxford Univ. Press, New York, USA.

Schluter D. 2001. Ecology and the origin of species. *Trends in Ecology and Evolution*, 16: 372 – 380.

Schluter D. 2009. Evidence for ecological speciation and its alternative. *Science,* 323: 737 – 741.

Schoener T.W. 2011. The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. *Science*, 331: 426 – 429.

Scott L., Anderson H.M., Anderson J.M. 1997. Vegetation history. *In* Vegetation of Southern Africa, R.M. Cowling, D.M., Richardson and S.M. Pierce, eds., pp. 62 - 69. Cambridge University Press, Cambridge, UK.

Scott G.R., Schulte P.M., Egginton S., Scott A.L.M, Richards J.G., Milsom W.K. 2011. Molecular Evolution of Cytochrome c Oxidase Underlies High-Altitude Adaptation in the Bar-Headed Goose.

*Molecular Biology and Evolution*, 28: 351 – 363.

Seehausen O. 2006. African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B,* 273: 1987 – 199.

Sexton J.P., McIntyre P.J., Angert A.L., Rice K.J. 2009. Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution and Systematics*, 40: 415 – 436.

Slatkin M. 1987. Gene flow and the geographic structure of natural populations. *Science*, 236: 787 – 792.

Slatkin M. 1985. Gene flow in natural populations. *Annual Review in Ecology and Systematics*, 16: 93 – 430.

Smith F.A., Betancourt J.L., Brown, J.H. 1995. Evolution of body size in the woodrat over the past 25 000 years of climate change. *Science*, 270: 2012 – 2014.

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

Smith T.B., Schneider C.J., Holder K. 2001. Refugial isolation versus ecological gradients: testing alternative mechanisms of evolutionary divergence in four rainforest vertebrates. *Genetica*, 112: 383–398.

Slabbekoorn H., and Smith T.B. 2002. Bird song, ecology and speciation. *Philosophical Transactions of the Royal Society B,* 357: 493 – 503.

Smouse P.E., Long J., Sokal R.R. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, 35: 627 - 632.

Smouse P.E., and Peakall R. 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity,* 82: 561 – 573.

Shi Y., and Yokoyama S. 2003. Molecular analysis of the evolutionary significance of ultraviolet vision in vertebrates. *Proceedings of the National Academy of Sciences USA*, 100: 8308 - 8313.
Sobel J.M., Chen G.F., Watt, L.R., and Schemske, D.W. 2009. The biology of speciation. *Evolution,* 64: 295 – 315.

Stenseth N.C., and Lidicker W.Z. 1992. Animal dispersal: small mammals as a model. Chapman and Hall, London, UK.

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

Stephens M., and Donnelly P. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal Human Genetics*, 73: 1162 – 1169.

Storz J.F., and Dubach J.M. 2004. Natural selection drives altitudinal divergence at the albumin locus in deer mice, *Peromyscus maniculatus*. *Evolution,* 58: 1342 - 1352.

Storz J.F., Sabatino S.J., Hoffmann F.G., Gering E.J., Moriyama H., Ferrand N., Monteiro B., Nachman M.W. 2007. The molecular basis of high-altitude adaptation in deer mice. *Plos Genetics*, 3: 448 - 459.

Storz J.F., Runck A.M., Sabatino S.J., Kelly J.K., Ferrand N., Moriyama H., Weber R.E., Fago A. 2009. Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. *Proceedings of the National Academy of Sciences USA*, 106: 14450 – 14455.

Strasburg J., and Rieseberg LH. 2009. How robust are ‘‘Isolation with Migration’’ analyses to violations of the IM model? A simulation study. *Molecular Biology and Evolution*, 27: 297 – 310.
Strasburg J., and Rieseberg L.H. 2011. Interpreting the estimated timing of migration events between hybridizing species. *Molecular Ecology*, 20: 2353 - 66.

Sutherland W.J. 1996. From individual behavior to population ecology. Oxford University Press, Oxford, UK.

Swart B.L., Tolley K.A., Matthee C.A. 2009. Climate change drives speciation in the southern rock agama (*Agama atra*) in the Cape Floristic Region, South Africa. *Journal of Biogeography*, 36: 78 - 87.

**T**

Tajima F. 1983. Evolutionary relationships of DNA sequences in finite populations. *Genetics*, 105: 437 – 460.

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

Temple H.J., Hoffman J.I., Amos W. 2006. Dispersal, philopatry and intergroup relatedness: fine-scale genetic structure in the white-breasted thrasher, *Ramphocinclus brachyurus*. *Molecular Ecology*, 15: 3449 - 3458.

Theron E., Hawkins K., Bermingham E., Ricklefs R.E., Mundy N.I. 2001. The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanic plumage morph of the bananaquit, *Coereba flaveola*. *Current Biology*, 11: 550 - 557.

Thorpe W.H. 1945. The evolutionary significance of habitat selection. *Journal of Animal Ecology*, 14: 67 - 70.

Thorpe R.S., Malhotra A., Stenson A.G., and Reardon, J.T. 2004. Adaptation and speciation in Lesser Antillean anoles. *In* Adaptive speciation. U. Dieckmann, H. A. J. Metz, M. Doebeli and D. Tautz, eds., pp. 324 – 335. Cambridge University Press, Cambridge, UK.

Thorpe R.S., Surget-Groba Y., Johansson H. 2008. The relative importance of ecology and geographic isolation for speciation in anoles. *Philosophical Transactions of the Royal Society B*, 363: 3071–3081.

Tieleman B.I., Williams J.B., Bloomer P. 2003. Adaptation of metabolism and evaporative water loss along an aridity gradient. *Proceedings Royal Society B*, 270: 207 – 214.

Tieleman B.I., Versteegh M.A., Fries A. Helm B., Dingemanse NJ. Gibbs, H.L., Williams J.B. 2009. Genetic modulation of energy metabolism in birds through mitochondrial function. *Proceedings of the Royal Society B,* 276: 1685 - 1693.

Tingley M.W., Monahan W.B., Beissinger S.R., Moritz C. 2009. Birds track their Grinnellian niche through a century of climate change. *Proceedings of the National Academy of Science USA*, 106: 19637 – 19643.

Tolley K.A., Burge M., Turner A.A., Matthee C.A. 2006. Biogeographic patterns and phyogeography of dwaf chamaleons (*Bradypodion*) in an African biodiversity hotspot. *Molecular Ecology*, 15: 781 - 793.

della Torre A., Costantini C., Besansky N.J., Caccone A., Petrarca V., Powell J.R., Coluzzi M. 2002. Speciation within *Anopheles gambiae* – the glass is half full. *Science*, 298: 115 - 117.

Toews D.P.L., and Irwin D.E. 2008. Cryptic speciation in a Holarctic passerine revealed by genetic and bioacoustic analyses. *Molecular Ecology,* 17: 2691 – 2705.

Turelli M., Barton N.H., and Coyne J.A. 2001. Theory and speciation. *Trends in Ecology and Evolution,* 16: 330 – 343.

Turner T.L., Hahn M.W., Nuzhdin S.V.. 2005. Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, 3: e285.

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

**V**

Van Bocxlaer I., Loader S.P., Roelants K., Biju S.D., Menegon M., Bossuyt F. 2010. Gradual adaptation toward a range-expansion phenotype initiated the global radiation of toads. *Science*, 327: 679 – 682.

Van Wyk A.E. 1996. Biodiversity of the Maputaland Centre. *In* The biodiversity of African plants: proceedings XIVth AETFAT congress 22-27 August 1994, pp.198 – 207.van der Maesen L.J.G., van der Burgt X.M., van Medenbach de Rooy J.M., eds., Kluwer Academic, Wageningen, The Netherlands.

Via S. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology and Evolution*, 16: 381 – 390.

Via S. 2009. Natural selection in action during speciation. *Proceeding of the National Academy of Sciences USA*, 106: 9939 – 9946.

Vines T. H., and D. Schluter. 2006. Strong assortative mating between allopatric sticklebacks as a by-product of adaptation to different environments. *Proceedings of the Royal Society B*, 273: 911 – 916.

**W**

Waits L.P., Luikart G., Taberlet P. 2001. Estimating the probability of identity among genotypes in natural populationns: cautions and guidelines. *Molecular Ecology*, 10: 249–25.

Waterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* 7: 256–276.

Watterson G.A .1986. The homozygosity test after a change in population size. *Genetics*, 112: 899–907.

Watson A., Moss R., Parr R., Mountford M.D., Rothery P. 1994. Kin landownership, differential aggression between kin and non-kin, and population fluctuations in red grouse. *Journal of Animal Ecology*, 63: 39–50.

Werger M.A. 1986. The Karoo and southern Kalahari. pp. 283-354. *In* Hot Deserts and Arid Shrublands. M. Evenari, I. Noy-Meir and D. W. Goodall, eds., Elsevier, Amsterdam.

White C., Blackburn T., Martin G., Butler P. 2007. Basal metabolic rate of birds is associated with habitat temperature and precipitation, not primary productivity. *Proceeding of the Royal Society B*: 274: 287 – 293

Wiens J.J. 2004. Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution,* 58:193–197.

Wiersma P., Munoz-Garcia A., Walker A., Williams J.B. 2007. Tropical birds have a slow pace of life. *Proceedings of the National Academy of Sciences USA,* 104: 9340 – 9345.

Wilson G.A., and Rannala B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, 163: 1177 – 1191.

Williams J.B, and Tieleman B.I. 2005. Physiological adaptation in desert birds. *BioScience*, 55: 416-425.

Wright S. 1931. Evolution in Mendelian populations. *Genetics*, 16: 97 - 159.

Wright S. 1943. Isolation by distance. *Genetics*, 28: 114 – 138.

Woxvold I.A., Adcock G.J., Mulder R.A. 2006. Fine-scale genetic structure and dispersal in cooperatively breeding apostlebirds. *Molecular Ecology*, 15: 3139 – 3146.

# POSTFACE

## Harmony typewriter

*In music the only thing that matters is whether you feel it or not. You cannot intellectualize music.*
| Ornette Coleman |

During the loneliness associated with the writing period of my dissertation and manuscripts, time started to be measured in lines, paragraphs, chapters and also melodies, tracks and albums, such that sometimes they confound themselves. I listened to some musicians compulsively, without even trying to intellectualize it. The providers of rhythm and perhaps inspiration were: Virgínia Rodrigues \ Nós; Maria Bethania \ Pirata and Rosalia de Souza \ Garota moderna while writing the Introduction and Conclusion. Cristina Branco \ Abril and Orchestra Baobab \ Specialist in All styles are the melody undelying chapter two. Bad Plus \ Give | Suspicious Activity, Dobet Gnahoré \ Na Afriki and Joshua Redman \ Freedom in the Groove provided the silence while writing chapter three. Beirut \ Live at the Music Hall of Williamsburg), Andrew Bird \ Armchair Apocrypha | Soldier On and Tinariwen \ Amassakoul were my hourglass for chapter four. Concha Buika \ El último trago, Ali Farka Touré \ Talking Timbuktu | Red & Green and, Ballaké Sissoko and Vincent Segal \ Chamber music marked the cadence for chapter five. Bon Iver \ For Emma, Forever ago | Bon Iver and Gretchen Parlato \ In a dream | Lost and Found were always 'never enough' while writing chapter six.