Dispersal and Gene Flow in the Southern African Endemic Lacertid, *Pedioplanis lineoocellata,* Based on Microsatellite and Capture-Mark-Recapture Data.

Ryan Joseph Daniels[12]

Thesis is presented in partial fulfilment of the requirements for the degree of Master of Science in the Faculty of Science at Stellenbosch University.

Principle Supervisor: Dr Krystal Tolley[12]

Co-supervisor: Dr Res Altwegg[3] & Dr Susana Clusella-Trullas[2]

December 2014

[1]Applied Biodiversity Research, South African National Biodiversity Institute, Private Bag X7, Claremont, Cape Town, 7735, South Africa.

[2]Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland, 7602 South Africa

[3]Statistics in Ecology, Environment and Conservation, Department of Statistical Sciences, University of Cape Town, Private Bag X3, Rondebosch, Cape Town, 7701 South Africa

Dispersal and Gene Flow in the Southern African Endemic Lacertid, *Pedioplanis lineoocellata*, Based on Microsatellite and Capture-Mark-Recapture Data

SANBI
Biodiversity for Life

UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot · your knowledge partner

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Friday, 02 September 2014

Signature: …………………………………………………..

Location: SANBI, Cape Town

ABSTRACT

Dispersal determines connectivity between populations within a species and is a regulator of genetic differentiation through gene flow. Although the necessity of dispersal for gene flow is clear, for many taxa the relationship between the two is not well understood. Gene flow, or a restriction thereof, may be inferred from population-level genetic divergence estimates. These measures are averages of contemporary and historic gene flow and as such they are not necessarily easily compared to measures of real-time dispersal. Changes in dispersal have been inferred from present day spatial genetic structure for many southern African taxa and further associated with environmental change events. *Pedioplanis lineoocellata* is a southern African endemic lacertid with a mitochondrial DNA structure that may have been the result of Plio-Pleistocene glacial climatic oscillations. As a wide-spread, open habitat species, *P. lineoocellata* is an excellent study species for examining the relationship between dispersal and gene flow. In the first data chapter, Chapter 2, nine new microsatellite markers are described for several populations for the purpose of examining gene flow and genetic structure in the species. The possibility of null alleles, population bottlenecks and high inbreeding are investigated as possible explanations for the detected deviation from Hardy-Weinberg equilibrium (HWE). The presence of null alleles and, at one population, relatively high inbreeding best explains the HWE deviations. While null allele frequencies were not excessively high, this caveat should be borne in mind when interpreting results. In Chapter 3 the microsatellite markers were used to assess the geographic genetic patterns for *P. lineoocellata* across the distribution of the two most wide-spread mitochondrial lineages and to test for evidence of hybridization at a point of clade contact in the Loeriesfontein area. Microsatellite genetic clusters did not match the mtDNA lineages, a possible result of gene flow between clades. However, measures of genetic differentiation and recent migration indicate only weak contemporary long distance gene flow. There was no evidence of genetic admixture at the Loeriesfontein area despite sympatric mtDNA lineages. The complexity of the geographic arrangement of the microsatellite clusters may be attributed to historic range contraction and expansion events for the species. In the last data chapter, evidence for an isolation-by-distance (IBD) pattern was examined within the most widespread mtDNA clade. Sampling over hundreds of kilometres produced an IBD pattern when using spatial autocorrelation while failure to detect IBD using the Mantel test was likely a result of the complex arrangement of microsatellite clusters. A combination of genetic data and demographic data was used to estimate the annual dispersal distances based on the neighbourhood size concept. Results indicated high levels of dispersal that covered distances of a few hundred metres, greater than is expected for a lacertid lizard. Strong dispersal

propensity would have influenced gene flow and genetic structure found in this thesis and will further influence future responses to environmental changes for the species.

Opsomming

Verspreiding (beweeglikheid) bepaal die verbinding tussen populasies van 'n spesie en is 'n reguleerder van genetiese differensiasie deur middel van gene vloei. Alhoewel die noodsaaklikheid van verspreiding vir baie taksa duidelik is, word die verhouding tussen die twee nie goed verstaan nie. Gene vloei, of 'n beperking daarvan, kan vanaf populasie-genetika divergensie skattings afgelei word. Hierdie maatreëls is gemiddeldes van die huidige en historiese gene vloei, en dus is dit nie maklik vergelykbaar met hedendaagse verspreiding nie. Veranderinge in die verspreiding is afgelei van die huidige geografiese genetiese struktuur vir baie Suider-Afrikaanse taksa en verdere veranderinge wat verband hou met omgewingsgebeurtenisse. *Pedioplanis lineoocellata* is 'n Suider-Afrikaanse endemiese sand-akkedis, met n mitochondriale DNA struktuur wat die gevolg is van Plio-Pleistoseen glacial klimaat ossillasies. As 'n wydverspreide oop habitat spesie, is *P. lineoocellata* 'n geskikte studie spesies om die verhouding tussen die verspreding en gene vloei te ondersoek. In die eerste data hoofstuk, Hoofstuk 2, word nege nuwe mikrosatelliet merkers vir verskeie populasies beskryf met die doelwit om gene vloei en genetiese struktuur in hierdie spesie te ondersoek. Die moontlikheid van nul allele, populasie knelpunte en hoë-frekwensie inteling word ondersoek as moontlike verklarings vir die afwyking vanaf Hardy-Weinberg ewewig (HWE) wat opgemerk was. Hardy-Weinberg ewewig afwykings word die beste verduidelik deur die teenwoordigheid van nul allele en die relatiewe hoë inteling binne een spesifieke populasie. Alhoewel, alleelfrekwensies nie buitensporig hoog was nie, moet die bogenoemde maatstaf steeds in ag geneem word, wanneer resultate geïnterpreteer word. In Hoofstuk 3 word die mikrosatelliet merkers gebruik om die geografiese genetiese patrone oor die verspreiding van die mees wydverspreide mitochondriale linies te evalueer. Verdere toetse vir die bewys van verbasterigng by n geografiese kontakpunt van twee genetiese groepe in Loeriesfontein was gedoen.Resultate toon dat genetiese groepe nie ooreenstemmend is tussen mikrosatelliet en mtDNA data nie, en dat dit heelmoontlik n gevolg van genevloei tussen klades kan wees. In teenstryding, toon die maatstawwe van genetiese differensiasies en onlangse migrasie swak gene vloei oor langafstande. Daar is geen genetiese vermenging in die Loeriesfontein area nie, ongeag van die simpatriese verspreiding vir twee klades. Die kompleksiteit van die geografiese indeling van die mikrosatelliet groepe kan toegeskryf word aan historiese inkrimping en uitbreiding gebeuternisse van die spesie. In die laaste data hoofstuk, word 'n isolasie-deur-afstand (IBD) patroon binne die mees wydverspreide mtDNA klade ondersoek. Opnames oor honderde kilometres het n IBD patroon getoon wannner ruimtelike outokorrelasie gebruik was, terwyl die gebruik van n Mantel toets gevaal het om `n IBD patroon op te tel, en kan moontlik toegeskryf word aan die komplekse rangskikking van die

mikrosatelliet groepe. 'n Kombinasie van genetiese en demografiese data was gebruik om die jaarlikse verspreiding afstande wat gebaseer is op die "buurt-omgewing" grootte konsep te skat. Resultate het hoë vlakke van verspreiding wat afstande van 'n paar honderd meter, groter as wat verwag word vir 'n san-akkedis getoon. Sterk verspreiding geneigdheid sou gene vloei en genetiese struktuur beïnvloed en sal ook verdere impak maak op hoe hierdie spesies in die toekoms sal reageer op omgewingsveranderinge.

ACKNOWLEDGEMENTS

I would also like to thank the staff and students at the South African National Biodiversity Institute (SANBI) for their friendliness and hospitality during my time at SANBI over the last three years. I have deep gratitude for my supervisors Dr Krystal Tolley, Dr Res Altwegg and Dr Susana Clusella-Trullas for the assistance with administration, finances and, most importantly, their patience in wading through my numerous, wordy drafts to help me find what needed to be said. Without their experience I would not have been able to complete this work. I extend sincere thanks to my lab-mates and colleagues for the years of field assistance, lab advice, guidance and laughter over silly lab jokes; Paula Strauss, Zoë Davids, Keshni Gopal, Buyi Makhubo, Hanlie Engelbrecht and Shandre Dreyer. Lastly, on a personal note, I would like to thank my statistics beer-buddies, Neo Mohapi and Greg Duckworth, for all the advice and encouragement that lead, and other times dragged, me through my MSc. Without their drunken words of wisdom I would not have made it through this with my sanity in-tact.

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

TABLE OF APPENDICES

Appendix A: Tables, Figures and Information

Appendix B: Tables, Figures and Information

Chapter 1 GENERAL INTRODUCTION: ECOLOGY OF DISPERSAL AND GENE FLOW

1 DISPERSAL ECOLOGY

Ecology and evolution are connected through a web of processes and feedback loops. At the most fundamental level, the genes of an individual control a myriad of idiosyncrasies and life history traits (Shields 1987). Life history traits, including age-specific survival, reproduction, and dispersal determine populations dynamics through population-level demographic processes such as population growth rates, meta-population sink-source dynamics, rates of population turnover and fluctuations in effective population size (Bauwens et al. 1997). The processes influence the reproductive output, health and social interactions of individuals within the populations at a somatic and genetic level and thus affect the individual's evolutionary fitness. This completes the link between genetics and demographic processes (Bullock et al. 2002; Clobert et al. 2012).

Life history traits may vary with geography, particularly in species with wide-spread distributions. Variation is the result of phenotypically plastic responses to the local environment and the result of genetic divergences generated through mutation, genetic drift and selective forces (Clobert et al. 1994; Niewiarowski 1994; Hartl et al. 1997). Genetic differentiation is of particular interest because it is the means by which populations become locally adapted, evolutionarily divergent and ultimately the process by which they speciate. Gene flow- the movement of genes across a landscape- may act to reduce genetic divergence between populations by sharing alleles among populations (Wright 1935; Bohonak 1999). It is through dispersal that this genetic connectivity between populations is maintained. Dispersal is, however, a demographic process because it directly affects population size and density without necessarily influencing the genetic composition of any population. Only with successful reproduction in the new environment can the dispersal of an individual change genetic structure and allele frequencies between populations, and thus contribute to gene flow. Understanding the spatial connections between populations facilitates understanding genetic differences, making a thorough understanding of dispersal fundamental to population ecology and conservation efforts (Wiens 1997; Hanski 1999).

1.1 DISPERSAL DEFINITIONS

The term "genetic dispersal" describes the movement of individuals relative to the breeding sites and includes "ecological movements", which simply describe the movements of an individual, and are not necessarily related to reproduction (Johnson et al. 1990). An example of an ecological movement is daily movement undertaken for foraging, grooming, mate searching and territorial activities (Shields 1987). The space covered during these activities constitutes an individual's 'home range' (Lawson et al. 1997; Perry et al. 2002). Movements within the home range (HR) are frequent and regular as compared to the once-off nature of a dispersal event and the individual does not stray beyond the boundary of the range; these

characteristics distinguish daily activities from dispersal. Natal dispersal is the permanent movement of an individual from a natal site to the place of reproduction, provided it survives and successfully reproduces (Johnson et al. 1990). Breeding dispersal is the movement from one home range to another between attempts at reproduction and occurs after the initial natal dispersal event (Johnson et al. 1990). The magnitude of dispersal may vary between species, within species between populations, between sexes and phenotypes, and even across the life time of an individual. The frequency distribution of dispersal distances for species is described using a dispersal kernel. For many species dispersal is strongly skewed toward shorter distance movements (Nathan *et al.* 2012) and is typically only a few home range diameters (Shields 1987).

Migration is yet another term in population genetics used to refer to the movement of individuals. It is typically used as a synonym for dispersal rate (m) -the proportion of individuals in a population that were in a different population in the preceding generation (Shields 1987; Hartl et al. 1997). "Population", as with "dispersal", has several different definitions, each applied in different contexts (Waples et al. 2006). I define population for this thesis according to the evolutionary (genetic) paradigm as a group of conspecifics that live close enough to one another to potentially mate with any member of the group and exhibit reproductive continuity from one generation to the next.

Despite the synonymous use of "migration" with "dispersal" in genetic papers (e.g. Johnson & Gaines 1990; Collingham & Huntley 2000; Bowie et al. 2006), migration is more commonly thought of as temporally-predictable ecological movements such as seasonal migration, which are completely different in form and function to dispersal. Despite this ambiguity, the terms 'immigrate' and 'emigrate' have clear definitions in population genetics. For the remainder of this thesis I use 'dispersal' to mean natal and/or breeding dispersal and dispersal rate or migration to represent the proportion of 'migrants' in a particular population. Immigrant refers to an individual who was born in a different population to the one within which it currently resides and emigrant refers to an individual who is in the process of or has already moved out of the population of interest to take up residence in another.

1.2 EVOLUTION OF DISPERSAL BEHAVIOUR

The evolution of dispersal has been investigated using theoretical modelling and species-specific empirical investigations but the great variation in dispersal biology between species hinders the formulation of over-arching generalisations (reviewed by Johnson et al. 1990; Clobert et al. 2012). Instead there appears to be several possible suites of drivers that may vary taxonomically.

Several studies have demonstrated how dispersal propensity is strongly influenced by a number of life history traits including size, sex, age and ontological phase, and daily movement (vagile or sedentary), in addition to external cues like the attractiveness of the destination populations and the distance between

populations (Shields 1987; Clobert *et al.* 1994; Bohonak 1999; Schowalter 2006; Clobert 2012). Dispersal has been identified across-the-board in all species examined to date thus there may be a common and significant benefit for undergoing some dispersal as opposed to none at all. The competition for free space through colonisation, even in temporally and spatially consistent environments, may have resulted in the ubiquity of dispersal (McPeek et al. 1992; Holt et al. 2001; Clobert et al. 2012). Furthermore, as environments and phenotypes vary temporally and spatially, so does life history traits as a result of varying selective pressures (McPeek et al. 1992). Dispersal may be an evolutionary bet-hedging strategy that allows individuals to adjust conditions temporally and spatially and, consequently, their relative fitness (McPeek et al. 1992; Holt et al. 2001). Thus dispersal is an evolutionary trait which would be beneficial for all organisms except those in the most stable environments. However, other factors likely contribute toward or regulate dispersal more strongly as indicated by the variation in the form of dispersal between species.

The evolution of dispersal behaviour is driven primarily by the genetic or somatic implications of dispersal. Genetic explanations consider inbreeding and outbreeding avoidance as possible drivers. Inbreeding depression should increase dispersal regardless of the costs of dispersing because of the inclusive fitness gained (McPeek et al. 1992). Conversely, outbreeding depression should decrease dispersal especially as the cost increases. Although genetics has been the traditional explanation for dispersal, there is much evidence that better supports somatic drivers in some species. For some lizards such as *Lacerta agilis*, however, inbreeding has demonstrable costs (Olsson et al. 1996) and evidence for selective pressures against close kin mating (Léna et al. 1998a, 1998b).

Somatic implications of dispersal consider competition for resources and the benefits and costs of dispersal (Shields 1987; Clobert *et al.* 1994; Bohonak 1999). Somatic drivers may have greater selective force than genetic drivers because of the direct cost for the disperser, though not necessarily the offspring, as well as the greater immediacy at which they affect the individual (Shields 1987; Andreassen et al. 2002). For selective pressures to favour dispersal, the benefits of dispersal need to balance the excessive, non-fatal costs. Alternatively, should benefits to the offspring in the receiving population outweigh the non-fatal cost for the dispersing parent, selection should favour dispersal irrespective.

Temporal and spatial variation in competition for food, shelter and mates are important for determining dispersal behaviour. Kin competition at the natal site selects for dispersal as the inclusive fitness gained by avoiding competition between kin may offset the costs of dispersing (Hamilton et al. 1977; McPeek et al. 1992). Asymmetrical competition between kin, such as overt oppression by dominant individuals, promotes dispersal because it offers a means to retreat from competition (McPeek et al. 1992). When environments are unstable, subordinate individuals should disperse as there is greater chance of encountering newly 'freed-up' territory elsewhere. The influence of social structural on dispersal has been particularly well studied for mammals and birds because of the important role it plays in the evolution of life history and the influence it

has on breeding systems (Greenwood 1980; Shields 1987). In contrast, most reptiles, barring a few exceptions, have no known social structures and are predominantly solitary, making kin competition only a minor determinant of dispersal (Shields 1987; Ronce et al. 1998; Stow et al. 2001, 2004; Milne et al. 2002; Chapple et al. 2005).

For many species age and sex have been identified as important determinants of how an individual within a population experiences resource competition and for both categories differential resource competition may lead to biases in dispersal. Particularly well studied is the different dispersal distances for sexes, because separating sexes can retain the benefits of genetic outbreeding but also the somatic advantages of philopatry for either of the sexes (Greenwood 1980). The form of competition either for mates or for other resources determines which sex disperses further. For example, females as a resource may be clumped in time and space for various reasons including habitat, food availability and social groupings. When competition between males for females is strong because of the clumped distribution (Greenwood 1980; Shields 1987), male-biased dispersal is predicted as males may have reduced competition for mates elsewhere. In contrast, when competition for other resources is stronger, such as when territories become available infrequently, the sex that is most likely to inherit the territory of the senescing parent should be less likely to disperse (Greenwood 1980; Ronce *et al.* 1998).

As is the case with many taxonomic groups, age is an important determinant of dispersal. Juveniles of many lizard species may disperse further than adults (Olsson et al. 1996, 2003; Sumner et al. 2001; Ujvari et al. 2008; Clobert 2012). In the prickly forest skink, *Gnypetoscincus queenslandiae*, the dispersal distance moved between successive years decreases with age (Sumner *et al.* 2001). For many lizard species natal dispersal typically does not exceed breeding dispersal by much (Olsson *et al.* 1996; Clobert 2012). Sex biases in dispersal may also affect juvenile dispersal, for example male juvenile Swedish sand lizards, *Lacerta agilis,* and Common lizards, *Zootoca vivipara,* dispersed twice as far as females (Clobert *et al.* 1994; Olsson *et al.* 1996).

There is also a strong relationship between mating systems and the type of dispersal. Polygyny is often associated with male biased dispersal because of the patchy occurrence of females and competition for access to mate. Many lizard species are polygynous and, as expected, male-biased dispersal is relatively common (Clobert et al. 1994; Olsson et al. 1996; Chapple et al. 2005; Ujvari et al. 2008; Zani et al. 2009). Female biased dispersal may occur when females perform mate searching for multiple mating opportunities (e.g. Olsson & Shine 2003). In truly monogamous systems, no bias in dispersal is predicted without either resource competition or inbreeding avoidance, and empirical evidence supports this (Shields 1987; Chapple et al. 2005). Not much is mentioned of polyandrous systems in the literature, assumedly because of their infrequency in nature. Patterns of sex-biased dispersal are less apparent for lizards compared to mammals and birds because many species of lizard have female-biased dispersal (e.g. *Niveoscincus microlepidotus,*

Olsson & Shine 2003; *L. agilis,* Olsson et al. 1996), while others have no consistent sex-bias (e.g. *Z. vivipara*, Clobert et al. 1994; *Egernia whitii*, Chapple & Keogh 2005) and still others have male-bias dispersal (e.g. Ujvari et al. 2008). This may well be the result of the ubiquity of the multiple mates mating system and absence of well-defined social structure (Uller et al. 2008).

1.3 PROXIMAL CAUSES FOR DISPERSAL

Many possible cues and causes for dispersal have been identified to date and there is large variation between species and between populations within species. A dynamic dispersal behaviour is favoured in temporally variable environments (Holt et al. 2001). Abiotic environmental change, including habitat fragmentation (Driscoll 2004; Driscoll et al. 2005), resource enrichment or depletion, climatic change and habitat degradation (Halpin 1987; Clobert *et al.* 1994; Schowalter 2006), will influence dispersal behaviour. Intra- and inter-population dynamics may change as a result of changes in resource availability. Because population demographic attributes are important regulators of dispersal through fitness implications, dispersal behaviour will change as well (Clobert et al. 1994; Sorci et al. 1994; Milne et al. 2002; Meylan et al. 2002; Cote et al. 2007; Vercken et al. 2012).

Changes in the biotic environment may also occur freely of the abiotic environment. For example, the invasion fronts for alien invasive species may be led by novel phenotypes that facilitate dispersal (e.g. Cane toad, Urban et al. 2008) or as a result of facilitation by other organisms (Lescano 2010; Karsten *et al.* 2013) and indirectly via ecosystem modification (Urban et al. 2008; Nuzzo et al. 2009; Lescano 2010; Duckworth et al. 2010; Aplet 2011). Furthermore, a range of population dynamics concerning, for example, social interactions, individual health and personality, maternal conditions and social history can also influence dispersal (Greenwood 1980; Shields 1987; Sorci et al. 1994; Ronce et al. 1998; Gardner et al. 2001; Cote et al. 2007). There are various possible proximal cues for dispersal and as such, causes often need to be considered case-by-case.

2 DISPERSAL AND GENE FLOW

Most species do not have a continuous and widespread distribution but instead occur as meta-populations across heterogeneous environments (Wiens 1997; Hanski 1998; Holt et al. 2001). The connectivity between populations has important implications for genetic structure because it mitigates divergence caused by local adaptation and genetic drift (Wright 1931; Hanski 1998). Population genetic structure therefore, can be a good indicator of the degree of gene flow and is often used to make inferences regarding migration rates, both contemporary and historic (Sokal et al. 1989; Whitlock et al. 1999; Raybould et al. 2002; Hey 2010). Gene flow estimates as an indirect measure of dispersal have become popular because of the ease with which genetic data can be collected compared to real-time dispersal measures (Bossart et al. 1998). Furthermore,

genetic data covers a temporal span allowing it to include rare long distance dispersal events and historic signals which are difficult to detect with real-time measures (Whitlock et al. 1999; Nathan et al. 2012). The earliest estimates of genetic connections between populations were based on Wright's F statistics for genetic differentiation (Wright 1931). Although such estimators have received a fair deal of criticism due to the plethora of assumptions required (Whitlock et al. 1999; Raybould et al. 2002) they remain useful for gauging dispersal rates. On-going improvements to statistical methods have relaxed some of the previous assumptions needed, thus making models more versatile (e.g. Cockerham & Weir 1993; Beerli & Felsenstein 1999; Paetkau et al. 2004; Faubet et al. 2007). For example, initially the relationship between genetic structure and migration rates could only be explained by one of two mathematical models, the Island model (Wright 1943; Whitlock et al. 1999) and the Isolation-by-distance model (Wright 1943, 1946). A noteworthy assumption of both models is that all populations may only exchange individuals directly, though rates may vary depending on geographical separation (Raybould *et al.* 2002). The stepping stone model (Kimura et al. 1964) is a subsequent development that highlights indirect gene transfer between populations and is more useful for species with continuous distributions (Raybould *et al.* 2002).

Although describing genetic structure for many species has assisted in suggesting historic causes for changes in gene flow, it still fails to describe the mechanisms that 'translate' dispersal into gene flow. To discuss mechanisms one would need to compare real-time measures of ecological dispersal and genetic measures of dispersal. As a result of the different data types and analytical models used for investigating dispersal from a top-down or bottom-up approach, there is some disjunction between estimates (Smouse et al. 1999; Peakall et al. 2003). A suite of individual-based genetic statistics have been developed that, in part, address the disjunction between migration rate estimates from gene flow and estimates of dispersal from real-time measures of movement. By using the individual as the unit of measure, the analyses allow investigations to be conducted on small temporal and spatial scales- the same scale at which ecological movements occur- while still under a genetic framework. Such analyses include parentage and kinship analyses for discussing mating systems and parent-offspring distances (Keogh et al. 2007; Uller et al. 2008), population assignments for assessing the movement of individuals between populations (Pritchard et al. 2000; Wilson et al. 2003; Keogh et al. 2007; Hoehn et al. 2007), and the use of neighbourhood size estimates with an individual-based genetic difference regression to estimate dispersal distances from genetic data (Rousset 1997, 2000, 2004; Watts *et al.* 2007). Although the disjunction between large scale and local-scale methods persists somewhat, individual-based analyses are now frequently used and are proving extremely useful for addressing ecological questions regarding dispersal and population connectivity.

3   STUDY RATIONALE

Globally, biodiversity is threatened by increasing levels of habitat fragmentation and degradation, and the additional concerns of anthropogenic global climate change (Collingham et al. 2000; Fahrig 2003; Driscoll 2004; Driscoll et al. 2005; Midgley et al. 2006; Deutsch et al. 2008; Sinervo et al. 2010; Sutherland et al. 2011). Effective conservation measures, including decision making, implementation and management are dependent on accurate and robust information on the biology of the species concerned, yet for the vast majority of species such information is grossly lacking. Dispersal and gene flow are important facets of both the genetic and the behavioural responses that a species may exhibit under changing environments, making them extremely important considerations for conservation. For example, patterns in genetic structure could inform conservation priorities within a species and knowledge of dispersal propensity could influence plans for conservation areas and be used in predictions of range contractions or expansions under different environmental change scenarios (Midgley et al. 2006; Da Silva 2013; Duckett et al. 2013).

While southern Africa has the third most biologically rich lizard fauna worldwide (Branch 1998; Bates *et al.* 2014), it still lacks the research necessary for the effective conservation of many of its species. Studies of demographic parameters, including dispersal, are notably uncommon but are necessary for addressing current and future conservation concerns. An appropriate understanding of the species dispersal biology of any species, firstly, requires the identification and delimitation of species within species complexes and the ear-marking of evolutionary distinct lineages within species (Fraser et al. 2001). Evaluation of contemporary genetic connectivity may be addressed only when these lineages have been identified. Secondly, gene flow between populations within clades and between clades needs to be evaluated. Lastly, local-scale and real-time measures of dispersal are needed to compare to the regional, intra-population levels of migration. With such information, the influence of dispersal on genetic structure may be examined. In this thesis, a southern African lacertid lizard is used to examine the relationship between genetic structure and dispersal by considering the latter two areas of interest discussed above, i.e. the gene flow between populations and between evolutionary lineages within the species.

4   STUDY SPECIES: *PEDIOPLANIS LINEOOCELLATA*

The southern African endemic Spotted sand lizard, *Pedioplanis lineoocellata*, is a small (~44-57mm snout-vent length) lacertid lizard with a wide-spread distribution across South Africa and Namibia, possibly extending far into Botswana as well. The species was selected because in comparison to other southern African lizards, there is a fair amount of information available on the phylogenetic history, foraging behaviour and population demographic parameters (discussed below). This provides some biological context for proposing hypotheses and discussing the results obtained.

4.1 PHYLOGENETIC HISTORY OF *PEDIOPLANIS LINEOOCELLATA* (FAMILY: LACERTIDAE)

Lacertids are common in much of the Old world and occupy diverse habitats (Fitzsimons 1943; Branch 1998; Harris et al. 1998). Lacertids of the sub-family Lacertinae include the African-Arabian and the more ancestral Eurasian clades. African taxa are thought to be more recently derived (Harris *et al.* 1998) and the southern African lizards, specifically, have an evolutionary history strongly affected by the aridification during mid-Miocene (Lamb et al. 2003). There are at least eight genera and 37 species of lacertid lizards in southern Africa (Branch 1998; Conradie *et al.* 2012) and the relationship of many of the species has been investigated (Harris *et al.* 1998; Makokha *et al.* 2007; Hipsley *et al.* 2009; Conradie *et al.* 2012; Edwards *et al.* 2013).
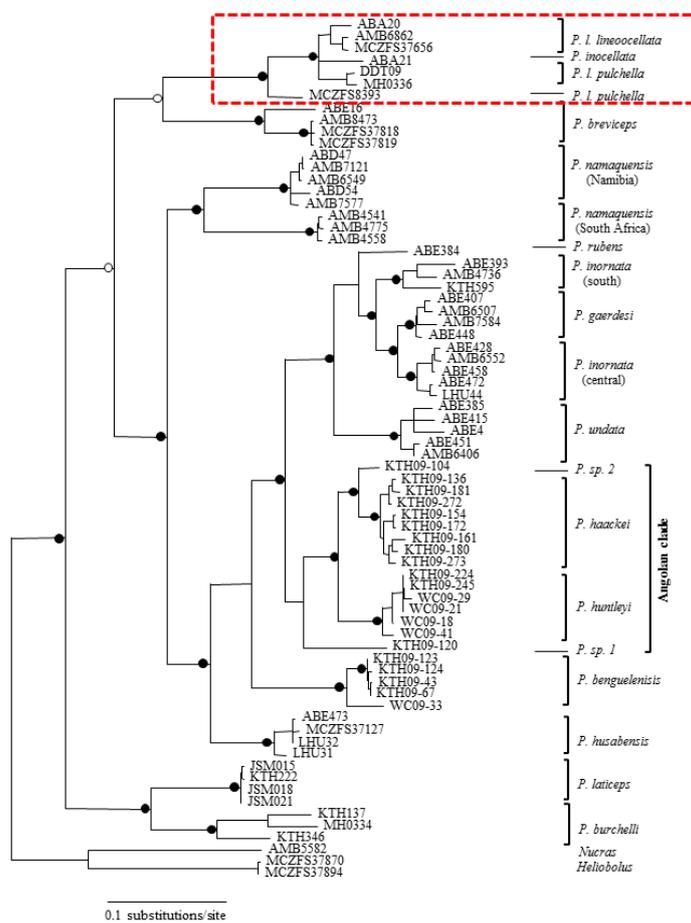
Figure 1.1: Bayesian inference phylogram for *Pedioplanis* species reproduced from (Conradie *et al.* 2012). Posterior probabilities [0.95] and maximum likelihood bootstrap [75%] support indicated by black circles at the nodes. Open circles indicate only maximum likelihood support. *Pedioplanis lineoocellata* are indicated by the red block.

Pedioplanis is one of the few genera for which taxonomy has been investigated (Makokha *et al.* 2007; Conradie *et al.* 2012; Edwards *et al.* 2013), the others being *Meroles* (Harris et al. 1998; Lamb et al. 2003; Edwards et al. 2013) and *Nucras* (Edwards 2013). *Pedioplanis* forms a well-supported monophyletic clade (Figure 1.1), within which *P. lineoocellata* is thought to be more recently derived than *Pedioplanis laticeps, P. burchelli* and *P. breviceps* (Makokha *et al.* 2007; Edwards *et al.* 2013). Within *P. lineoocellata* the former subspecies, *P. lineoocellata lineoocellata, P. l. pulchella* and *P. l. inocellata*, have been synonymised due to paraphyly (Edwards 2013). Edwards (2013) found evidence for four, potentially five, previously unknown divergent clades within *P. lineoocellata* based on mitochondrial DNA (mtDNA) but it is unclear if the clades represent separate species (Edwards 2013).

Morphological convergence has occurred as a result of similar dietary and local environmental selective pressures for some southern African lacertids (e.g. *Meroles squamulosus* and *Ichnotropis* species, *Australolacerta australis* and *Vhembelacerta rupicola*) (Edwards *et al.* 2012, 2013). In *P. lineoocellata* different morphologies have been associated with the former sub-species, including differences in colouration and scalation (Fitzsimons 1943; Branch 1998; Edwards 2013; Bates *et al.* 2014) but morphological differences within the species do not seem to relate to broad-scale habitat or phylogenetic lineages and is as yet unexplained (Edwards 2013).

4.2 FORAGING, MOVEMENT AND TERRITORIALITY OF *PEDIOPLANIS LINEOOCELLATA*

The genus *Pedioplanis* has active, ambush and mixed foraging modes (Cooper et al. 1999; McBrayer et al. 2009) and *Pedioplanis lineoocellata* has been characterised as a mixed foraging mode lizard (McBrayer et al. 2009). Variation in foraging mode is tightly correlated with morphology and physical performance such as speed and movement mechanics (e.g. McBrayer & Wylie 2009). Sit-and-wait foragers and those with mixed foraging modes are associated with evolutionary transitions toward long limbs to improve sprint speed while active foragers have shorter limbs and narrower pelvises, which improves manoeuvrability and endurance (e.g. Mcbrayer & Wylie, 2009). *Pedioplanis lineoocellata* has a lower endurance (as cited in Huey et al. 1984) but greater capacity for bursts of speed and a substantially lower maximal oxygen consumption compared to a sympatric active forager, *Heliobolus lugubris* (Bennett et al. 1984). Work by Bennett et al.(1984) found no significant differences in muscle physiology between the two species but did demonstrate a greater anaerobic capacity in *P. lineoocellata* associated with the capacity for bursts of speed. They found a smaller heart mass and hematocrit volume as well, which explains the lower endurance performance. This would suggest a physiology optimised for short burst of activity rather than endurance. More recent work indicates that at the whole-organism level, it is the combination of muscle fatigue resistance and, respiratory and circulatory function that determine stamina (Vanhooydonck *et al.* 2014). Although I could not find any literature on comparisons between foraging modes and dispersal biology, different physiological adaptations

may result in different dispersal biology based on endurance and daily movement. In birds and mammals, species with large home ranges often have greater dispersal distances (Bowman et al. 2002; Bowman 2003), while in lizards active foragers have greater home ranges compared to ambush foragers (Verwaijen et al. 2008b). From this, it might be expected that dispersal propensity is related to foraging mode.

Most movement studies for *P. lineoocellata* have focused on the influence of foraging behaviour on home range dynamics. Wasiolka and colleagues (Wasiolka 2007; Wasiolka et al. 2009b, 2009a) found foraging movements to be more frequent in degraded habitats (open structure because of proportionally greater bush/shrub cover and bare ground) compared to non-degraded habitats. Blumroeder et al. (2012) found that only an interaction of prey-habitat structure affected movement behaviour in *P. lineoocellata*. The difference in behaviour could indicate the interaction of costs and benefits of foraging modes under different habitat structures (Blumroeder *et al.* 2012). Individuals may adjust their home range in response the availability of shelter, food, water and mates within an environment, and this would occur primarily through modification of daily activities especially foraging movements (Perry et al. 2002; Blumroeder et al. 2012). Home range size estimates by radio-tracking adult male *P. lineoocellata*, agreed with a more active foraging mode in degraded habitats (Wasiolka *et al.* 2009b) as lizards travelled greater distances daily and home-range sizes were three-fold larger (209 m^2 vs. 646 m^2) (Wasiolka *et al.* 2009b). Such changes indicate behavioural plasticity and the potential for behaviour to buffer against environmental change. Other dynamics of the switch such as changes in energy expenditure, prey encounter rates, predation risk etc. may have important fitness implications but are less obvious and have yet to be investigated.

Considering that *Pedioplanis lineoocellata* has relatively limited endurance and aerobic scope, the costs of traveling great distances may have large fitness costs. Transitions across long stretches of unfavourable habitat, such as might be encountered during dispersal, might incur similar costs. While selection should favour philopatry in a situation where the somatic costs are large, behavioural flexibility could facilitate movement across unfavourable terrain and facilitate dispersal. Other factors important for dispersal remain to be considered; such as territory inheritance, breeding systems, social interactions and inbreeding avoidance.

In this study, I use microsatellite data and demographic data to examine dispersal and levels of gene flow for *Pedioplanis lineoocellata*. The thesis is divided into three data chapters. The first data chapter, Chapter 2 is an extension of preliminary work describing nine microsatellite markers used in this thesis, however the chapter examines additional populations. Chapter 3 focuses on examining the geographic patterns of genetic differentiation across the distribution of the two most wide-spread mtDNA lineages within the species. Gene flow between populations is estimated and the possibility of hybridization at the Loeriesfontein area is addressed. In Chapter 4 annual dispersal distances are estimated using demographic and genetic estimates of population density, respectively, and the results are compared. Genetic data are examined for evidence of an isolation-by-distance pattern that should result from limited dispersal. I conclude the thesis by summarising

the results of the three data chapters and discussing further developments needed to better understand the genetic connection between clades and the possible uses for such information.

Chapter 2 Isolation and Testing of Novel Microsatellite Loci for the Wide-spread Southern African Endemic Spotted Sand Lizard, *Pedioplanis lineoocellata*

1 INTRODUCTION

Pedioplanis lineoocellata is a lacertid lizard endemic to southern Africa but with a wide-spread distribution across several countries in the region, including Namibia, Botswana and South Africa (Branch 1998; Bates *et al.* 2014). The species also occupies a variety of habitat types, ranging from fynbos to Nama Karoo and savanna (Branch 1998). Recent phylogenetic work has invalidated the formally recognised sub-species but has also found no correspondence of the observed morphological variation to biome type or mitochondrial lineage (Fitzsimons 1943; Makokha *et al.* 2007; Edwards 2013). The most detailed phylogenetic examination within the species by Edwards (2013) uncovered four mitochondrial lineages and highlighted a possible region of contact between the two most wide-spread lineages, Clade A and Clade B (Chapter 1). The discordance between morphological variation in scalation, ocelli patterns and body dimensions to the mitochondrial lineages remains to be explained. Similarly, further detail regarding the area of contact between clades and the possibility of hybridisation requires more genetic work at population level in these areas.

Nuclear and mitochondrial gene regions sequenced for phylogenetic studies are often slow evolving and serve better to answer questions of deep-rooted evolutionary history rather than contemporary genetic scenarios at a population level (Avise 1989; Cho *et al.* 1995; Fritz *et al.* 2008). Microsatellite markers, in contrast, have high mutation rates and undergo crossing over and independent assortment during sexual reproduction making them hyper-variable and popular for small temporal and spatial scale genetic analyses (Gardner et al. 2000; Ellegren 2004; Selkoe et al. 2006). Microsatellites and similar hyper-variable genetic markers have been successfully used to address a diverse range of evolutionary questions including those regarding introgression and hybridisation (Tokarska et al. 2009; Schulte et al. 2012; Miraldo et al. 2013), contemporary gene flow, migration rates, sex and age biases in dispersal (Sumner et al. 2001; Gardner et al. 2001; Gold et al. 2001; Peakall et al. 2003; Clark et al. 2008; Schregel et al. 2012), mating systems (Laloi et al. 2004; Chapple et al. 2005; Holzman et al. 2009) and selective sweeps (Li et al. 2010).

Microsatellites are not without their shortcomings. One of the most common problems with microsatellite markers is the presence of non-amplifying alleles or null alleles. Null alleles may arise due to poor primer annealing because of nucleotide sequence divergence of flanking regions caused by point mutations or insertion-deletion events (Dakin et al. 2004). Most authors have simply ignored the presence of null alleles in their datasets but the presence of null alleles may produce an excess of homozygotes above that predicted from the HWE expectation (Dakin et al. 2004; Björklund 2005). Heterozygosity estimates are affected in two ways. Firstly, when null alleles occur in a heterozygote, the null allele does not amplify making the individual appear as a homozygote. Secondly, when null alleles are present in the homozygote state, both alleles fail to amplify resulting in missing data (Wagner et al. 2006; Lemer et al. 2011). Consequences of

using markers that have null alleles include increased inaccuracies for estimates of genetic differentiation and significant bias in results of parentage analyses (Dakin et al. 2004; Lemer et al. 2011). Some taxa have demonstrated high levels of microsatellite priming site variability (e.g. lepidopterans, nematodes and marine molluscs; discussed by Lemer et al. 2011) and thus null allele susceptibility. Lacertid lizards appear to be yet another group where null alleles are a common problem (e.g. Richard & Thorpe 2000; Bloor et al. 2006, 2010; Li et al. 2012; Miraldo et al. 2013).

The first steps in the application of microsatellite markers in a genetic study should include examining the dataset for the presence of null alleles and other possible causes for deviations from Hardy-Weinberg equilibrium (HWE). Eleven microsatellite markers were recently developed for *Pedioplanis lineoocellata* specifically for the present study (Tolley et al. 2014b, Appendix Information A 1). Screening the microsatellite library using a single population showed that five of the eleven loci were out of HWE and for two loci there was evidence of relatively elevated null allele frequencies. After the removal of the two loci, inbreeding and random genotyping failures better explained the observed deviations from HWE expectation (Appendix Information A 1). Although there was evidence for null alleles, deviations from HWE at several loci could also suggest a population-level cause. It is therefore necessary to test for other possible causes such as high levels of inbreeding and a recent genetic bottleneck.

In this chapter the new microsatellite markers were used to examine genetic diversity at additional populations of *P. lineoocellata* from within Rooipoort Nature Reserve and several other localities from across South Africa that were not included in the preliminary work. Data from each sample site were examined for evidence of inbreeding, the presence of null alleles and genetic bottleneck events. Furthermore, the basic analyses of testing for linkage disequilibrium and estimating allelic diversity were also performed. These preliminary steps are a necessary pre-requisite for the use of microsatellites in addressing ecological and evolutionary questions, particularly those pertaining to within- and between-populations processes, such as the questions addressed in the remaining chapters of this thesis.

2 METHODS AND MATERIALS

2.1 SAMPLE COLLECTION

Samples for the screening of null alleles, Hardy-Weinberg equilibrium and inbreeding were obtained from field collections and from the HerpBank (Herpetological DNA Bank, South African National Biodiversity Institute), which included multiple sampling areas across South Africa and Namibia. The available DNA samples were pooled into 'sample sites' based on local proximity of ≤ 2km pairwise geographical distance between samples. This criterion created eight 'sample sites' with at least five individuals each, resulting in a dataset with a total of 120 (out of 191 available individuals, Figure 2.1 and

Appendix Table A 1). Seventy-four samples were from four sites at Rooipoort Nature Reserve (RNR) and 46 samples from four other localities across South Africa. The remaining 71 samples from 37 localities did not group according to the criterion but were retained for other analyses in the next chapter (Chapter 3). Of the eight sample sites, only four sites had more than 15 samples each (HART, GANS, RNR1 and RNR2); sample sizes better suited for microsatellite analyses. These four sites were used for simultaneous estimation of null alleles and inbreeding and for the bottleneck tests (see below) while the remaining four sites were primarily used for descriptive statistics.
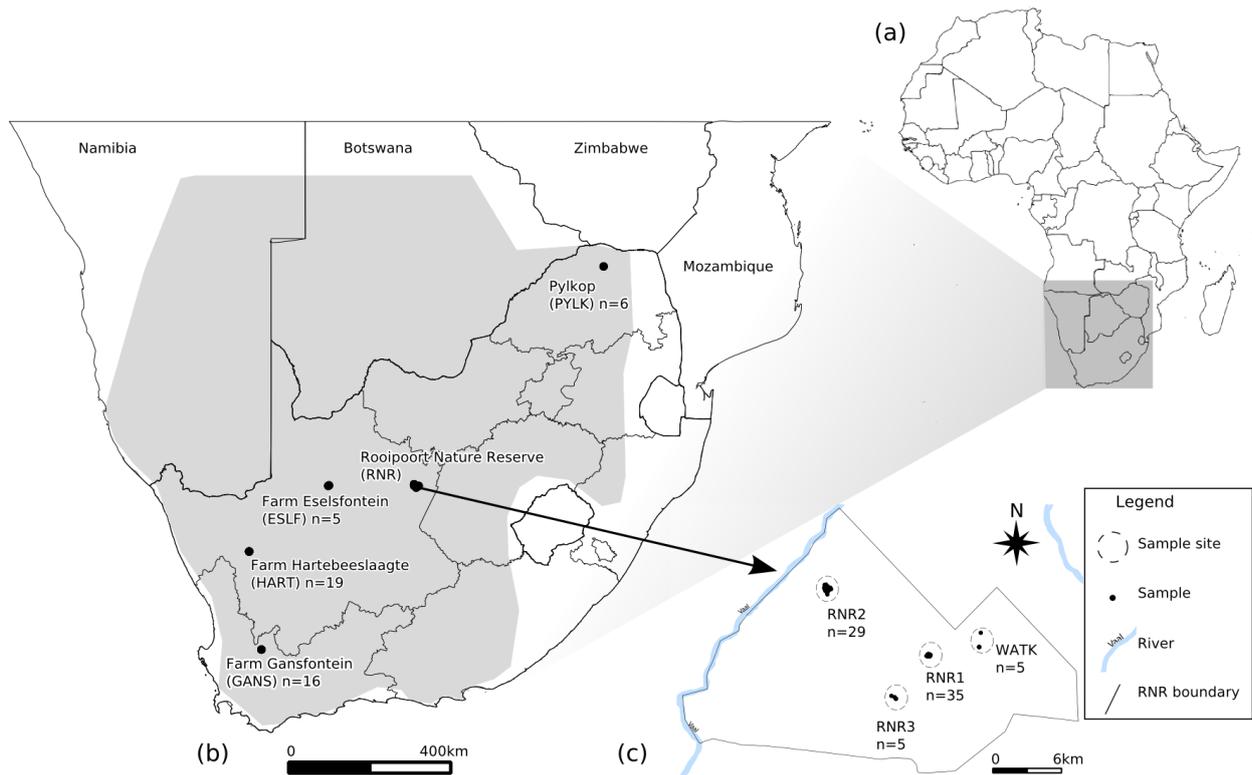
Figure 2.1: Map of the eight sample sites for *Pedioplanis lineoocellata*. (a) Inset indicating the position of the study region within Africa. (b) Map of the Southern African study region with the eight primary sample sites indicated (as black dots) including their abbreviated names (in brackets) and the respective sample sizes (n). Species distribution polygon (shaded grey) defined by occurrence records from GBIF, HerpBank (SANBI) and SARCA (Bates *et al.* 2014).(c) Sampling sites of 2km diameter (circled) within Rooipoort Nature Reserve (RNR), Northern Cape, South Africa.

2.2 LABORATORY WORK

Only nine of the eleven microsatellite loci included in the preliminary work were considered here (Appendix Information A 1) because Peli021 and Peli030 were removed due to inconsistent amplification success across the sample sites. Total genomic DNA was extracted by salt extraction and Proteinase K protocol (Aljanabi et al. 1997). Standard Polymerase chain reaction (PCR) protocols were followed in 10μl

reactions according to the general protocol of Feldheim et al. (2010), but with locus-specific annealing temperatures and MgCl$_2$ concentrations (Table 2.1). The PCR cycling protocol included a denaturation step of 94 °C for 1 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at primer specific temperatures for 30 s and an extension at 72 °C for 45 s then a final extension at 72 °C for 10 min. Two μl of PCR product were visualized on an ethidium bromide stained 2% agarose gel and verified with a Gelpilot 100bp ladder (Qiagen). Products were profiled (monoplexing or poolplexing; Guichoux et al. 2011) at the Central Sequencing Facility, Stellenbosch University using an ABI 3130 Prism (Applied Biosystems, Foster City, California, USA). Alleles were scored using GeneMapper v 4.0 (Applied Biosystems) and binned with MSatAllele v.1.03 package (Alberto 2009) under the R environment v. 3.0.0 (R Core Team 2013). All loci except Peli018, Peli039 and Peli050 showed signs of fragment sizes inconsistent with the expected repeat lengths, presumably either due to insertion-deletion events (indels) in the flanking regions or scoring/profiling errors (see Appendix Figure A 1). Where these deviations were observed, the closest "expected" allele size was chosen to be the repeat sequence. All other values were pooled around these 'expected' repeat sequences. Alleles were grouped using $Px = \{x - 1, x, x + 1, x + 2\}$ for tetranucleotide repeats and $Px = \{x - 1, x\}$ for dinucleotide repeats, where Px refers to the pooled alleles that were labelled size 'x'.

2.3 ANALYSES

Allele frequencies, size range and diversity were estimated in Arlequin v3.5 (Excoffier et al. 2010). The modified Exact test was used to test for Hardy-Weinberg equilibrium (HWE) (Guo et al. 1992) as implemented in Genepop on the Web v.4.2 using the default settings for the Markov chain (Raymond et al. 1995; Rousset 2008). Evidence for linkage disequilibrium between loci was examined using a log likelihood ratio test in Genepop on the Web v.4.2 (Raymond et al. 1995; Rousset 2008) with sequential Bonferroni corrections (Rice 1989).

The level of inbreeding within sites was estimated using Wright's Inbreeding coefficient, F_{IS}, for all eight sample sites (Wright 1943; Slatkin 1991; Hartl et al. 1997) as estimated in Arlequin. However Wright's F_{IS} may be substantially upward biased if null alleles or genotyping errors are present in the dataset (e.g. Björklund 2005). Likewise, high levels of inbreeding or the inclusion of closely related individuals in a study sample can cause an upward bias in estimates of null allele frequencies (Chybicki et al. 2009; Campagne et al. 2012). Standard tests for null alleles and other genotyping errors do not allow for this inter-dependence (Van Oosterhout *et al.* 2004; Campagne *et al.* 2012). Therefore inbreeding coefficients were estimated and data were screened for the presence of null alleles by simultaneous estimation of inbreeding coefficient (F), null allele frequencies (n), and random genotyping failure (b) with model selection under Bayesian Inference using the program INEst v2 (Chybicki et al. 2009, 2011). An individual inbreeding model (IIM) was

employed with models including all combinations of parameters F, n, and b. Models were compared based on deviance information criterion (DIC) values (Burnham & Anderson 2002; Mason *et al.* 2012). The Monte Carlo Markov chain (MCMC) was run with 500,000 cycles, a 20% burn in, saving parameters every 500 cycles. The analysis was only performed on sample sites with more than 15 samples each, i.e. HART, GANS, RNR1 and RNR2.

I tested for a population bottleneck or founder effect as a cause for the deviations from HWE using two different tests. Firstly, the M-ratio (or G-W statistic) was estimated in Arlequin and values smaller than 0.68 are indicative of a recent bottleneck event (Garza et al. 2001; Excoffier et al. 2010). The possibility of a bottleneck was further tested by examining for heterozygosity excess. Heterozygosity excess is expected when heterozygosity is estimated from the number of alleles assuming a mutation-drift equilibrium, compared to the heterozygosity observed from the actual data, should there be a bottleneck signal (Cornuet et al. 1996; Piry et al. 1999). The heterozygosity test was performed using a Wilcoxon's sign-rank test and a Sign test. Default settings were retained including 1000 iterations. Both simple mutation model (SMM) and infinite allele model (IAM) mutation models were included as they represent two extremes of the mutation models between which the most appropriate model is likely to fall (Cornuet et al. 1996). Analyses were implemented using BOTTLENECK v1.2.02 (Cornuet et al. 1996; Piry et al. 1999). All tests were performed with and without samples containing missing data but there were no meaningful differences between the results, thus results presented are that including samples with missing data.

3 RESULTS

The proportion of missing data per locus ranged from 0.04 (Peli001) - 0.22 (Peli050), and the total proportion (166 of 1818 alleles) was only 9.0% ±6.0 across sample sites (mean ± S.D.). Site PYLK had the largest proportion (0.20) whereas GANS had the lowest (0.007) (Table 2.2). All nine loci were polymorphic across the combined dataset for all eight sample sites and the number of alleles ranged from eight to 22 (Table 2.1). Within sample sites, the number of alleles varied from one to 14 and Peli050 was monomorphic at three of the eight sites; PYLK, HART, WATK (Appendix Table A 3). The number of alleles per locus increased with increasing sample size (results not shown) indicating insufficient samples at the sites with only five samples.

Five of the eight sites were not in HWE (Table 2.2), notably only the sites with five samples were in HWE but sample sizes may be too small to determine if the results are valid. There was homozygote excess at six sites and across all but one (Peli001) locus. No loci were found to be in linkage disequilibrium after sequential Bonferroni corrections within each sample site (results not shown).

Table 2.1: Summary data for the nine microsatellite loci tested for *Pedioplanis lineoocellata* for 120 individuals combined. Information given: name and primer GenBank accession number (Locus), primer sequence (F: forward, R: reverse), fluorescent label used (Label), microsatellite repeat motif (Motif), annealing temperature (°C), MgCl$_2$ (given in mM), type of Taq (Taq: GOLD; Super-Therm Gold DNA Hot-Start polymerase (Southern Cross Biotechnology), SST; Super-Therm DNA Polymerase (Southern Cross Biotechnology), GT; Promega Gotaq ® DNA Polymerase (Promega Corp.), number of alleles (N_A) and the size range in base pairs (*r*).

Locus	Primer sequence (5'-3')	Label	Motif	(°C)	MgCl$_2$	Taq	N_A (*r*)
Peli001 KF302081	F: TGAGATGGGTATGCTGTTGC R: CCCCCTACTTAGCATTTATTCAG	FAM	(TAAC)5	60	1	SST	13 (109-161)
Peli005 KF302082	F: GGGTGGATTTTAGTCCACATT R: CGTGTTTGTTCCATGACAGA	FAM	(TGTA)10	56	1.5	SST	15 (135-211)
Peli018 KF302083	F: TGGTTTCTCCTCGCACATA R: CAACGTGTAACAGACATGCAG	HEX	(TGTC)11	60	1.5	SST	15 (76-132)
Peli020 KF302084	F: AGTCCCCCAACAACAGCTT R: TGGTTAGTGGATCTGGGAAAA	NED	(TAAG)5	58	1.5	GOLD	11 (117-157)
Peli022 KF302086	F: GAGAAAGCAAACGGAGTGAG R: GATTGTTGTCTGAGGCTTGG	HEX	(TATC)11	56	1.5	SST	18 (165-237)
Peli034 KF302088	F: AGATGGCTGCTGAAAACCA R: CCTCCATCAATGCTTCCTTC	HEX	(AC)18	56	1.5	SST	22 (133-185)
Peli039 KF302089	F: AGAAGTCAATGCAAGAGCAG R: ATTGCACATAACACTCAAGAGC	FAM	(TATG)9	58	1.5	GT	10 (105-173)
Peli048 KF302090	F: GCATCCACCATTTCTGACC R: CCAAAATTAGGTTGCCAGTG	HEX	(AGAT)10	56	1.5	GT	16 (130-218)
Peli050 KF302091	F: ACCTGGAACCCAAACCAAC R: GGCTGTAGATTGCATTTTTGC	HEX	(TG)6	58.5	1.5	SST	8 (110-162)

Simultaneous estimation of inbreeding coefficients, null allele frequencies and random genotyping failure rates showed that the best model contained a parameter for random genotype failure for all sample sites examined (Table 2.3) and the estimated proportions of genotyping failures were moderate across loci and sites (<0.10) with the exception of Peli018 for RNR1 and 2 (Figure 2.2).

INEst null allele frequency estimates for each locus were similar for sites RNR1, RNR2 and HART (Figure 2.2), particularly for HART and RNR2. These results were broadly comparable to the Van Oosterhout estimates (Appendix Table A 4) as in both cases Peli020, Peli034 and Peli039 had high null allele frequencies. Only GANS had no support for the presence of null alleles.

Table 2.2: Results from the Exact Test for Hardy-Weinberg equilibrium (p-value) for *Pedioplanis lineoocellata* per locus and per sample site. Values in **bold** indicate significant deviations from Hardy-Weinberg equilibrium at $\alpha = 0.05$. Sample sizes per population (*n*) and proportion missing data (Prop. MD) in each dataset are indicated. Chi-squared test statistics for overall Hardy-Weinberg proportions across loci indicated at the bottom of the table.

Locus\Sites	ELSF	GANS	HART	PYLK	RNR1	RNR2	RNR3	WATK
(*n*)	(5)	(16)	(19)	(5)	(35)	(29)	(5)	(5)
Prop. MD	0.088	0.007	0.029	0.201	0.071	0.050	0.044	0.022
Peli001	0.15	0.89	0.37	0.33	0.99	0.77	0.70	0.69
Peli005	**0.03**	**0.01**	**<0.001**	0.30	0.59	0.32	1.00	0.24
Peli018	0.69	0.26	**<0.001**	1.00	0.93	0.78	1.00	0.15
Peli020	0.36	1.00	1.00	1.00	**0.01**	**<0.001**	0.62	1.00
Peli022	0.62	0.98	0.87	1.00	0.95	**0.04**	0.29	0.06
Peli034	0.66	**0.02**	0.06	1.00	0.06	**<0.001**	0.24	0.37
Peli039	1.00	**<0.001**	**<0.001**	0.24	**<0.001**	**<0.001**	1.00	0.06
Peli048	0.62	**0.04**	0.35	0.08	0.15	**0.04**	0.15	**0.04**
Peli050	1.00	**0.01**	-	-	1.00	1.00	1.00	-
p-value across all loci	0.59	**<0.001**	**<0.001**	0.72	**<0.001**	**<0.001**	0.82	**0.05**
X²	16.08	49.52	∞	12.41	39.87	∞	10.89	26.65
d.f.	18	18	16	16	18	18	16	16

Inbreeding was an important parameter only at site RNR2 and GANS (Table 2.3), with comparatively elevated levels of inbreeding at GANS (0.11, 0.02 - 0.19 95% highest posterior density- HPD). Additionally, with no support for null alleles at GANS there would be no inflation of inbreeding coefficients. In contrast, Wright's Inbreeding coefficients were greatly inflated when estimated independently of null allele frequencies (Table 2.4). The average F_{IS} inbreeding coefficients varied strongly between loci and between sites but without obvious or consistent patterns. With the exception of PYLK, all sites had estimates that would have suggested inbreeding (F_{IS}>0.13) contributes to deviations from HWE, had this parameter been estimated without considering the effect of null alleles and random amplification failure. Instead, the combined estimates suggest that random genotyping failure, inbreeding and null alleles all likely contributed to the deviation from HWE observed.

Table 2.3: Model selection results from INEst analysis for the four *Pedioplanis lineoocellata* sample sites (a-d) with *n*>15. Model selection was based on the Δ Deviance Information Criterion (ΔDIC). Models with <2 unit differences from the best model are highlighted. Model parameter abbreviations: b- Random genotyping failures, f-Inbreeding, n-Null alleles. Average Log likelihood (LogL) and sample mean inbreeding coefficient; Avg(Fi) (lower-upper 95% highest posterior density) indicated.

(a) GANS

Model:	LogL:	ΔDIC	Avg(Fi)
fb	-382.0	-	0.108 (0.02-0.19)
nfb	-382.7	1.191	0.065 (0.00-0.16)
nb	-383.2	1.574	-
nf	-383.4	2.145	0.065 (0.00-0.15)
b	-385.9	3.034	-
n	-384.3	3.67	-

(b) HART

Model:	LogL:	ΔDIC	Avg(Fi)
nb	-354.9	-	-
n	-355.7	0.119	-
nfb	-355.5	2.054	0.056 (0.00-0.18)
nf	-356.2	2.418	0.063 (0.00-0.16)
fb	-359.9	14.164	0.258 (0.15-0.36)
b	-372.6	34.298	-

(c) RNR1

Model:	LogL:	ΔDIC	Avg(Fi)
nb	-906.4	-	-
nfb	-906.6	1.024	0.014 (0.00-0.05)
fb	-912.1	12.178	0.072 (0.02-0.12)
b	-916.8	17.237	-
nf	-925.2	34.592	0.014 (0.00-0.05)
n	-925.4	35.308	-

(d) RNR2

Model:	LogL:	ΔDIC	Avg(Fi)
nfb	-797.5	-	0.059 (0.00-0.13)
nb	-801.0	3.243	-
nf	-803.8	9.358	0.059 (0.00-0.12)
fb	-801.7	10.553	0.195 (0.13-0.26)
n	-808.3	14.13	-
b	-823.9	45.594	-
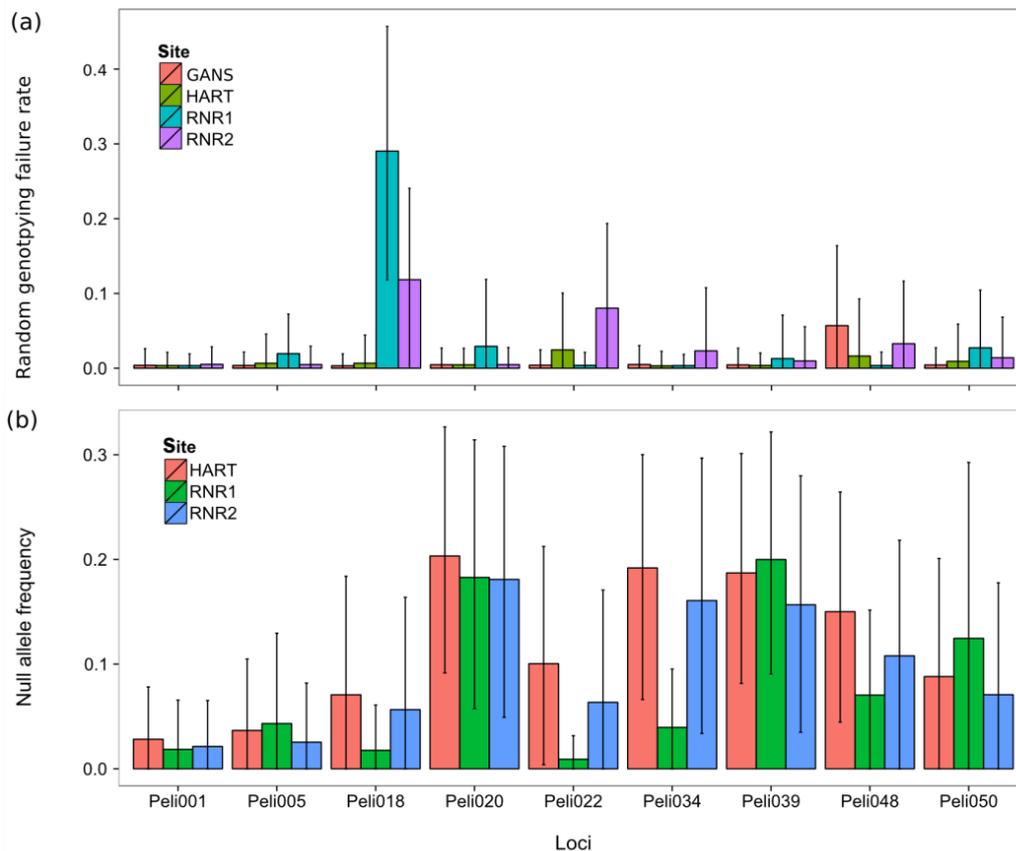
Figure 2.2: (a) Mean random PCR failure rate and (b) null allele frequency estimates across loci by sampling site for *Pedioplanis lineoocellata* determined using the best INEst models. The 95% highest posterior density (HPD) intervals indicated as error bars. GANS is excluded from (b) because there was no support for the presence of null alleles by the best INEst model.

There was no evidence for a genetic bottleneck from either the G-W statistics or heterozygosity excess at those sample sites with more than 15 samples each. The M-ratio for the four sites straddled the 0.68 threshold (0.64-0.77; Appendix Table A 3) suggesting possible bottleneck signals, however all two-tailed t-tests for means different from $M = 0.68$ were not significant (p>0.10, d.f.=8, t=-0.6-1.92). With fewer than ~25 samples per site the estimated M-ratio values may be lower than expected (Garza et al. 2001). This possibly explains the higher M-ratio values and lower variance for RNR1 and RNR2 sites compared to the remaining sites with lower sample sizes. For all sample sites, the M-ratios appear to have been influenced by strong between-loci variation and small sample sizes. The M-ratio for the remaining sites with less than 15 samples each was estimated as well. Values and standard deviations were comparable to that of the four sites with n>15 (Appendix Table A 3).

Table 2.4: Population specific F_{IS} values for *Pedioplanis lineoocellata* based on allele size variation (i.e. *Rho$_{IS}$*) per polymorphic locus. Information provided: sample size (*n*), average F_{IS} across all samples by locus (Across sites) and average F_{IS} across all loci by sample site (Across loci). Average F_{IS} across all samples and loci indicated in **bold**. "-" used to indicate monomorphic sites.

Locus	Across sites	ESLF	GANS	HART	PYLK	RNR1	RNR2	RNR3	WATK
	n	5	16	19	6	35	29	5	5
Peli001	0.110	0.825	0.108	0.368	-0.600	-0.354	0.325	0.143	0.221
Peli005	0.581	0.455	0.805	0.888	-0.689	0.251	0.711	-0.289	0.081
Peli018	0.194	-0.600	0.444	0.453	-0.143	-0.010	-0.007	0.143	0.950
Peli020	0.038	0.667	-0.186	-0.160	-0.175	0.318	0.447	0.200	-0.269
Peli022	0.066	0.019	-0.013	0.086	0.307	0.060	-0.115	0.501	0.715
Peli034	0.082	-0.053	-0.048	-0.099	-0.434	0.134	0.148	0.722	-0.353
Peli039	0.425	0.091	0.264	0.337	0.688	0.602	0.450	0.000	0.176
Peli048	0.143	0.009	-0.160	0.307	-0.404	0.209	0.321	0.516	0.245
Peli050	0.848	-0.667	0.963	-	-	-0.029	-0.149	-0.244	-
Across loci	**0.216**	0.311	0.158	0.287	-0.077	0.136	0.302	0.483	0.409

Bottlenecks were also examined using the heterozygosity excess test, which indicated a heterozygosity deficiency under simple mutation model (SMM) for all but one site (GANS). The Sign test was significant for heterozygosity deficiency for RNR1 and HART (Table 2.5) and the Wilcoxon sign-rank test was significant for RNR1, RNR2, HART, which is suggestive of a population expansion rather than a population bottleneck.

Under the infinite allele model (IAM), results were not significant for all sites, except GANS, which had significant heterozygosity excess for the one-tailed Wilcoxon test (p=0.014) suggesting a population bottleneck. Results from the M ratio were congruent with the heterozygosity excess estimates. The M-ratio was >0.68 for RNR1 and RNR2 and M<0.68 for GANS. Only HART had an M value (0.64) that did not agree with the evidence for heterozygosity deficit. Therefore there is no evidence of a population bottleneck at RNR1, RNR2 and HART, and only weak evidence for an event at GANS, possibly due to higher inbreeding coefficients.

Table 2.5: Results from the heterozygosity excess analysis (p values) for a bottleneck for *Pedioplanis lineoocellata* at the four sample sites with *n*>15. Analyses were performed under each of two mutation models (Model), Infinite Allele Model (IAM) and Simple Mutation Model (SMM). Results presented are for the Sign test and Wilcoxon sign-rank test for a one-tailed Heterozygote (H) deficit and excess and a two-tailed test for both. Significant (p<0.05) figures in **bold**.

Sites	Model	Sign Test	Wilcoxon sign-rank Test		
			H deficit	H excess	Two tailed
GANS	IAM	0.062	0.990	**0.014**	**0.027**
	SMM	0.097	0.212	0.820	0.426
HART	IAM	0.562	0.727	0.320	0.641
	SMM	**0.009**	**0.010**	0.994	**0.020**
RNR1	IAM	0.235	0.936	0.082	0.164
	SMM	**0.027**	**0.010**	0.993	**0.020**
RNR2	IAM	0.249	0.752	0.285	0.570
	SMM	0.114	**0.019**	0.986	**0.037**

4 DISCUSSION

Eleven microsatellite loci were screened for *Pedioplanis lineoocellata* at several sample sites in preparation for utilising these markers in population studies (Tolley et al. 2014b). In this study two loci (Peli021 and Peli030) were removed because of inconsistent amplification success across sample sites. Five of the remaining nine loci deviated from Hardy-Weinberg equilibrium (HWE) most likely due to the influence of null alleles, random genotyping failures and, in one population, inbreeding. There was however no support for a genetic bottleneck event at any site.

This study found evidence of null alleles at several loci and sample sites. This included two loci (Peli020 and Peli039) at RNR1, only one of which (Peli039) also had evidence for null alleles in the preliminary work (Tolley et al. 2014b; Appendix Information A 1). The detection of null alleles at Peli020 for RNR1 in this study is most likely due to the inclusion of additional samples not included in the preliminary work. In the preliminary work Peli039 and Peli021 were removed because of null allele frequency estimates with lower 95% HPD bounds above 0.00. After their removal, the best INEst model did not include null alleles as a parameter. Peli021 was absent in both this chapter and the preliminary work, and Peli030 was already removed from the data set in this chapter because of inconsistent amplification across sample sites. Further removal of the two loci with evidence for null alleles at RNR1 in this chapter would seriously reduce the available data. Furthermore, Peli039 did not have unusually high null allele estimates compared to other loci as the 95% HPD intervals for almost all estimates overlapped. More generally, the detection of null alleles at

loci was also inconsistent across sites so that some sites had additional loci with evidence of null alleles, namely, Peli034 and Peli048 for RNR2 and HART. Therefore, deciding which loci to remove based on the relatively elevated null allele frequencies at some loci may prove difficult because of the similar frequencies between loci but differing frequencies across sites.

Although the presence of null alleles is one explanation for the HWE deviation, population-level events such as genetic drift, non-random mating, selection at certain loci, sex-linkage of loci, recent migration or the Wahlund effect may have caused the HWE deviation (Hartl et al. 1997). The Wahlund effect- the loss of heterozygosity observed when samples from populations of different allele frequencies are pooled- was ruled out as a possible cause. Population substructure seems unlikely considering the small geographic scale at which the samples were collected at each site, <2km pairwise geographic distance between samples. There was also no support for a bottleneck at any population with both the statistics used. GANS had weak support for a population bottleneck however inbreeding coefficients were relatively high for GANS and may have biased result towards a bottleneck signal (Cornuet et al. 1996). Inbreeding, in general, was not supported either as only two of four sites in this chapter had inbreeding as a parameter in the best INEst model. Moreover inbreeding coefficients are relatively low at the latter two sites compared to estimates from other lacertids. The uncorrected estimates of inbreeding coefficients are comparable to upper estimates in many other lacertids [e.g. *Lacerta lepida*, *Lacerta nevadensis* (Miraldo *et al.* 2013), *Iberolacerta monticola* (Remón *et al.* 2013), *Lacerta viridis* (Böhme et al. 2011) and some *Podarcis muralis* populations (Schulte *et al.* 2012)]. Despite the comparable estimates, these other studies share issues with this study such as null alleles, possible hybridization and small sample sizes and therefore the other estimates are likely inflated as well (Olsson et al. 1996; Böhme et al. 2011; Miraldo et al. 2011, 2013; Remón et al. 2013). This further supports that inbreeding for *P. lineoocellata* is not unusually high and is unlikely to be a primary driver for the HWE deviations.

The results best support null alleles as the cause for the deviations from Hardy-Weinberg equilibrium. Correcting for the presence of null alleles should mitigate the biases null alleles introduce into a dataset (Wagner *et al.* 2006). In ~10% of studies, the affected loci are removed from the dataset but only a small fraction of the remaining 90% use statistical corrections to account for null allele presence (Dakin & Avise 2004). Because there were only nine loci available in this study, replacing any of the loci that had evidence for null alleles was not possible. The removal of loci without replacement could cause a reduction in statistical power and possibly the introduction of different/additional bias (Wagner et al. 2006; Chybicki et al. 2009). In cases such as this one, retaining the affected loci is preferable though it may result in increased error margins for estimates (Wagner *et al.* 2006). Some researchers advocate strongly for further laboratory work because of the affects that null alleles can have in parentage analysis and similar methods (Dakin et al. 2004; Lemer et al. 2011). However, the decision needs to be made according to the case at hand, and in

studies where divergence estimates are high, finding common microsatellite loci for all sample sites may not be practical or cost effective. Corrected allele frequencies may be used to mitigate any bias from null alleles. Corrections are performed by changing the excess homozygote profiles into heterozygotes with an unidentified allele, thus adjusting the allele frequencies to match that predicted under Hardy-Weinberg equilibrium (Van Oosterhout *et al.* 2004). However, there is no way of deciding which apparent homozygous individuals in a data set are truly homozygous and which possess null alleles. This precludes certain types of analyses from being used. Moreover, some analyses rely on a true deviation from HWE to detect the signal of interest, thus making data corrected to reflect a HWE misleading (Cornuet et al. 1996; Luikart et al. 1998a; Björklund 2005; Chybicki et al. 2009). For the remaining chapters, analyses that are not reliant on HWE assumptions were used when possible, however for some analyses these alternative options were not available. Any results of analyses that rely on HWE assumptions in this thesis are tentative and results are interpreted keeping in mind the caveat that null alleles may be present (see Chapter 3).

Chapter 3 COMPLEX SPATIAL GENETIC PATTERNS AND EXTENSIVE SECONDARY CONTACT IN THE WIDE-SPREAD SPOTTED SAND LIZARD (*PEDIOPLANIS LINEOOCELLATA*)

1 INTRODUCTION

A species' distribution may consist of partially isolated populations connected to some degree by dispersal and possibly gene flow, forming what is known as a meta-population (Wiens 1997; Hanski 1998). Meta-populations are ubiquitous in spatially heterogeneous habitats where suitable conditions are fragmented (Hanski 1998). The genetic differentiation between these populations may increase gradually as a result of interactions between local genetic drift, selective pressures and gene flow, further resulting in population structure (Hartl et al. 1997). Gene flow is an important regulator of genetic differentiation because it can counter-act the effects of genetic drift and selection. Although ecological dispersal, the physical movement of individuals, is important for establishing genetic exchange between populations, gene flow further requires individuals to survive and reproduce in their new environment (Raybould *et al.* 2002). Thus high genetic differentiation between populations may indicate low gene flow despite high levels of ecological dispersal or vice versa (Whitlock et al. 1999; Raybould et al. 2002; Nathan et al. 2012).

Present-day geographic patterns in genetic differentiation reflect the interplay of genetic processes, historic and contemporary. Contemporary influences of population genetics can be measured by various methods (Luikart et al. 1998a; Rooney et al. 1999; Paetkau et al. 2004; Coulon et al. 2004; Wang et al. 2009) while historic influences can only be inferred based on correlations between present-day genetic patterns and known historic environmental events (e.g. Holzhauer et al. 2006; Fontaine et al. 2010; Barlow et al. 2013). Knowing what limits dispersal will be important for identifying at what scales populations are expected to be of limited contribution to connectivity and undergo genetic differentiation or to identify what environmental change will most strongly affect connectivity. Furthermore, a better understanding of the contemporary relationship between ecological dispersal and gene flow can be useful for inferring historic genetic events.

The role of historic climate change on the evolution of many taxa is undisputed. In the northern hemisphere, glacial-interglacial cycles during the Plio-Pleistocene period (beginning ~2.8Ma) had notable influence in the divergence and subsequent speciation of many taxa (e.g. Pfenninger et al. 2003; Svenning et al. 2008; Miraldo et al. 2011, 2013; Zhao et al. 2011; Remón et al. 2013). During cooler glacial periods, ranges contracted into refugia, areas thought to have been more climatically hospitable during these times. There may have been several range expansion and contraction events because of climatic oscillations, which would have resulted in several opportunities for secondary contact (Pinceel et al. 2005; Remón et al. 2013; Miraldo et al. 2013). For some species, the periods of isolation were sufficient for speciation and even following secondary contact there has been limited evidence of hybridization (Fuchs et al. 2011; Miraldo et al. 2013).

Patterns of isolation are also present in southern hemisphere tropical and sub-tropical taxa but our understanding of the evolutionary response pales in comparison to the northern hemisphere temperate

Page | 28

species. There are large taxonomic gaps and even within the species examined, sampling is often only at the regional scale. Patterns in phylogeographic breaks are often similar between taxa, as has been found for sub-Saharan open-habitat species (e.g. Arctander et al. 1999; Lorenzen et al. 2010; Fuchs et al. 2011) and for many species from the south-western parts of South Africa [e.g. Rock Hyrax, *Procavia capensis* (Prinsloo et al. 1992; Visser 2013); Southern Scrub hare*, Lepus saxatilis* (Kryger et al. 2004); rock elephant shrew, *Elephantulus edwardii* (Smit et al. 2007); Karoo bush rat, *Myotomys unisulcatus* (Edwards *et al.* 2011); Southern Rock agama, *Agama atra* (Swart et al. 2009); Clicking stream frog, *Strongylopus grayii* (Tolley et al. 2010a) and Burchell's sand lizard, *Pedioplanis burchelli* (Tolley *et al.* 2009)]. In the open-habitat Puff adder, *Bitis arietans*, refugial areas have been identified and they too correspond to these common phylogenetic breaks in some regions. The similar pattern would suggest common positions for refugial areas and thus climatically stable areas, though few have actually been identified (Barlow *et al.* 2013).

Pedioplanis lineoocellata is a small southern African endemic lacertid with a range in open habitats extending from south-west Namibia through Botswana and into much of South Africa (Figure 3.1), a rather similar distribution to the puff adder (Branch 1998; Makokha *et al.* 2007; Conradie *et al.* 2012; Edwards 2013; Bates *et al.* 2014). *Pedioplanis lineoocellata* forms a strongly supported monophyletic clade (Makokha *et al.* 2007) including both the previously recognised subspecies *P. l. lineoocellata* and *P. l. pulchella* (Branch 1998), however recent genetic work suggests a more complicated relationship and the sub-species have been invalidated due to lack of morphological or genetic support (Edwards 2013). The species complex occupies a diverse set of habitats including montane grassland, coastal fynbos, succulent and Nama Karoo, and arid and moist savanna (Branch 1998; Edwards 2013). Because of the species' occurrence in several types of open-habitat vegetation and its wide distribution, it is an excellent study system for examining and comparing geographic patterns in genetics.

At least four previously unknown evolutionary distinct lineages or clades have been identified for *P. lineoocellata* based on mitochondrial DNA (mtDNA) in South Africa (Figure 3.2). This includes the identification of the isolated Waterberg (Limpopo Province) and south coast clades as being genetically distinct (Edwards 2013). The most geographically extensive of the mtDNA clades (hereafter Clade A) occurs across the arid savanna and Karoo environments and along the western parts of the distribution within South Africa but could possibly occur in the moister savanna environments in Botswana as well. The second most extensive clade identified was the Karoo interior clade (hereafter Clade B), which occurs in the south-western region of South Africa. Many southern African taxa have similar genetic breaks between identified lineages as found for *P. lineoocellata*, particularly in the south-western coast of South Africa and around the Groot Swartberg area. Edwards (2013) proposed a possible scenario of range contraction of a wide-spread ancestor into isolated refugia due to climatic shifts and the lineages would likely have expanded their range during favourable warmer and wetter periods of climatic oscillations. This has been found for other southern

African species, in some cases with possible secondary contact on several occasions (e.g. Kryger et al. 2004; Lorenzen et al. 2012; Barlow et al. 2013). Clade A has some distribution overlap with Clade B around the western interior at the Loeriesfontein area (Figure 3.1) based on evidence of sympatric mitochondrial lineages within a single sampling site (Farm Hartebeeslaagte). Their co-occurrence begs the question as to whether hybridization may be occurring in this area as well as the possibility of other contact zones.

Understanding gene flow dynamics within and across clades can assist in describing how possible scenarios of isolation and expansion may have affected *Pedioplanis lineoocellata*. Additionally, understanding gene flow and geographic genetic structure would be useful for investigating the possible genetic admixture at the contact zone. In this study I investigate the genetic relationship between the previously identified mitochondrial clades, Clade A and B, of *P. lineoocellata* using nuclear DNA microsatellite markers and examine the relationship between populations within the most wide-spread mtDNA clade, Clade A. Additionally, I examine the possibility of hybridization at the Loeriesfontein area.

Although mtDNA clades have been identified, an investigation of microsatellite DNA patterns is yet to be undertaken. Previous work on genetic sequences from three lizards, *Agama atra*, *Bradypodion pumilum* and *Pedioplanis burchelli*, from the Cape Floristic Region of South Africa found limited evidence for contemporary gene flow between sub-specific clades (Swart *et al.* 2009; Tolley *et al.* 2009). Should across-clade gene flow be similarly low, hyper variable microsatellites should have greater between population variation than mtDNA. Additionally, as nuclear regions, microsatellites undergo crossing-over and independent assortment during sexual reproduction, further increasing their variability. The additional resolution gained in using microsatellite markers also allows one to track contemporary genetic changes in populations and over fine temporal and spatial scales.
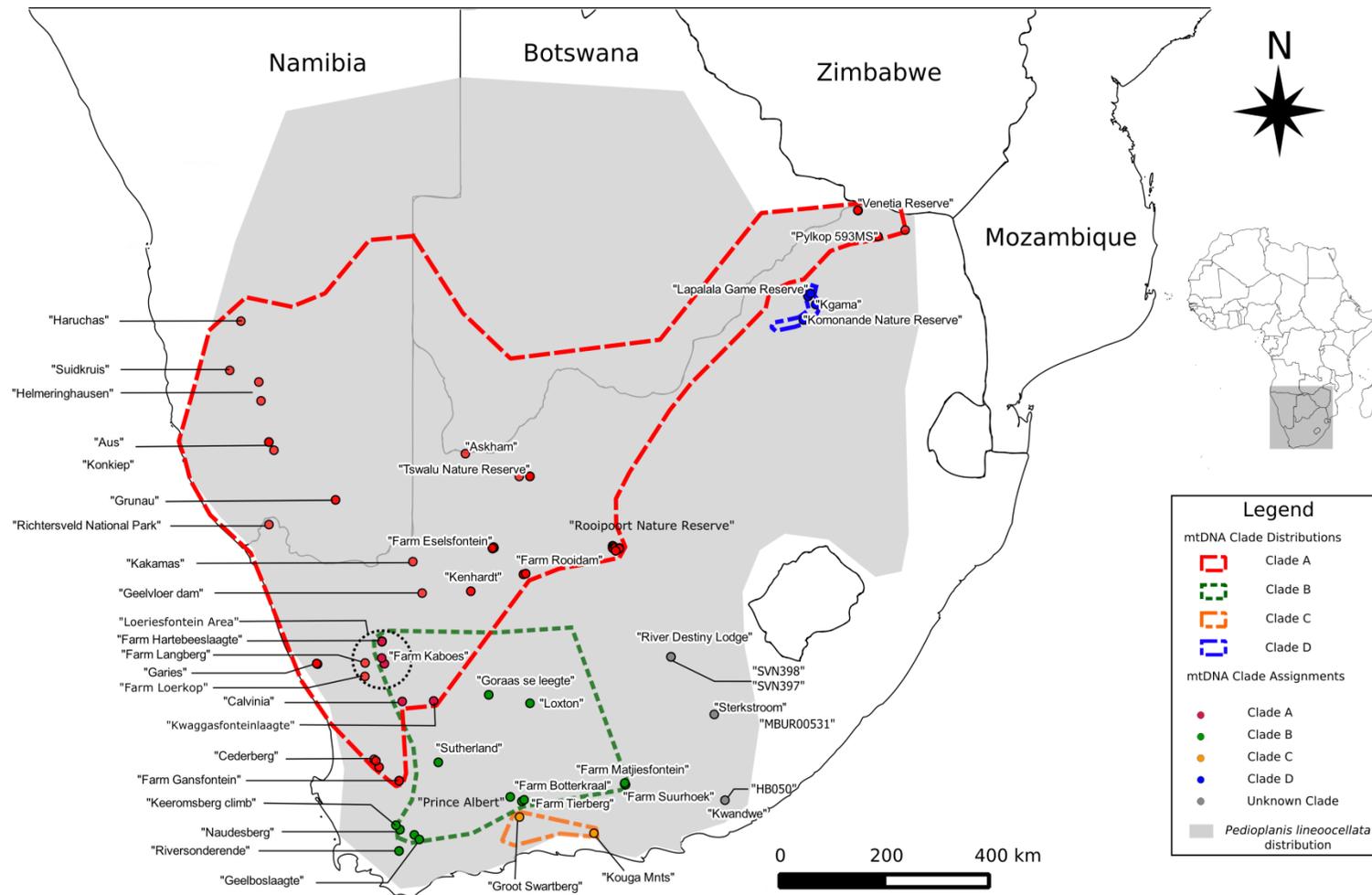
Figure 3.1: Map for *Pedioplanis lineoocellata* sample sites within southern Africa. All samples used in this study (coloured dots) overlain by mtDNA clades A-D (dotted polygons) as genotyped by Edwards (2013). Species distribution (grey shading) estimated using occurrence records from GBIF, HerpBank (SANBI) and SARCA (Bates *et al.* 2014). Locality names for sample sites labelled on the map. The Loeriesfontein area is indicated by the broken circle. Labels for samples with unknown mtDNA clades include sample name. Inset of Africa in top right position. Legend of clade colours and borders in bottom right position.
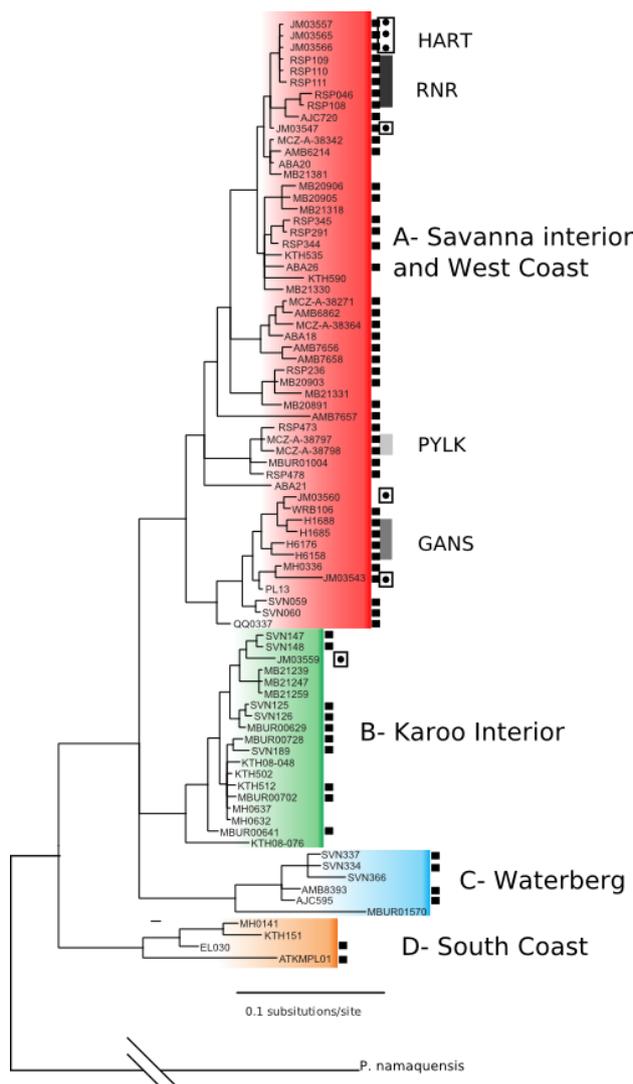
Figure 3.2: Phylogeny of *Pedioplanis lineoocellata* based on ND4 and 16S mitochondrial gene regions, adapted from (Edwards 2013). Clades A-D indicated above samples. • indicates samples from population HART and ▪ indicates samples in common with the present study. Bars above the sample name indicate to which of the primary sites sampled for this study the individual belongs. Clade C and D samples discussed in Appendix A .

Pedioplanis lineoocellata of Clade A occupies homogenous habitat types and may be abundant where it occurs (see Chapter 4). *Pedioplanis lineoocellata* has also demonstrated behavioural flexibility (Wasiolka et al. 2009b, 2009a; Blumroeder et al. 2012) that may allow it to exist in structurally diverse habitats and with different food availabilities. Although the habitat occupied by *P. lineoocellata* is seemingly continuous and behavioural flexibility could facilitate movement across unfavourable areas, dispersal may still limit gene flow over great distances because dispersal distances for many lacertids and other similar sized lizards are seldom greater than 10-100m (Olsson et al. 1996, 2003; Ronce et al. 1998; Chapple et al. 2005; Clobert 2012). I predicted that within Clade A, which spans hundreds of kilometres, there would be a strong

prevailing Isolation-by-distance pattern because of the geographic scale of the clade's distribution and the likely limited dispersal over great distances. I also predict the presence of strong population structure and limited dispersal between populations as reflected by low migration rates estimates.

The distinction between clades might be clearer when investigated using microsatellite markers provided there has been sufficient divergence time and limited gene flow. Given this, I predict that microsatellite data will produce the same clade groupings as did mtDNA except at the site of possible contact between clades. The result should corroborate the genetic distinctiveness of the different mtDNA clades. The co-occurrence of two different mtDNA lineages at the Loeriesfontein sampling site, Farm Hartebeeslaagte, suggests a possible contact zone between Clade A and Clade B with the possibility of hybridization as well. Alternatively the lineages may be sympatric with little to no genetic exchange, which could potentially suggest divergence in sympatry and the potential for ecological speciation. Hybridization between previously isolated lineages has been identified in some Eurasian lacertids taxa (Zhao *et al.* 2011; Miraldo *et al.* 2013) but there is no work on southern African lacertids in this regard. The genetic distinction between Clade A and Clade B should be less obvious in the contact zone if there is hybridization. I predicted that individuals from areas within the contact zone will show genetic admixture between Clade A and Clade B indicating hybridization of nuclear DNA.

2   METHODS AND MATERIALS

2.1 SAMPLING

A total of 185 samples from over 45 localities across South Africa and parts of Namibia were sampled during previous field work (2009-2011) and during surveys at Rooipoort Nature Reserve (RNR), near Kimberley (2012-2013). The analyses were performed using only samples from Clade A and B because low sample sizes for Clade C and D precluded their use. The dataset analysed included eight primary sample sites with more than five samples each from Chapter 2 and the remaining 37 sample sites (71 individuals) with fewer than five samples each (see Chapter 2 for further details). Seven of the primary sites; ESLF, GANS, PYLK, RNR1, RNR2, RNR3, WATK fall inside the mtDNA Clade A distribution while the eighth site, HART, falls within the Clade A-B overlap region (Figure 3.1). The four sample sites from within RNR, RNR1-3 and WATK, were collected for this study during April-May and again in October-November 2012.

Lizards were caught using a pan-fishing pole with a noose attached, which is an effective and commonly used method for catching lizards (e.g. Bertram & Cogger 1971; Verwaijen & Herrel 2002; Doughty & Sinervo 2013). Lizards were discovered by performing random walks across suitable sites between 8am - 6pm. Once a lizard was seen, it was caught and measured. At each capture point GPS coordinates (GPSmap 60CS, Garmin Series with ±3m accuracy) were recorded. The animals' body size (Snout-vent length; SVL)

and tail length (TL) were measured using a digital calliper (e.g. Bauwens & Diaz-Uriarte 1997; Damme & Vanhooydonck 2002). Sexes were recorded as well (Appendix Table A 1, described in Chapter 4). A small (~1cm) segment of tail tip was removed from caught lizards and stored in 98% ethanol. This is common practice as a non-invasive biopsy for lacertid species (e.g. Richard & Thorpe 2000; Laloi et al. 2004; Miraldo et al. 2013) as they naturally autotomize and regenerate their tails as a defence mechanism (Cooper et al. 2004). Tail clippings also served as a batch mark to ensure that the same individual was not sampled twice. All samples were profiled at nine microsatellite loci. Details of the laboratory procedure, the analyses and results for allelic diversity, Hardy-Weinberg equilibrium and linkage disequilibrium for the eight primary sample sites can be found in Chapter 2.

2.2 ANALYSIS

2.2.1   POPULATION STRUCTURE WITHIN CLADE A

Population structure within Clade A was examined using a subset of the total available samples that consisted of only samples within the known mtDNA Clade A distribution and included the eight primary sample sites.

2.2.1.1 Analysis of molecular variance

Population level variation for the eight primary sample sites was assessed using a hierarchical analysis of molecular variance (AMOVA), which divides the variance into additive components. Because all sites from outside of Rooipoort Nature Reserve (RNR) were over 200km from each other, only the four RNR sites were grouped together as a region whereas the remaining four sampling sites (ESLF, GANS, HART & PYLK) were each assigned to separate regions. Pairwise genetic differentiation between the eight sites was examined using F-statistics estimated using F_{ST} (Cockerham et al. 1993) and R_{ST} (Slatkin 1995) estimators. F_{ST} uses allele identity and frequency to estimate genetic differentiation (Cockerham et al. 1993; Raybould et al. 2002) while R_{ST} includes information on the allele size difference following an expected simple single-step mutation model and thus can be more informative provided the time-scale of interest exceeds 10-100 generations (Slatkin 1995). The AMOVA was run with 10 000 permutations to estimate significance p-values in Arlequin v3.5 (Excoffier et al. 2010).

2.2.1.2 Isolation-by-distance

An isolation-by-distance (IBD) pattern between populations was examined using a Mantel test (Mantel 1967), which tests for a correlation of pairwise genetic differentiation with geographical distances between populations. Analyses were performed in the online program IBDWS v2.23 (Jensen et al. 2005). Correlations

of both F_{ST} and R_{ST} estimates against linear and log-transformed distances (km) were examined and analyses were performed with 1000 randomizations. Because of possible hybridization at site HART, which may produce peculiar results, I ran analyses with and without HART. The DAPC analysis suggested that GANS may be particularly divergent from the other Clade A sites (see results), thus IBD analyses were also repeated excluding this site.

2.2.1.3 Population assignment

Genetic clustering of individuals and the sample site membership assignment probabilities based on microsatellite data were examined using a Discriminant Analysis of Principle Components (DAPC) (Jombart et al. 2010). Such multivariate genetic statistics rely on using synthetic variables built as linear combinations of genetic variables (for microsatellites these would be allele frequencies), to form combinations that reflect as best as possible the genetic variation among the studied individuals (Jombart 2013). This analysis was selected over others such as STRUCTURE (Pritchard *et al.* 2000) as it does not make any assumptions about mutation models, Hardy-Weinberg equilibrium or mutation-drift equilibrium and therefore potentially avoids bias induced by deviations from HWE (Chapter 2). Unlike an ordinary Principal Component Analysis (PCA) and similar approaches, DAPC optimizes between-group variation whilst minimising within-group (or between individual) variation by employing discriminant function analysis of the principle components (Jombart et al. 2010). The analysis was performed using the adegenet package (Jombart 2008) under R environment v.3.0.2 (R Core Team 2013) following the guidelines provided in the DAPC vignette (Jombart 2013).

Including prior sample site assignments in the DAPC analysis may optimise principal components (PC) and discriminant functions (DA), and improve the accuracy of the results. However, if prior assignments do not correspond to the actual microsatellite patterns of relatedness, forcing prior groupings would be misleading. To avoid this potential problem, samples were first examined without the restriction of *a priori* groups to assess the appropriateness of using prior group assignments. I performed a k-means analysis using the find.cluster() function in adegenet to identify the number of groups that best summarise the genetic variation. The maximum number of possible clusters was set to 10 as it seemed unlikely for there to be more clusters than well sampled sites. The number of PCs retained (n.pca) was set to 200 as an arbitrarily large number. Aside from computational time, there is no reason to exclude principle components at the cost of retaining discriminatory power in a k-means analysis (Jombart 2013). The optimal clustering is judged using a Bayesian Information Criteria (BIC), which is similar to other commonly used criteria such as the Akaike Information Criteria (AIC) in that the lowest BIC indicates the best fit (Jombart 2013).

The second round of analyses included the use of prior group assignments. Samples were split into training data and supplementary data. The training data were samples collected within a well-defined

sampling areas with a fair sample size (n>5) and were used to define the PCs and DAs. Thus training data included seven of the eight sample sites; ESLF, GANS, RNR1, RNR2, RNR 3, WATK, PYLK and two additional sample sites from within RNR, RNR4 and HRBK. Supplementary data were defined as all other samples from within the distribution of Clade A but without well-defined sampling areas. The DAPC analysis assigned supplementary individuals to one of the prior groups from the training data based on the optimised PCs and DAs. Samples from HART were included as supplementary data to identify the Clade A populations which may have contributed to the genetic composition of HART. Clade B samples were left out of this analysis because the variation would be uninformative without the context of a well sampled Clade B site and may result in loss of discriminatory power in the discriminant functions.

The DAPC was performed using the function dapc(). The na.replace() function was used to replace missing data with uninformative global mean values, (method="mean"). Unlike the k-means analysis, DAPC benefits from only retaining necessary PCs. The appropriate number to retain was determined by the optimum a-score, which is a measure of the difference between the proportion of successful re-assignments to prior groups and to those obtained from a randomised dataset (Observed discrimination- random discrimination) corrected for the number of PCs retained (Jombart 2013). The function optim.a.score() was used with 40 simulations (n.sim=40) and the median optimum n.pca was set. I retained all axes in the discriminant analysis step. The colorplot() function was used to plot data points in geographic space with the first three DAs presented as colour intensity in RGB channels indicating genetic similarity. Membership probabilities analogous to the admixture coefficients of STRUCTURE (Pritchard *et al.* 2000), were obtained and interpreted as indicators of genetic similarity of individuals to a particular clusters or as evidence of admixture (Jombart 2013). This was used to produce a plot of membership probabilities to evaluate the assignments of each individual to each of the prior groups.

2.2.1.4 Migration rate estimates

Estimates of proportion migrants (m) at each site was used to indicate migration between sites by using the Bayesian inference method implemented in BayesAss v.3.0 (Wilson et al. 2003). This method was preferred to those of IMA2 (Hey 2010) and Migrate (Beerli et al. 2010) as it allows the use of individuals with missing data and relaxes the assumptions around HWE (Wilson et al. 2003). Parameter settings recommended by Faubet et al. (2007) were used, including five independent runs with random number seeds, 50×10^6 iterations with 10% discarded for burn-in and sampling performed every 100 000 iterations. The mixing parameter for the inbreeding coefficient (-f), migration rates (-m) and the allele frequencies (-a) were tested between 0.01-1.00, however, even with all parameters set at 1.00 (the largest possible move between proposed values) the acceptance rates remained rather high (~50-60%). All values were thus set at 1.00 to maintain acceptance rates between 20-60% as recommended by Rannala (2013). High acceptance rates may

be suggestive of weakly informative data and a very flat likelihood surface (Rannala 2013). Trace files from all runs were examined for evidence of convergence using R.

2.2.2   CLADE DELIMITATION

The correspondence of microsatellite groups to the previously identified mtDNA clades (Edwards 2013) was assessed by performing the DAPC analysis using the k-means analysis and by using clades as prior groups as described above. Training data consisted of samples that had been genotyped by Edwards (2013) (i.e. the exact same samples), thus had confirmed *a priori* mtDNA clade assignments (n=49). Samples from HART were excluded from the training data because of possible hybridization (Table 3.1). Supplementary data consisted of all samples not genotyped for mtDNA (n=136) and were assigned to clades based on their distribution within the known range of the two mtDNA clades (see Figure 3.1) thus giving them an assumed *a priori* mtDNA clade. This was necessary for evaluating successful 're-assignment' probability of supplementary data to the correct clades. The k-means analysis was performed twice to assess microsatellite clusters without PC and DA optimisation. Firstly, the analysis was performed using only the training data and secondly, using all the available data. This was done to evaluate if training data, for whatever reason, better discriminated between clades without supplementary samples. The DAPC analysis was then repeated with prior clade assignments. In total the analysis was run three times.

2.2.1   CLADE A –CLADE B HYBRIDIZATION

The possibility of hybridization between Clade A and Clade B at the overlap region was examined using the population HART as supplementary data in the DAPC analyses described in the sections above for population assignment within Clade A, Section 2.2.1, and Clade delimitation, Section 2.2.2. Specifically, the genetic similarity of the samples from the clade overlap region, including HART samples, to sample sites with known *a priori* groupings was examined. The population-level DAPC analysis within Clade A was used to identify which of the Clade A sample sites were the most similar to HART. The analysis using Clade A and B as prior groups assessed the level of between-clade admixture at HART. Mixed group assignments at <0.85 assignment probability to any population were interpreted as an indication of genetic admixture with the respective populations. Migration rate estimates provide an additional measure of contribution as one would expect evidence of greater migration from the contributing sample sites than the other sites.

Table 3.1: Sample sizes for training and supplementary samples of *Pedioplanis lineoocellata* with Clade A or B assignments used for the DAPC analysis. Supplementary data were assigned to clades based on geographic position.

mtDNA Groups	Training	Supplementary
Clade A	40	108
Clade B	9	9
Unknown	0	2
Clade A-B Overlap	-	17
Total	49	136
Clade overlap region (*a priori* assignment)		
HART (Unknown)	0	15
Farm Loerkop (Unknown)	0	1
Farm Kaboes (Unknown)	0	1
Calvinia (Clade A)	1	0
Kwaggasfonteinlaagte (Clade A)	1	0

3 RESULTS

3.1 POPULATION STRUCTURE WITHIN CLADE A

3.1.1 ANALYSIS OF MOLECULAR VARIANCE

Population structure within the widely distributed Clade A was evident from hierarchical AMOVA results at multiple levels (sites and regional groups) using both F and R statistics, and the distribution of variation across levels of organisation was similar between statistics (Table 3.2). Very little of the variation was partitioned between populations within regions (i.e. sites within RNR). There is support for grouping all RNR sites into a single 'sample site' dataset as compared to the remaining sample sites. The R statistic values were larger than F statistics for variation between groups indicating an influence of mutation model.

Table 3.2: Hierarchical analysis of molecular variance (AMOVA) results for *Pedioplanis lineoocellata* based on the primary eight sample sites. Percentage variation by source based on both F and R statistics. Information provided: SS- sum of squares; VC- variance components; % var- percentage variation.

Source	F			R		
	SS	VC	% var	SS	VC	% var
F_{CT}/R_{CT}	85.0	0.5	13.6	4305.5	23.8	21.6
F_{SC}/R_{SC}	14.4	0.0	1.0	1044.6	8.7	7.9
F_{IS}/R_{IS}	390.0	0.5	13.6	10124.4	16.8	15.3
F_{IT}/R_{IT}	304.5	2.7	71.8	7045.0	61.0	55.3
Total	793.9	3.7	100	22519.5	110.3	100

The patterns in the pairwise estimates of F_{ST} and R_{ST} were largely similar to each other. Pairwise comparisons between sample sites within RNR showed small genetic differentiation (F_{ST} = 0.00 to 0.046) but results were not significant. In some cases, small sample sizes may have led to the low statistical power and this might explain the lack of statistical significance for some of the large F-statistics. F-statistics for the two better sampled sites (RNR1 and RNR2) were of similar magnitude to the other RNR sites but were significant (Table 3.3). Considering the small geographic distances between these samples, these F-statistics suggest restricted gene flow between sample sites leading to genetic structure even on this small scale. All the remaining between-region comparisons showed significant genetic differentiation (F_{ST} = 0.105-0.289; p<0.01) offering strong support that individual sample sites are differentiated. HART, on average, had the highest levels of differentiation when compared to other sites (F_{ST} 0.149-0.289, R_{ST} 0.089-0.333) and thus seemed particularly divergent.

3.1.2   ISOLATION-BY-DISTANCE

The Mantel test results were not significant regardless of whether F_{ST} or R_{ST} was used (r =-0.28 & 0.24 ; p>0.3) (Figure 3.3) and there was no change in the pattern when the test was re-run without HART (F_{ST} r = 0.181, Z = 4.823, p = 0.417; R_{ST} r = 0.218, Z = 6.999, p = 0.287) or without GANS (F_{ST}-r = -0.126, Z = 6.404, p = 0.681; $R_{ST} - r$ = 0.183, Z = 10.082, p = 0.450). This indicates that isolation-by-distance is not the most significant process influencing population structure across Clade A.
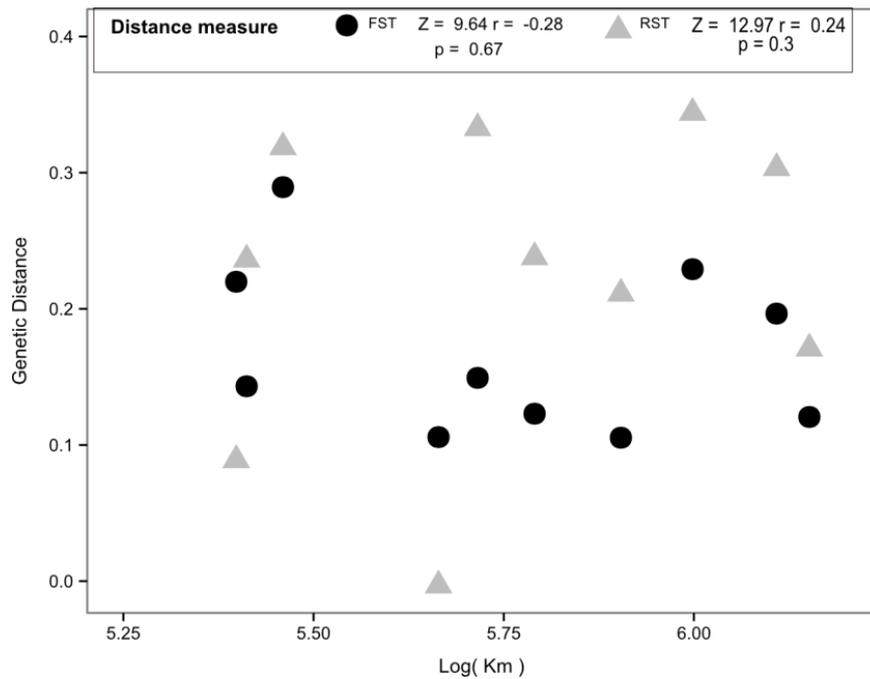
Figure 3.3: Isolation-by-distance plot for *Pedioplanis lineoocellata* as indicated by the genetic distance (F_{ST} or R_{ST}) with the logarithm of geographical distance (km) for all pairwise comparisons of all sites within Clade A.

3.1.3   POPULATION ASSIGNMENT WITHIN CLADE A

For the DAPC analysis assessing only Clade A individuals without prior sample site assignment, 55 principle components (PCs) were retained, which accounted for 92.5% of the genetic variation. The k-means analysis indicated 3-5 clusters fit the data equally well (Table 3.4). The hierarchical splitting of clusters with increasing k values did not have any clear geographic pattern (Figure 3.4). At k=3, sites HART and PYLK were grouped into cluster 1 but the cluster included RNR samples as well. Sites GANS and ESLF and a variety of other samples were grouped into cluster 2 while Cluster 3 was predominantly RNR samples (Figure 3.5). With k=4, the RNR and PYLK individuals from cluster 1 were separated from HART samples, the latter now forming cluster 4. With k=5, GANS was separated out from the rest of cluster 3 and very few individuals showed evidence of mixed assignments. By the sequence in which clusters split, genetic similarity between sites was not based on geographic proximity.

The DAPC analysis was then performed with geographically defined sample sites used as prior groups (Appendix Information A 2). Membership probabilities of training data support sample site integrity with >90% correct re-assignment of samples to their *a priori* sample site but the k-means analysis results suggest that RNR may consist of individuals from several genetic clusters. Should RNR consist of individuals from several microsatellite clusters, the use of RNR as training data in the DAPC with prior group assignments

would have led to the formation of a genetically diverse 'umbrella' group within which several other potential groups were pooled. Therefore, the use of prior group assignments was considered inappropriate.

Table 3.3: Sample site pairwise comparisons of F_{ST} and R_{ST} values for *Pedioplanis lineoocellata*. R_{ST} values below the diagonal and F_{ST} values above the diagonal. Values on the diagonal are inbreeding coefficients estimated ($R_{IS} \backslash\backslash F_{IS}$). Significant values in **bold**; p<0.05*, p<0.01**, p<0.001***. *n* indicates sample sizes per sampling site.

Sites (n)	RNR1 (35)	RNR2 (29)	RNR3 (5)	WATK (5)	ESLF (5)	GANS (16)	HART (19)	PYLK (6)
RNR1	0.14\\0.09	**0.013***	0.016	-0.004	**0.158***	**0.125***	**0.158***	**0.117***
RNR2	**0.129***	0.3\\0.2	0.017	-0.001	**0.134***	**0.124***	**0.161***	**0.108***
RNR3	0.047	**0.125***	0.48\\0.07	0.046	**0.132**	**0.116***	**0.208***	**0.170**
WATK	0.059	-0.014	0.083	0.41\\0.3	**0.164**	**0.127***	**0.133***	**0.099***
ESLF	**0.250***	**0.294***	**0.191***	**0.242***	0.31\\0.25	**0.106***	**0.289***	**0.229**
GANS	**0.261***	**0.239***	**0.266***	**0.158**	-0.00318	0.16\\0.11	**0.220***	**0.121***
HART	**0.351***	**0.391***	**0.451***	**0.332***	**0.319***	**0.089**	0.29\\0.28	**0.196***
PYLK	**0.170**	**0.351***	**0.339**	**0.263**	**0.344**	**0.171***	**0.303***	-0.08\\-0.03
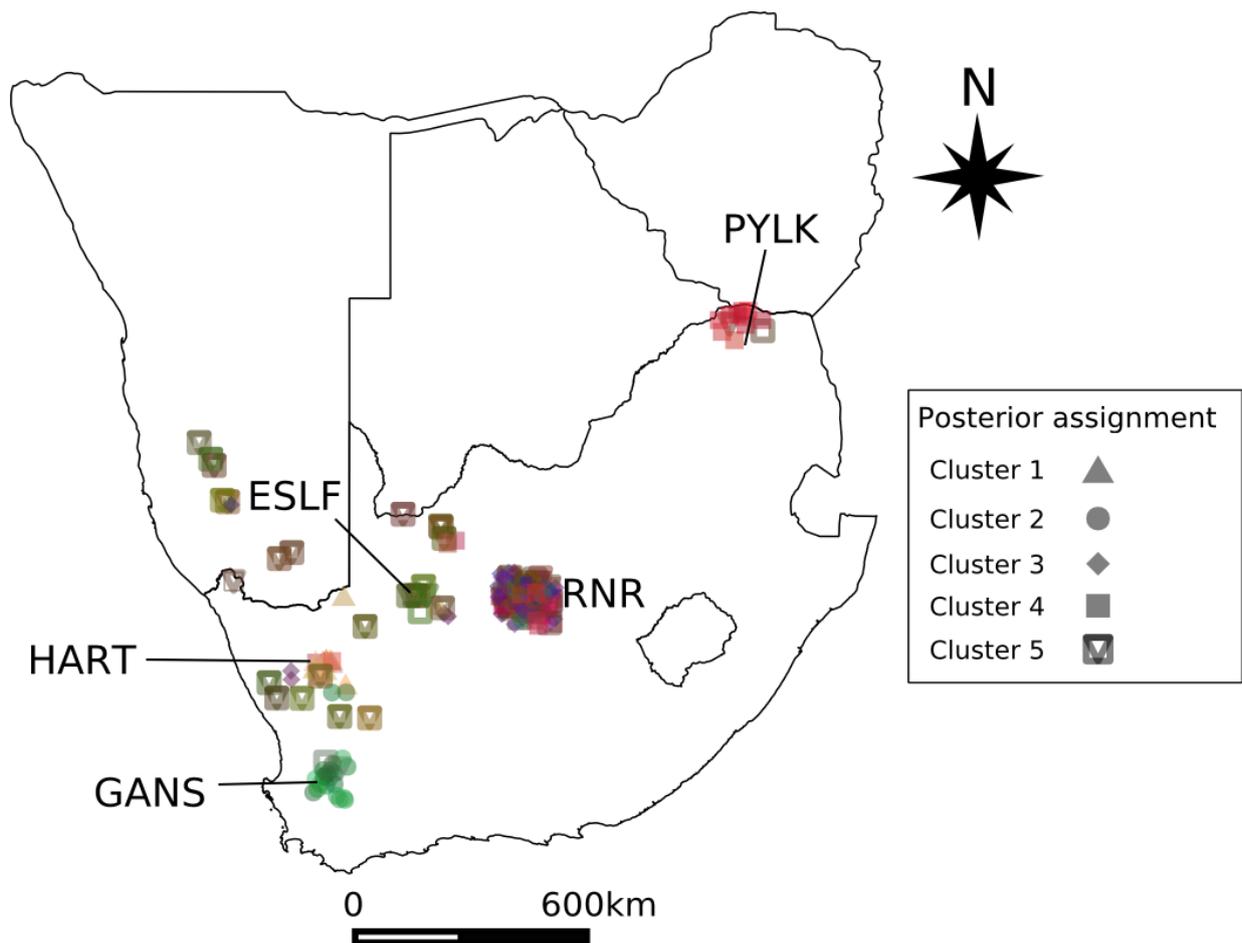
Figure 3.4: Microsatellite DNA clusters (k=5) based on DAPC for *Pedioplanis lineoocellata* from within the distribution of mtDNA clade A. Cluster numbers match that of the membership assignment plots for k=5. Colours of individual symbols represent genetic similarity and are determined by weighting the first three colour channel intensities in RGB according to the co-ordinates from the first three discriminant functions from the DAPC analysis.

Table 3.4: BIC values from k-means analysis for various cluster numbers (k) near the turning point of the BIC curve. Data used included all data (no prior groups) within Clade A and, training data (prior groups) and all data (no prior groups) across Clade A and B, respectively. Values selected for use indicated in **bold**.

Within Clade A		Across clades	
	All		Training	All
k=2	221.469	**k=1**	**82.972**	287.028
k=3	**217.236**	k=2	83.210	281.539
k=4	**216.298**	k=3	84.065	**278.453**
k=5	**215.783**	k=4	85.262	**278.329**
k=6	216.431	k=5	87.305	278.885

3.2 MIGRATION RATE ESTIMATES

The estimated migrant proportion of each sample site was negligible in most cases (m<0.01; Table 3.5). Low migrant proportions correspond well with large and significant F-statistics and the few individuals with mixed membership assignments in the DAPC analysis. There was no evidence of migration between HART and any other site. There was evidence for non-symmetrical dispersal from GANS to ESLF and among all the RNR sites. Point estimates for the migrant proportions for these sites were high but CI bounds were very large, effectively stretching from a low non-zero value to the maximum m value (0.3) allowed by BayesAss. Furthermore, point estimates may be tentative when F_{ST} values are low (<0.01), such as those among RNR sampling sites, but the true migration rates are still expected to lie within the 95% CI under most circumstances (Faubet *et al.* 2007). Therefore, the most conservative conclusion would be that there is a non-zero migrant proportion for the respective sites. The low genetic differentiation would support gene flow in any case.

3.3 CLADE DELIMITATION

Training data examined by DAPC without prior group assignments (i.e. only samples genotyped for mtDNA) produced the lowest BIC at k=1 clusters (Table 3.4) indicating that microsatellite (nuclear DNA) patterns do not reflect the same patterns found for the mtDNA clades. When using all of the available data, the k-means analysis had the lowest BIC values at k = 3-4 (Table 3.4). Both k-values produced clusters that had notable geographic overlap (Figure 3.6), but no useful resolution was gained from using k=4 (Figure 3.7). With k=3, 60 PCs and two discriminant functions (DAs) were saved, which retained 81.8% of the variation. As a consequence of large geographic overlap, genetic groups did not correspond to the mtDNA clades of Edwards (2013). At k=3, the three clusters formed (clusters 1-3) were composed of a subset of *a*

priori Clade A samples, *a priori* mtDNA Clade B samples and the samples from the Clade A-B overlap region, respectively. However all clusters also included many other samples from the Clade A distribution (Figure 3.7).

Geographically, samples from cluster 2 were the most widespread while samples from cluster 1 were largely restricted to RNR (Clade A) (Figure 3.6). Samples from cluster 3 were primarily from two sites, PYLK (Clade A) and HART (Clade A-B overlap). Evidently, mtDNA clades do not reflect the same geographic pattern as microsatellites and there is microsatellite genetic structure that was not detected using mtDNA.

Very few samples showed mixed assignment probabilities and most were strongly assigned to one cluster. Only seven samples (six from RNR) had <0.8 assignment probability to any cluster (k=3), and only four samples when k=4. Surprisingly, with k=4, samples from the A-B overlap region still strongly clustered together with no evidence of mixed assignment. High assignment probabilities to a cluster, as opposed to mixed assignment probabilities, does not support genetic admixture between clades at HART.
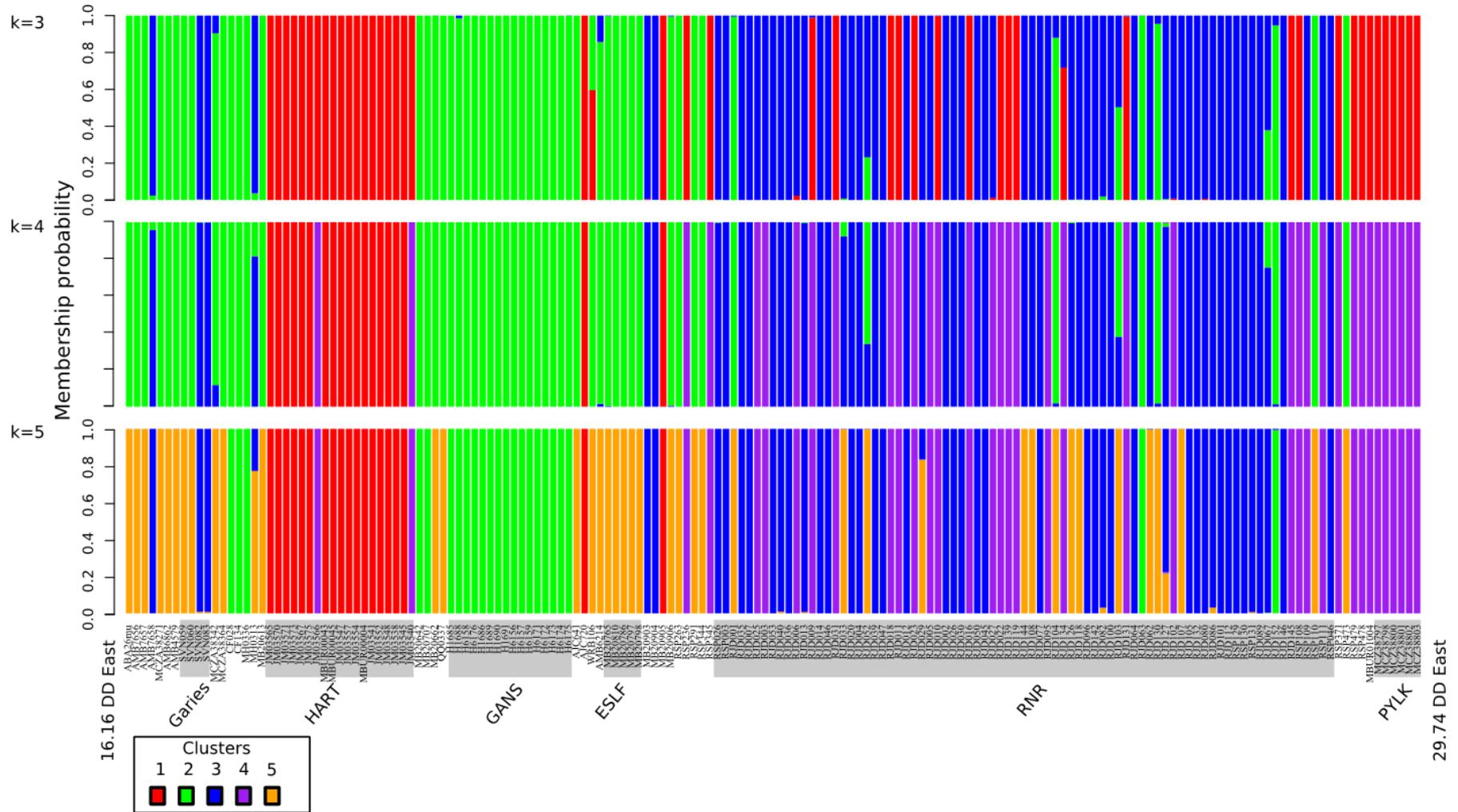
Figure 3.5: Assignment proportions for individual *Pedioplanis lineoocellata* samples within mtDNA Clade A to one of the DAPC clusters (k=3-5). Names of sampling areas indicated below the bars and samples arranged according to longitude, from west to east. Colours of membership assignment proportions match cluster colours in the legend.

Table 3.5: Estimations of migrant proportions (m) as a result of recent migration between sample sites of *Pedioplanis lineoocellata* across mtDNA Clade A. The 95% confidence intervals indicated in brackets (Lower CI - Upper CI). Sites in columns are donor sites and sites in the rows are the receiving sites, i.e. migration occurs from column to row. Values with lower CI >0.00 are in ***bold italics,*** and estimates with the upper confidence limit >0.1 are in **bold.** Values in the diagonal are the proportion of individuals in each population with ancestry from the current population.

Receiving\donor	ESLF	GANS	HART	PYLK	RNR1	RNR2	RNR3	WATK
ESLF	0.694	*0.152*	0.026	0.026	0.026	0.026	0.026	0.026
	(0.645 - 0.743)	*(0.066 - 0.237)*	(-0.021 - 0.072)	(-0.021 - 0.072)	(-0.021 - 0.072)	(-0.021 - 0.073)	(-0.021 - 0.072)	(-0.021 - 0.072)
GANS	0.014	0.901	0.014	0.014	0.017	0.014	0.014	0.014
	(-0.012 - 0.039)	(0.84 - 0.961)	(-0.012 - 0.04)	(-0.012 - 0.04)	(-0.014 - 0.048)	(-0.012 - 0.039)	(-0.012 - 0.04)	(-0.012 - 0.039)
HART	0.012	0.013	0.91	0.012	0.016	0.012	0.012	0.012
	(-0.011 - 0.035)	(-0.011 - 0.037)	(0.855 - 0.966)	(-0.011 - 0.035)	(-0.013 - 0.045)	(-0.011 - 0.036)	(-0.011 - 0.035)	(-0.011 - 0.035)
PYLK	0.024	**0.048**	**0.061**	0.694	**0.089**	**0.037**	0.024	0.024
	(-0.02 - 0.068)	**(-0.047 - 0.143)**	**(-0.05 - 0.172)**	(0.641 - 0.746)	**(-0.032 - 0.21)**	**(-0.055 - 0.128)**	(-0.02 - 0.067)	(-0.02 - 0.068)
RNR1	0.008	0.014	0.009	0.008	0.938	0.009	0.008	0.008
	(-0.007 - 0.023)	(-0.008 - 0.036)	(-0.008 - 0.026)	(-0.007 - 0.023)	(0.898 - 0.977)	(-0.009 - 0.027)	(-0.007 - 0.023)	(-0.007 - 0.022)
RNR2	0.009	0.015	0.009	0.009	*0.264*	0.677	0.009	0.009
	(-0.008 - 0.026)	(-0.009 - 0.038)	(-0.008 - 0.026)	(-0.008 - 0.026)	*(0.218 - 0.309)*	(0.657 - 0.696)	(-0.008 - 0.026)	(-0.008 - 0.026)
RNR3	0.026	0.026	0.026	0.026	*0.152*	0.026	0.694	0.026
	(-0.021 - 0.072)	(-0.021 - 0.074)	(-0.021 - 0.072)	(-0.021 - 0.072)	*(0.065 - 0.238)*	(-0.021 - 0.072)	(0.645 - 0.742)	(-0.021 - 0.072)
WATK	0.026	**0.043**	0.027	0.026	*0.133*	0.026	0.026	0.694
	(-0.021 - 0.072)	**(-0.019 - 0.106)**	(-0.022 - 0.075)	(-0.021 - 0.072)	*(0.046 - 0.22)*	(-0.021 - 0.073)	(-0.021 - 0.072)	(0.645 - 0.742)

Figure 3.6: Microsatellite DNA clusters (k=3) based on DAPC analysis for *Pedioplanis lineoocellata* across Clade A and B. Cluster numbering matches that of the membership assignment plots for k=3. Colours of individual symbols represent genetic similarity and are determined by weighting the first two colour channel intensities in RGB according to the co-ordinates from the first two discriminant functions from the DAPC analysis.

To assess if forcing prior mtDNA clade assignments on samples for PC and DA optimisation improved discrimination between clades, the DAPC analysis was re-run on training data using the prior clade assignments. The optimum number of PCs was set at 14 and one DA was saved as there were only two clades included. The discriminant functions retained 58.2% of the variance, far less than that retained for all the data under the k-means analysis. The two clusters defined using the training data separated well and the proportion of training data samples correctly re-assigned to their prior groups was high (>75%) for both mtDNA clades (Figure 3.8). Several individuals with prior Clade A assignment but not from the overlap region had a high Clade B membership probabilities indicating an intermediate genetic position between the samples used to define the clades and suggestive of gene flow. These individuals were largely from the western south-west region of South Africa and proximate to the Clade B distribution (GANS, Cederberg and Calvinia). The discriminant function performed poorly on the supplementary data for samples with Clade B

Page | 48

prior assignment. Only 11% of individuals were correctly "re-assigned" to the posterior Clade B cluster. Most samples from the Clade A-B overlap region were assigned to the posterior Clade A cluster with >0.9 assignment proportion, which does not support hybridization at these sites. This was surprising because, samples from well within the Clade A distribution, such as RNR, had higher posterior Clade B assignment proportions than samples with prior Clade A-B assignment, as much as 0.2 in some cases. Geographically, the distinction between the mtDNA clades based on microsatellite markers is weak though a trend is evident as indicated by genetic similarities where samples from the south-western parts of southern Africa are marginally different from the rest of the distribution (Figure 3.9).
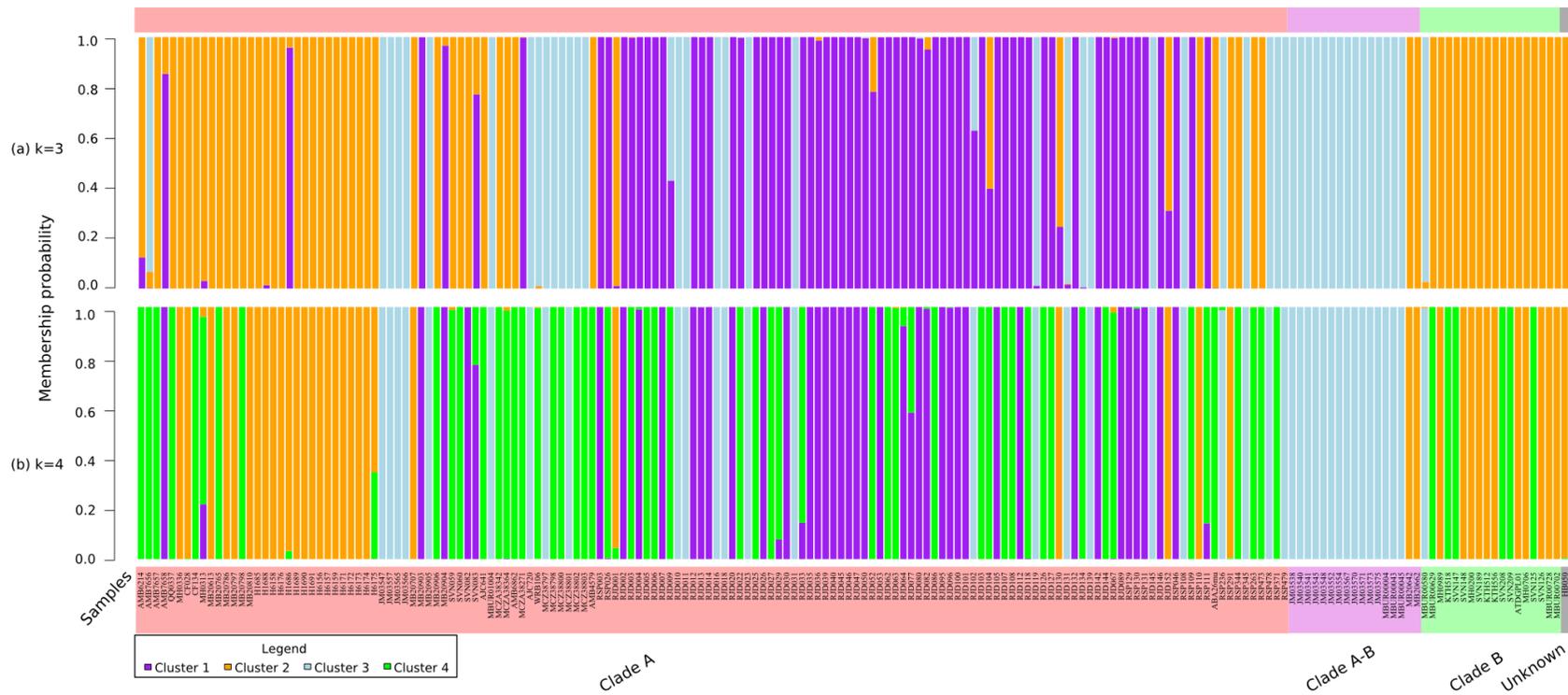
Figure 3.7: Assignment probabilities of all *Pedioplanis lineoocellata* samples from Clade A and B to the posterior clusters formed in the k-means analysis using microsatellite markers for three and four clusters, respectively. Bars above and below the figure indicate the mtDNA clade to which each sample belongs based on their geographic position within the clades' distribution. Colours of membership assignment proportions match cluster colours in the legend.
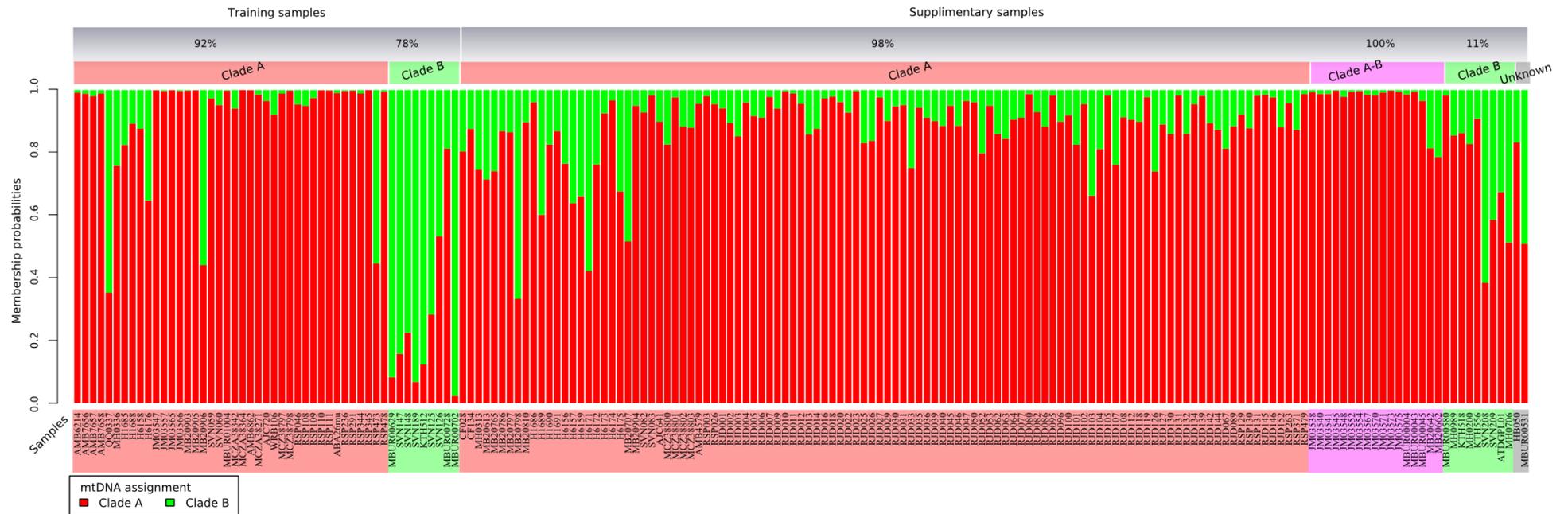
Figure 3.8: Assignment probabilities of *Pedioplanis lineoocellata* individuals to one of the two *a priori* mtDNA clades, Clade A and Clade B. Principle components and discriminant functions optimised using the training data. Proportion of individuals successfully re-assigned to the *a priori* group indicated above each group as a percentage. For supplementary data, *a priori* groups were based on geographical position with the known distribution of the clades with the exception of two samples that did not fall within any of the known clades- "Unknown". Assignment proportions above the 'Clade A-B overlap' region are for Clade A assignments. Bars above and below the figure indicate the prior mtDNA clade. Colours of membership assignment proportions match clade colours in the legend.
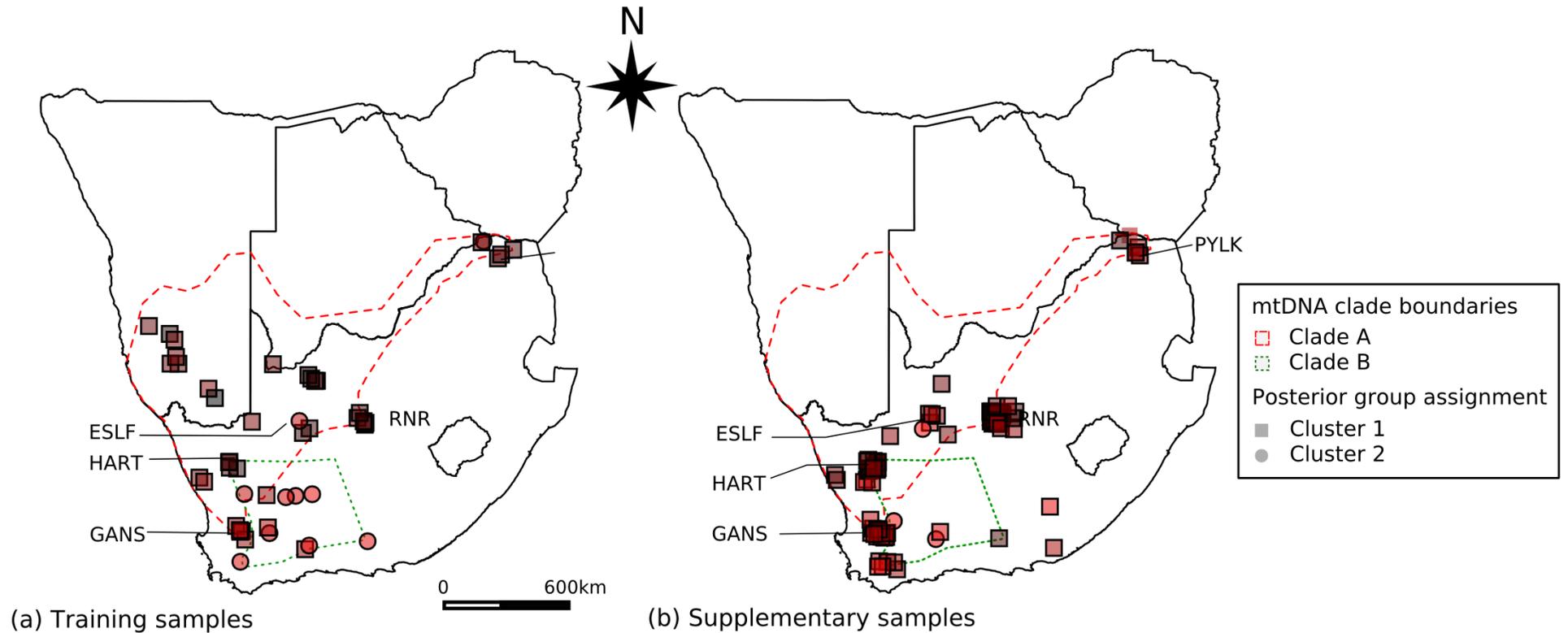
Figure 3.9: Clade boundary comparison between mtDNA Clade A and B, and nuclear microsatellite clusters for *Pedioplanis lineoocellata*. Posterior clade assignments determined using DAPC analysis based on training data with *a priori* groups. (a) Cluster assignment for training data and (b) cluster assignment for supplementary data. The colour intensity of individual symbols indicates genetic similarity as determined by weighting the colour intensity of red according to the co-ordinates from the Discriminant Analysis.

4   DISCUSSION

The Spotted sand lizard, *Pedioplanis lineoocellata*, shows marked genetic structuring on a large geographic scale covering most of its distribution across South Africa and Namibia, as indicated by analysing nine newly developed microsatellite nuclear markers. Contrary to the predictions, the geographic arrangement of the clusters did not correspond to the mitochondrial DNA (mtDNA) clades, Clade A and Clade B using ND4 and 16S markers (Edwards, 2013). The incongruence suggests more recent nuclear gene flow between the mtDNA clades and a contact/overlap zone that is potentially far greater than previously anticipated.

As should be expected, levels of gene flow between sites were greater when geographic separation was low (i.e. a few kilometres across RNR) compared to when separation was large (hundreds of kilometres). Gene flow within RNR was supported by non-zero migration estimates and low pairwise F-statistics. Studies of other taxa over similar geographic scales have found comparable results (e.g. Wang et al. 2009; Dubey et al. 2011). Albeit there is support for gene flow at this scale, BayesAss estimates of m can be inaccurate when population differentiation is very low ($F_{ST} < 0.01$) therefore assessing the contemporary connections between sub-populations within RNR in more detail may not be possible (Faubet *et al.* 2007). For sample sites examined across the distribution of Clade A, sites in close proximity were not necessarily more genetically similar than those further apart, as reflected by the failure to detect the isolation-by-distance (IBD) pattern across the range and DAPC clustering. However, significant F-statistics, zero migration estimates between sites and the near-absence of mixed membership assignments do support the genetic distinctiveness of each of the eight focal sample sites as a population with restricted gene flow between sites. Genetic differentiation was comparable to that between populations of various other lacertids such as *Lacerta viridis* (Böhme & Moravec 2011), Swedish and Hungarian *Lacerta agilis* (Gullberg et al. 1997), between strongly divergent populations of *Iberolacerta monticola* occurring along the continuous Iberian Peninsula and as isolated and reduced populations (Remón *et al.* 2013) and, lastly, between species of *Lacerta lepida* and *Lacerta nevadensis* (Miraldo *et al.* 2013).

The nuclear microsatellite markers revealed an unexpected complexity in the geographic arrangement of the microsatellite clusters, which was likely the reason no IBD pattern was detected. The discriminant analysis of principle components (DAPC) indicated that there were geographically wide-spread microsatellite clusters and that individuals from some sites, particularly RNR, were not assigned to the expected clusters. The most wide-spread cluster included individuals from the entire sampled range within Clade A. This cluster included ESLF but collapsed to include GANS samples when the number of allowed clusters was reduced, indicating greater similarity between ESFL and GANS, than between other sites.

HART and PYLK also clustered together but only with low cluster numbers. The result suggests a more recent genetic link between the former than the latter. Site ESLF had the lowest divergences from GANS based on F-statistics and there was evidence of migration from GANS to ESLF, despite that the geographical distance from ESLF to HART is half that of the distance from ESLF to GANS.

Contrary to the prediction that HART would have signals of genetic admixture, there was no evidence of genetic admixture between clades at the sample site and it clustered closer to PYLK, the furthest site geographically, than it did to more proximal sites. For example, samples from Garies were only ~150km east of HART, yet clustered with ESLF and RNR. HARTs affinity with PYLK may be a relict connection as the sites are geographically distant from each other and, together, are most divergent from other sites.

RNR did not form a well-defined group despite having the largest sample size of any of the sites, instead the site consisted of individuals assigned to various other genetic clusters, including an unidentified cluster (i.e. no site consistently represented that cluster, not even sites within RNR). Sites with comparable sample sizes, and even those with fewer samples, had formed better defined clusters than RNR. This might result from secondary contact at RNR between various genetic clusters but with the low frequency of individuals with mixed posterior membership assignments, genetic exchange between these different clusters is not supported. The sympatric occurrence of individuals from different microsatellite clusters without evidence of admixture would suggest reproductive barriers and cryptic species. However, this seems unlikely as Edwards (2013) found genetic divergence estimates between mtDNA lineage Clade A and Clade B to be below the typical between-species level for the *Pedioplanis* genus therefore divergence between members within the Clade A lineage should be even lower. The most likely cause for the peculiar result may be the analysis and sampling. The DAPC analysis relies on highlighting variability between populations over variation between individuals and the precision with which clusters are formed may be negatively affected by low per-population sampling and geographically patchy sampling. However, this explanation fails to explain why the better sampled site showed the greatest mis-assignment. The DAPC does not make assumptions regarding genetic evolution when summarising genetic data and therefore the results may be different when analyses are performed under more stringent genetic framework such as in the program STRUCTURE or Geneland (Pritchard et al. 2000; Guillot et al. 2005). A post-hoc analysis using STRUCTURE produced results that were not difference from that of the DAPC and the overall conclusions remained unchanged. The integrity of RNR was better supported (Information A 2) but a better developed conclusion awaits more comprehensive local-scale sampling at RNR including populations some tens of kilometres away. Additional sites between the eight focal sampling areas as well as in parts of the region poorly represented in this study should be considered. Using additional analyses such as least-cost-path analysis may shed additional light regarding from which direction lizards enter each population.

When examining the relationship between Clade A and Clade B, the results show that even when forcing prior group assignments, the training data (samples genotyped for mtDNA) did not cluster into recognised mtDNA clades. When largely unaffected by genetic processes such as selection, genetic drift and gene flow, the utilised microsatellite markers should corroborate the genetic distinctiveness of mtDNA clades or produce additional resolution, yet this is not seen from the results (Dakin et al. 2004; Glenn et al. 2005). Mitochondrial gene regions are considered to reflect older genetic events compared to microsatellites because the former have lower mutation rates allowing them to retain better historic signal (Ellegren 2004; Selkoe et al. 2006). Provided the distribution of the mtDNA clades were not biased by chance sampling and low sample sizes (discussed below), the discrepancy between the different genomes would indicate that microsatellite patterns reflect more recent genetic events than do the mtDNA markers.

Gene flow between the mtDNA clades subsequent to their initial formation could lead to different signals for mitochondrial and nuclear genomes. Many species with limited dispersal ability or those that occur as isolated populations have population-level genetic signals that correspond to deeper phylogenetic signals (e.g. Dolman & Moritz 2006; Duftner et al. 2006). As an open-habitat species with a wide distribution across several habitat types, I suspect *P. lineoocellata* has strong dispersal capabilities. Migration estimates and F-statistics do provide support for this across small geographic distances (i.e. within RNR) but not across the distribution of Clade A. The differences between conclusions drawn from DAPC compared to F-statistics and migration estimates may be a consequence of different geographic and temporal scales at which the analyses work. BayesAss migration estimates can only detect recent (~3 generations at most) immigrants by their divergent genetic signal relative to a defined population. With sampling sites separated by hundreds of kilometres, it seems unlikely that one would find recent migrants between sites. In contrast, the DAPC analysis summarises the variation between all included samples, thus reflecting patterns more strongly influenced by gene flow, both recent and historic. The F-statistics too reflect longer term gene flow but the results do not necessarily contradict the DAPC results as no Clade B populations were included in the calculations of the F-statistics. Therefore no comparisons can be made of between- and within-clade F-statistics. The gene flow that caused the loss of microsatellite distinction between Clade A and B may have been a rather historic event that is no longer reflected in the contemporary dynamics between populations. For example, gene flow between sample sites within Clade A may not reflect the gene flow dynamics between populations at the Clade A-B boundaries. From the DAPC analysis, Clade A sample sites near the boundary, in the western south-west region of South Africa, have a notably greater level of mixed membership assignments between clades compared to the Clade A sample sites further away. The result supports possible gene flow. Although the F_{ST} values for one of these sites, GANS, indicates significant

genetic differentiation when compared to other Clade A sites, gene flow from Clade B to GANS, and possibly other sites, may not be as restricted.

Differences in gene flow between the sexes of many organisms have been used to explain discordance in genomic signals between mtDNA and nuclear DNA (Seielstad et al. 1998; Arctander et al. 1999; Destro-Bisol et al. 2004; Tishkoff et al. 2007; Ujvari et al. 2008; Barlow et al. 2013; Miraldo et al. 2013). Genetic structure based on bi-parentally inherited nuclear genes will respond more rapidly to gene flow under male-biased dispersal than maternally inherited mtDNA would. Sex-biased dispersal in lacertids is well documented with strong evidence for male, particularly juvenile, biases (Olsson *et al.* 1996; Meylan *et al.* 2002; Miraldo *et al.* 2013) related to the prevalence of polygynous mating systems. Incongruence between nuclear and mitochondrial patterns based on sequences data for *B. arietans* (Barlow *et al.* 2013) and of microsatellite data for other reptile species (Chapple et al. 2005; Dubey et al. 2011) have similarly been explained through male-bias in dispersal. However, the complexity of the geographic arrangement of microsatellite clusters for *P. lineoocellata* cannot be explained only by male-biased dispersal and a more focused local-scale analysis examining patterns between sexes is needed to confirm biases in dispersal.

Selection is another possible cause for discordant genetic signals as it can create genetic patterns that reflect niche use or sexual selection rather than gene flow. Although microsatellite loci are generally thought to be selectively neutral, loci may be linked to gene regions that are under selection (e.g. Li & Merilä 2010). Similarly, mtDNA gene regions used for sequencing may also be subject to selective pressures and differential selection on genomes could create different signals. The microsatellite loci in this study could not be tested for selection due to the low number of well-sampled populations and few loci available therefore local selection cannot be ruled out as a potential cause for the differences between mtDNA and microsatellite signal.

Genetic drift could also create mismatch between genomes. As mtDNA has one-fourth the effective population size of nuclear DNA, erosion through genetic drift should be faster (Li et al. 2010) and result in the more rapid formation of distinct mtDNA lineages compared to microsatellite clusters (e.g. Hellborg et al. 2002). However, it seems improbable that genetic drift could generate such wide spread mtDNA clades without the need to evoke range expansion for each lineage from a more restricted geographic range.

Lastly, the discrepancy between the mtDNA clades and microsatellite clusters may be the results of more extensive geographic overlap than was previously found. Sequence-based phylogenetics can be accomplished with far fewer samples than is necessary for co-dominant population-level genetic markers such as microsatellites. For many of the sites sampled by Edwards (2013), there were less than five samples therefore some of the other mtDNA lineages present at several sites, by chance alone, may not have been

sampled. Several mitochondrial lineages could be present but at low frequencies at many of the sites and the relative proportion of each lineage within a population may vary as a cline across the Clade A-B distribution. This scenario does not preclude possible gene flow between lineages either. Low sample sizes at several populations for the microsatellite profiling will also have affected the resolution of the results. Therefore, distinguishing between the possible causes will require sequencing of mtDNA regions and microsatellite profiling for several (~10-30) individuals from several sites and the sampling in this study was not sufficient for this purpose. Succeeding studies may wish to consider sampling populations from Namibia, Botswana and the areas north and/or east of RNR, in addition to improving sampling at the sites included in this study. The other two mitochondrial clades may also be considered.

The presence of different mitochondrial lineages at HART was thought to indicate the possibility of genetic admixture between mtDNA clades at the Loeriesfontein sample site but this was not supported by the results. The individual genotyped as Clade B was not successfully profiled in this study and therefore the nuclear DNA cluster to which it would assign is still unknown. Based on the absence of mixed assignment probabilities, the zero migration estimates with other sample sites and the large pairwise F-statistics, there is little support for admixture between clades, or even between populations within the distribution of Clade A, at HART. The presence of different mtDNA lineages sympatrically at this site may be a consequence of better sampling compared to other sites. Alternatively, the individual sampled by Edwards (2013) may have been a recent migrant from a Clade B population. Migration rates between Clade A and Clade B populations were not estimated but estimates of recent migration between Clade A sites were zero over large distances. It seems unlikely that the individual would be a recent natural migrant given that the nearest Clade B population from HART may be as far south as the Klein Karoo (over 200km). It remains more likely that different mtDNA lineages occur sympatrically at HART and most likely at several other sites as well and the lineages were simply not sampled in previous surveys by chance.

Edwards (2013) proposed that climate oscillations of the Plio-Pleistocene may have generated the mtDNA lineages of *P. lineoocellata* however, there are no divergence date estimates with which to compare to the estimated dates for these events. Parapatric mtDNA lineages of many southern African species have been associated with glacial-interglacial climatic cycles, and in some cases secondary contact has been found between lineages (Tolley *et al.* 2010a; Edwards 2013; Barlow *et al.* 2013). Several southern Africa open-habitat taxa have had proposed refugial areas within which these taxa persisted during glacial periods (Lorenzen *et al.* 2010, 2012; Fuchs *et al.* 2011; Barlow *et al.* 2013). Many of the Eurasian lacertid populations examined have diverged and even speciation after prolonged separation in refugia during glacial periods (Zhao *et al.* 2011; Remón *et al.* 2013; Miraldo *et al.* 2013). The distribution and breaks between the *P. lineoocellata* mtDNA clades match those of other taxa, particularly in the western parts of South Africa

(e.g. Daniels et al. 2007; Swart et al. 2009; Tolley et al. 2010a; Smit et al. 2011). Furthermore, the most divergent clade within *P. lineoocellata* (Clade C) is found in the Fynbos along the south coast, a biome thought to be relatively climatically stable (Tolley *et al.* 2014a). Microsatellite DNA appears to reflect genetic events that are possibly much younger than climate oscillation of the Plio-Pleistocene and may be the result of recent expansions. Based on microsatellite data, the extent of the contact region between mtDNA Clades A and B may be far greater than previously thought, possibly including the central areas of South Africa as suggested by the most geographically extensive of the identified microsatellite clusters. Following the last glacial period the species would have expanded its range to occupy the current distribution subsequent to improvements in the conditions of neighbouring areas. If this is the case, then sympatric occurrence of mtDNA lineages might not be unusual. Such a scenario may result from secondary contact between clades or in cases of incomplete lineage sorting.

In contrast to the wide spread microsatellite cluster, populations from Loeriesfontein appear restricted to the moderately stable Succulent Karoo region, the west coast area south of the Orange River, including the Knersvlakte (Tolley *et al.* 2014a). The Loeriesfontein population, HART, also showed a deep connection to PYLK that was interpreted as a historic signal of separation from the remaining sample sites. HART may be an old population of the Succulent Karoo while PYLK may be a similarly historic population in the Limpopo Province. This could explain the lack of evidence for hybridization despite being geographically embedded within the wide-spread microsatellite cluster, which included ESFL. Furthermore, the overall lower F-statistics for GANS and ESLF and support for migration between the two sites would suggest a greater dispersal propensity for *P. lineoocellata* from these clusters compared to the others.

5 CONCLUSION

Across the distribution of *Pedioplanis lineoocellata* there is evidence for mismatch between the microsatellite clusters obtained and the patterns of mitochondrial DNA (mtDNA) clades (Edwards 2013). I suggest that the discordance is the result of gene flow subsequent to the formation of the mtDNA clades and a possible result of secondary contact with a contact zone that is larger than previously thought. Within Clade A, sample sites were genetically differentiated but there was no support for isolation-by-distance patterns, most likely because of the geographic arrangement of the microsatellite clusters. The identified microsatellite clusters may reflect range contraction and expansion events, however the signal differs between clusters. The cluster that included GANS and ELSF may have expanded its range recently to cover the large geographic distribution it currently occurs over. The cluster to which PYLK and HART belong appears more geographically restricted. Although Clade A and B lineages were found in sympatry at the Loeriesfontein population, HART, there appears to be no evidence to support genetic admixture between

clades. Sympatric clades may exist at other sites too, but have been unsampled to date, possibly due to disproportionate frequency of the two mtDNA lineages. The extent of contact between clades remains unknown. The relationship between populations and their geographic arrangement is still rather unclear and the present data set is too poorly sampled to provide the resolution needed to address questions further. The study highlights the need for large samples per site (>30) even within a few kilometres, and simultaneous sequencing of mtDNA and microsatellite profiles, particularly when addressing questions of hybridization.

Chapter 4 DISPERSAL ESTIMATES AND SPATIAL AUTOCORRELATION

INDICATE STRONG DISPERSAL IN THE SPOTTED SAND LIZARD

(*PEDIOPLANIS LINEOOCELLATA*)

1   INTRODUCTION

Population genetic structure is a consequence of biological processes including social behaviour, dispersal, mating system and demographic dynamics, all of which are inter-related (Halpin 1987; Shields 1987; Clobert et al. 1994; Andreassen et al. 2002; Chapple et al. 2005). Dispersal, in particular, is an important cause for, and consequence of, the evolution of various life history traits (Halpin 1987; Shields 1987). Understanding dispersal dynamics is important given that a species' response to contemporary conservation threats is determined, in part, by the dynamics of dispersal under these novel conditions (Midgley *et al.* 2006; Duckett *et al.* 2013).

Dispersal is a primary determinant of gene flow but individuals must survive the 'transfer' and immigration process, persist in the new environment and then successfully reproduce for gene flow to occur (Bohonak 1999; Andreassen *et al.* 2002; Raybould *et al.* 2002). Gene flow regulates genetic differentiation between geographically separated populations and is one of the processes which controls the development of population structuring. Much work has attempted to describe the theoretical relationship between dispersal and geographic genetic structure, but for many species the relationship remains to be investigated empirically (Wright 1935, 1946; Rousset 1997, 2000; Raybould *et al.* 2002).

Genetic or demographic measures of dispersal may be used to investigate dispersal biology empirically. However, there is a growing interest in addressing both in unison (e.g. Clobert et al. 1994; Sumner et al. 2001; Chapple & Keogh 2005; Dubey et al. 2011). Broadly speaking, measures of dispersal are either through observing physical movement of individuals using techniques such as capture-mark-recapture and telemetry, or genetic estimates made at population level using various markers to estimate gene transfer. Although demographic techniques directly measure dispersal, they are often only measured over relatively short time periods compared to the longevity of the study organism and are often geographically restricted relative to the species range. Consequently, direct measures likely miss important but rare events in dispersal, changes in inter-population dispersal dynamics over time and the variation across the species range (Bullock et al. 2002; Templeton et al. 2011; Clobert et al. 2012; Nathan et al. 2012). In contrast, measures of dispersal based on population genetics are mainly estimated through gene flow and migration rates using equilibrium based measures like Wright's F-statistics (Wright 1978; Raybould *et al.* 2002). These estimates are strongly influenced by evolutionary processes and are "averages" over an evolutionary time scale rather than current indicators of dispersal. Genetic estimates have seen increased use because they can be obtained relatively easily, allowing one to assess several populations in a short space of time (Bossart et al. 1998; Bohonak 1999).

When dispersal is estimated independently using either of the above approaches, the estimates can be compared to assess how well observed dispersal translates into measurable gene flow (e.g. (Whitlock et al. 1999; Sumner et al. 2001; Watts et al. 2007; Dubey et al. 2011). However, the differences in temporal scale at which the two types of estimates are made has raised concerns regarding comparability (Lebreton et al. 1992; Whitlock et al. 1999). There are several genetic methods of assessing gene flow and/or dispersal at small geographic or temporal scale that rely on smaller units of sampling than the conventional population-level analyses and could partially address the disjunction between demographic and population-level genetic measures. Individual-based methods such as parentage assessments, population membership assignments and spatial autocorrelation are examples of such methods (e.g. Sumner et al. 2001; Stow et al. 2004; Laloi et al. 2004; Chapple et al. 2005; Dubey et al. 2011).

Spatial autocorrelation of relatedness (SAC) and dispersal estimation from the neighbourhood concept are two individual-based genetic analyses that are not as commonly used as other methods but have demonstrated power (Sokal et al. 1989; Smouse et al. 1999; Sumner et al. 2001; Watts et al. 2007). Spatial autocorrelation of genetic relatedness is useful for testing for an isolation-by-distance pattern (IBD) especially where the signal may not be strong enough across the entire sampling scale to be detected using a Mantel test. SAC explores correlations at multiple lags (spatial distance classes) making it more powerful than a standard Mantel test and useful for identify barriers to gene flow and the spatial extent of genetic structuring (Peakall *et al.* 2003; Duckett *et al.* 2013). It has also been used to identify group-specific biases in dispersal (Peakall et al. 2003; Chapple et al. 2005; Keogh et al. 2007).

Estimates of dispersal can be obtained from genetic data using the neighbourhood concept (Wright 1943) and an individual-based genetic distance regression developed by Rousset (2000) as an alternative to direct measures of movement. An obvious advantage of this is that potentially dispersal estimates could be made from data collected within a single field session as opposed to several long-term studies. The method has been used for species of small mammal, lizard and insect, and found to produce estimates of dispersal comparable to real-time direct estimates (Rousset 2000; Sumner *et al.* 2001; Broquet *et al.* 2006; Watts *et al.* 2007; Duckett *et al.* 2013). Another appeal of the method is that it requires the inclusion of demographic parameters, which may be determined independently using either direct or indirect methods and the dispersal estimates based on the demographic parameters can be compared to assess agreement (Sumner *et al.* 2001; Watts *et al.* 2007).

The dispersal-gene flow relationship for terrestrial vertebrates is better studied for birds and mammals than for reptiles (Keogh *et al.* 2007).Yet lizards are good study organisms for testing basic hypotheses in dispersal biology because many of the complications of social behaviour and parental care are removed

(Clobert *et al.* 1994). Dispersal has been well studied only in a few species of lizard, particularly Eurasian taxa and some Australian species (e.g. Stow et al. 2001; Chapple & Keogh 2005; Clobert 2012; Miraldo et al. 2013). Southern African species, however, are poorly studied in this regard. *Pedioplanis lineoocellata,* a southern African lacertid, is a fairly well studied species although, most studies have focused on foraging mode and, more recently, phylogenetics (e.g. Wasiolka et al. 2009b; Blumroeder et al. 2012; Conradie et al. 2012). To my knowledge, only one study of demographic parameters has been carried out (Wasiolka 2007) but no work has addressed dispersal to date. The availability of other biological information for *P. lineoocellata* allows predictions to be made and the results to be supported or challenged based on known biology making this species a good candidate for dispersal estimation.

A dispersal movement is often defined as a permanent movement over a distance some-fold greater than the home range size diameter (Shields 1987). I approximate dispersal for *Pedioplanis lineoocellata* based on home range size estimates ($200 - 600m^2$) for adult males in the Kalahari area of the Northern Cape, South Africa (Wasiolka 2007; Wasiolka *et al.* 2009b). By assuming roughly circular isotropic home ranges, the home range diameter should range between 16-28m and thus dispersal (being a few fold greater) should range within a few tens of metres. Furthermore, neonatal and often adult dispersal in small lizards, such as most lacertids, is seldom greater than some tens of metres (Clobert et al. 1994; Olsson et al. 1996, 2003; Sumner et al. 2001), although estimates based on genetics have been as great as several kilometres for some species (Hoehn *et al.* 2007; Duckett *et al.* 2013). Overall, annual dispersal distances should be comparable to estimates for other lacertids.

Pedioplanis lineoocellata is widely distributed across southern Africa with evidence for at least four mitochondrial DNA (mtDNA) lineages within the distribution (Edwards 2013). Clade A, which occurs within the savanna and Karoo interior of southern Africa, is particularly wide-spread (Edwards, 2013). Based on the isolation-by-distance model, there should be a pattern of higher relatedness between individuals at close proximity compared to those at further distances assuming restricted dispersal and the absence of selection. Species widely distributed within homogenous macro-scale habitats, such as the Nama Karoo and savanna, are ideal for testing for an IBD pattern because, compared to species within more heterogeneous habitats, there should be fewer processes disrupting the IBD pattern. Yet no such pattern was found for *P. lineoocellata* across the distribution of Clade A (Chapter 3). The IBD pattern may have been disrupted by as yet unidentified landscape features or historic environmental factors resulting in the complicated geographic arrangement of microsatellite groupings found in Chapter 3. A spatial autocorrelation analysis is more sensitive to patterns of changes in relatedness and may detect the presence of patterns not detected by the Mantel test (Smouse et al. 1999).

In this chapter the NS-based estimator of annual dispersal and the individual-based spatial autocorrelation method are used to estimate dispersal distance and test for an IBD pattern, respectively. Other studies have found a fair amount of agreement between NS-based dispersal estimates and those made using direct measures, typically within a two-fold range (Sumner *et al.* 2001; Watts *et al.* 2007). Dispersal distances for *P. lineoocellata* are predicted to be within the range of a few tens to a hundred metres. Estimates of effective population density need to be included in the estimation of annual dispersal distances and the population density estimates may be determined using direct (demographic) or indirect (genetic) methods. This allows for the comparison of dispersal results based only on genetics and those including demographic estimates. The estimates of effective population density are predicted to correspond.

No spatial genetic structure is predicted for the SAC when sampling at a local geographical scale of a few kilometres because, at this scale, the distances between individual samples can easily be traversed over few generations. In Chapter 3, sites within Rooipoort Nature Reserve (RNR) were shown to have significant genetic differentiation but some between-site migration, which was interpreted as evidence of dispersal between sites at this scale. Therefore, the SAC analysis is predicted to detect spatial genetic structure in the form of an IBD pattern across the broad-scale within RNR. Finally, at a regional scale across the distribution of mtDNA Clade A there are very large geographic distances between sample sites and individuals as a consequence of the geographic scale of sampling. In Chapter 3 the F-statistics indicated clear genetic structure of sampling sites and very limited evidence for migration despite there being no isolation-by-distance pattern. The more sensitive SAC analysis is predicted to support the presence of an IBD pattern between individuals at the regional scale.

2 METHODS AND MATERIALS

2.1 STUDY SITES

Analyses were conducted at three different geographic scales; 1) local-scale, 2) broad-scale (within Rooipoort Nature Reserve) and 3) regional-scale (across the distribution of Clade A). At the local-scale, three different Northern Cape (South Africa) sampling sites were compared, two (RNR1 and RNR2) at Rooipoort Nature Reserve (RNR) and the third, in the Loeriesfontein area at Farm Hartebeeslaagte (HART) (Figure 4.1 and Table 4.1). Rooipoort Nature Reserve is approximately 60km west of Kimberley and situated within the savanna biome whereas farm Hartebeeslaagte is situated in the Nama Karoo biome approximately 520km south-west of RNR. Rooipoort Nature Reserve was chosen for field sampling because the study species is known to be abundant at this reserve. RNR data were collected specifically for the genetic and capture-mark-recapture work in this thesis whereas HART was not sampled for this thesis but was included

because it had a large sample size and GPS coordinates for each sample, allowing for both SAC analysis and NS estimation.
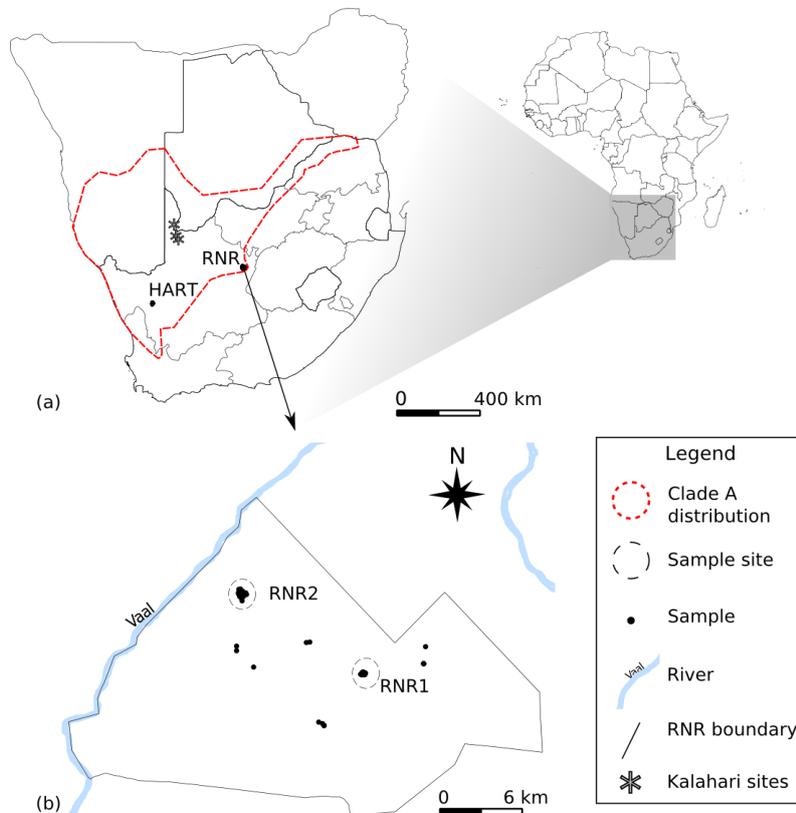
Figure 4.1: Local-scale study sites (a) HART and RNR in South Africa. (b) Capture-mark-recapture sites (RNR1 & RNR2) and the additional sites included in the broad-scale analysis within RNR. * indicates the study sites used by Wasiolka (2007) (see Annual dispersal distances below for details).

The original study design included repeated data collection from a single site (RNR1) to compare population densities between seasons. However, an unplanned fire burnt large parts of the reserve a few days prior to the second session in early October 2012, during which the primary site was completely burnt. The new site (RNR2) was selected because it was the closest (11km away) large and unburnt area of similar vegetation to RNR1. The two sites were identified by performing a two-day survey at the beginning of each field session that involved selecting areas where lizards were present and the habitat contained what was considered to be suitable vegetation and scattered rocks (Branch 1998; Plessis et al. 2011).

Table 4.1: Sampling sites and sample sizes (n) for the five datasets for *Pedioplanis lineoocellata*. Hull areas were based on a minimum convex hull around the individual samples. Total Area calculations include the site-specific mean maximum distance moved (MMDM) buffer zone as calculated from capture-mark-recapture (CMR) data. Sample sizes and area measures for RNR1 and RNR2 differ between genetic and capture-mark-recapture datasets (see Capture-mark-recapture below for details).

Site Name	n	Hull area (MMDM)	Total Area
RNR1	35	8.12 ha (30m)	11.87 ha
RNR2	29	48.02 ha (169m)	104.99 ha
HART	19	18.07 ha (169m)	71.46 ha
Rooipoort Nature Reserve	77	-	12 400 ha
Clade A distribution	84	-	108×10^6 ha

(a)   (b)

Figure 4.2: Vegetation at the two sites RNR1 (a) and RNR2 (b), in Rooipoort Nature Reserve, Northern Cape. Both fall within the *Eragrostis lehmanniana – Tarchonanthus camphoratus* Shrubland identified by Mucina and Rutherford (2006).

RNR1 was sampled during late summer/early autumn 2012 (9 April -5 May 2012) and RNR2 was sampled during spring 2012 (14 October- 03 November). The local vegetation at both sites was a patchy distribution of grass tussocks, bushes and compact, barren soil with plenty of rock fragments littered around, including igneous (andesite) but predominantly exposed calcrete (pers. obs, 2012) (Figure 4.2). The

vegetation falls under the *Eragrostis lehmanniana – Tarchonanthus camphoratus* Shrubland though grass layers do open with plenty of uncovered soils (Mucina & Rutherford, 2006). A more detailed description of the vegetation is available elsewhere (Bezuidenhout 2009).

At each RNR site we searched a square area but the RNR2 area was increased from the size used at RNR1 because of lower encounter frequencies at the former (10ha, RNR1; 100ha, RNR2). Surveys were conducted by random walk, covering the whole area of the study site, between 9am-6pm (April-May) and 8am-6pm (October-November) according to the available hours of sunlight. Tracks of the two surveyors were recorded on both field occasions every day on a GPS device (GPSmap CS60, Garmin Series with ±3m accuracy) to monitor the distribution of survey effort. Search efforts were not equal throughout the surveyed areas therefore only a subset of the original data was retained based on a survey effort criterion. Only capture points from areas that were within ~5m of a search transect (the maximum detection distance for these small lizards) at least once every four surveyed days were included in defining the final dataset and estimating the sampling areas (Figure 4.4 and Table 4.1). For RNR1 and RNR2 only a subset of all the lizards sampled were used for microsatellite profiling therefore the sample sizes and sampling areas differ between genetic and capture-mark-recapture work (discussed in the respective sections). Sample collection at HART occurred in August 2006 and November 2008 for another study. No capture-mark-recapture work was conducted and limited information is available regarding the habitat or survey effort at this site. Comparisons between sites are confounded with season and mean lizard age (see Results), most likely due to the differences in survey timing and site locality. Results from RNR were also possibly influenced by the large fire event in October 2012. Therefore, comparisons between sites are not possible.

The second geographic scale examined, the broad-scale, included all samples from RNR- samples from RNR1 and RNR2 as well as an additional 13 individuals collected across the reserve (Figure 4.1 and Table 4.1), which together covered a large portion of the reserve (12 400ha of 42 467 ha). Finally, the regional-scale was examined with all the available samples collected from across the entire known distribution of mtDNA Clade A (see Chapter 3 for sample collection details; Table 4.1).

2.2 ANNUAL DISPERSAL DISTANCES

The neighbourhood concept (Wright 1943) and an individual-based genetic distance linear regression developed by Rousset (2000) can be used to estimate dispersal. Where a population occurs continuously, the neighbourhood may be considered the population of a region from which parents of individuals born near the centre are a sub-sample or drawn at random (Wright 1946; Shields 1987). Neighbourhood area is the size of an area inclusive of 86.5% of the parents (Wright 1978, pp. 302-303) and neighbourhood size (NS) is the

number of individuals breeding in a neighbourhood area (Wright 1978). Neighbourhood size can be used to determine, σ, an estimate of distance between natal and breeding site (Rousset 1997, 2000, 2004; Sumner *et al.* 2001).

$$NS = 4\pi D_E \sigma^2$$

Where D_E is the time-scaled effective population density and σ^2 is the 'variance' of the distance between adult birth site and offspring birth site with a mean orthogonal value of zero. More precisely, σ^2 is the mean square of axial parent-offspring dispersal distance per generation. The $D_E\sigma^2$ units do not depend on the arbitrary choice of time interval as density and dispersal can be scaled according to any chosen interval (Sumner *et al.* 2001). The equation can be solved for σ and an accurate dispersal estimate may be obtained when NS and D_E are estimated independently. An indirect estimator of NS was developed by Rousset (1997) using the inverse of the regression slope of $\frac{F_{ST}}{1-F_{ST}}$ against the logarithm of the geographical distance for isotropic dispersal in two –dimensions. This was further developed for individual-based analyses using, $\hat{a}$, a multilocus estimator of $\frac{F_{ST}}{1-F_{ST}}$ (Rousset 2000). Estimates of effective population density (D_E) can be estimated directly using capture-mark-recapture data or indirectly using genetic data. In this study, annual dispersal distances for the local- and broad-scale sampling areas were determined following the individual-based regression procedure above.

2.2.1   NEIGHBOURHOOD SIZE ESTIMATION

Using the microsatellite profiles of individuals generated in Chapter 2, neighbourhood sizes were estimated as the inverse of the regression slope between genetic differentiation estimator $\hat{e}$ and natural logarithm of pair-wise geographic distances between samples (pgd). The estimator $\hat{e}$ can be used as an alternative to the original individual-based estimator, $\hat{a}$ (Rousset 2000; Watts *et al.* 2007). The $\hat{e}$ estimator is less biased than the latter when dispersal distances are large and when estimates are in two dimensional space making it better at detecting IBD patterns with weak genetic structure (Watts *et al.* 2007). Analyses should be performed with pgd greater than the (as yet unknown) σ (Rousset 2000; Watts *et al.* 2007). However, the resulting bias is only significant for very small inter-individual distances and the removal of data could be more detrimental to the analysis should the σ value be large (Watts *et al.* 2007), as would be the case when dispersal distances are great. For this reason only comparisons of individuals less than 1m apart were excluded. Isolation-by-distance was examined with Genepop v4.2 using a one-tailed Mantel test with 1000 permutations and the ABC bootstrap procedure (Diciccio et al. 1996) was used to estimate the 95% CI around the statistic *r* (Raymond et al. 1995; Rousset 2008). The lower 95% CI was estimated from

the $\hat{a}$ estimator as it has been shown to be consistently more accurate for the lower CI value than the $\hat{e}$ estimator (Watts *et al.* 2007).

2.2.2   GENETIC D_E ESTIMATES

Genetic estimates of effective population density were made for the three local sample sites. The effective population size, N_E, was estimated using a sib-ship assignment method implemented under COLONY v2.0.11 (Wang 2008) using full-likelihood with long-run, medium-likelihood precision and no sib-ship prior. The sib-ship analysis has been shown to out-perform temporal, heterozygosity excess and linkage disequilibrium methods and is less sensitive to violations of the random mating assumption (Wang 2009). The loci-specific mean genotyping error rates were set according to each sample sites' INEst estimates (Chapter 2). The mating system parameter was set to "male and female polygamy" as both are ubiquitous among non-avian reptiles (Uller et al. 2008) but only for RNR2 was inbreeding included (based on the best INEst model). Effective density, D_E was determined by dividing the N_E by the area occupied by the profiled samples as determined by a minimum convex hull in QGIS (QGIS Development Team 2013). An additional buffer was added to the hull to account for temporary immigrants. The buffer size was determined according to the mean maximum distance moved (MMDM) in each study site. As no CMR data were collected at the HART site, the RNR2 (adult individuals) buffer size was used (see Demographic DE estimates below). The MMDM values were determined using the MMDM() function in secr package (Efford 2011) in the R environment (R Core Team 2013).

Only across the whole of RNR was the NS regression significant (see Results). Therefore I estimated dispersal distances at this scale as well, despite that this broad-scale analysis did not have geographically continuous sampling and therefore was not amenable to density estimation. Constant adult densities across the reserve were assumed and the mean D_E values (and corresponding CI bounds) from RNR1 and RNR2 were used for D_E across RNR.

2.2.3   DEMOGRAPHIC D_E ESTIMATES

The duration of the study was not long enough to measure dispersal directly. To incorporate demographic information into the dispersal estimates based on NS, I used genetic and demographic D_E estimates and compared the two. I estimated D_E directly using capture-mark-recapture (CMR) data and a spatially explicit capture recapture (SECR) analysis, as it provides robust estimations of population density (Efford 2011). The CMR part of the study was only conducted at two of the three local-scale sites, RNR1 and RNR2 (detailed below).

2.2.3.1 Capture-mark-recapture

Sampling was conducted at the sites described under Study sites above. Whenever a lizard was observed it was slowly approached and captured using an extendable fishing pole with a noose attached. This approach avoids the need to chase the lizards and can be done with no harm to the animals. It is less invasive and stressful on the lizards compared to being chased or the use of pitfall traps (e.g. Wilson *et al.* 2007), which puts the animal at risk from predation, overheating and dehydration. At the point of capture, a GPS waypoint was recorded. Captured individuals were returned to their point of capture (within ~3m accuracy) by early evening on the same day. Where processing samples continued after dark or if there was an onset of unfavourable weather, the lizards were kept overnight in cloth bags and released the next favourable day. Releases occurred around 10am to reduce the risk of mortality as ectotherms may be especially vulnerable when ambient temperatures are low as it influences mobility (Huey et al. 1976; House et al. 1980).

Each newly captured lizard was marked with a small (2x4mm) white-with-black-text elastomer label glued with a 1-2 drops of tissue glue (B/Braun AESCULAP Histoacryl ®) to the lower back of the lizard for easy identification from a distance (1-3m). Each mark consisted of a letter and a two digit number that uniquely identified that lizard. Variables measured from each lizard included the animal size as the distance between the snout tip and the posterior vent opening (Snout-vent length; SVL) and tail length (TL) using a digital calliper, body weight (0.1g accuracy), sex and the unique tag number. Photographs of all lizards were taken against a 1cm^2 grid paper background for documenting the individual. Individuals with clipped tails (collected for DNA, see Chapter 2) were suspected re-captures that had lost their tags, as such dorsal spot patterns were cross-checked against photographs to identify the recaptured individuals. Sex was determined for adults based on the presence of a hemipenal bulge (KAT, pers. comm.) at the base of the tail. Where the bulge was not obvious, eversion of hemipenes was used to verify the sex. Adults were distinguished from immature individuals using a SVL threshold of 42-43mm following Goldberg, (2006) and Wasiolka (2007). During the surveys, marked lizards were not approached. Instead, we read the tag numbers through 10x25 binoculars or a 26x optical zoom Nikon Coolpix L810. Marked individuals were not physically re-captured to reduce interference with their normal movement. For each re-sighting the tag number, date, time and GPS position was recorded.

2.2.3.2 SECR density estimation

Population density estimates and detection probabilities at RNR1 and RNR2 were estimated using spatially explicit capture-recapture (SECR) analysis. Data from all full-day surveys were included and each day was considered a re-capture occasion in similar style to that of Royle & Young (2008). I estimated lizard

densities using the secr package (Borchers et al. 2008; Efford 2011) under the R environment. The analysis has two parameters; g(0), which estimates the detection probability at the centre of a home range, and τ, which is a scale parameter describing how quickly the detection probability declines with distance from the activity centre (in the original literature, the parameter is denoted σ but I refer to it as τ to prevent confusion with the dispersal distance). Together these parameters estimate the detection probability as a function of distance from the activity centre, g(d). The method estimates the effective sampling area and thus directly, population density.

The 'Polygon' detector type was used and I included multiple captures of the same individual within a single occasion. I decided to use the default halfnormal detection function as there was no particular biological reason to favour any of the other functions. The most popular alternative detection function, the Hazard rate function, is used when home ranges are expected to have a 'hard' boundary such as when animals are strongly territorial (Efford et al. 2009). Lacertids are typically non-territorial (Martins 1994), therefore I did not expect a sharp boundary to the home range.

The inclusion of covariates in model design can account for heterogeneity that otherwise would make parameter estimates less precise. Re-sighting probabilities may potentially vary with body size, age or sex of the individual as well as temporal variables such as the weather on any particular survey day. Therefore a suite of potential influential covariates were included in model design (Detailed in Appendix Information B 1). At the initial capture, individuals were handled whereas re-sightings did not involve any handling thus there may have been a trap shy response following the first capture event. A learned-behaviour or trap-response covariate (b) was included on g(0). Unfortunately no covariates could be included on τ because the data were too sparse. The conditional likelihood was maximised, which allowed the inclusion of individual covariates such as SVL. The parameter for population density (D) was held constant.

Sex ratios of ~1:1 were found for *P. lineoocellata* at three different study sites (Wasiolka 2007) and similar results were found for the RNR2 site (see Results) thus no adjustments were made to *D* estimates to account for unequal sex-ratios. However, I suspected that the number of captured adults at RNR1 did not truly reflect their actual proportion of the population therefore, the density of juvenile lizards was used to infer the number of adults based on adult: juvenile ratios found by Wasiolka (2007) for the February-April period in the Kalahari area (Figure 4.1). No information is available on the age structure within adult populations nor the reproductive variance for neither males nor females hence I could not make adjustments for either on density.

2.3 SPATIAL AUTOCORRELATION OF RELATEDNESS

2.3.1   SAMPLE SELECTION

Trends in relatedness across geographic space were assessed by spatial autocorrelation (SAC) using the microsatellite profiles of *Pedioplanis lineoocellata* individuals sampled across the distribution of mtDNA Clade A (see Chapter 2 and Chapter 3 for details on sampling). Analyses were carried out on the full dataset including individuals with missing data in their profiles, and again with a subset of only individuals with complete profiles. There was no difference between the results of these two datasets, therefore only the full dataset is discussed below. Three different spatial scales of sampling were used for SAC; a local-scale (<2km pairwise geographical distances – pgd - between individuals), which aims to assess IBD within sample sites, a broad-scale (<15km pgd) assessing the SAC over several kilometres but only within RNR was there sufficient sampling to perform the broad-scale analysis. Lastly, this analysis was also performed at a regional-scale across the distribution of Clade A (over 1600km pgd).

2.3.2   SPATIAL AUTOCORRELATION

The spatial autocorrelation (SAC) analysis produces a correlogram examining the correlation coefficients (*r*-values) of genetic similarity at incremental distance classes (or lags). The *r*-value is a proper correlation coefficient and is bounded by 1 and -1 in a very similar fashion to Moran's-I (Smouse et al. 1999). A value of 0 indicates the complete absence of spatial genetic structure. If *r* is 1 there is a perfect positive autocorrelation indicating high relatedness between individuals at this lag. If *r* is -1 there is a perfect negative autocorrelation and individuals at this lag have low relatedness. A decline in correlation coefficients with progressive lags indicates an isolation-by-distance (IBD) signal at that spatial extent of sampling (Duckett *et al.* 2013). Computer simulation studies of autocorrelation patterns predict an initial positive coefficient, which should decline to zero and then become negative and subsequently oscillate between positive and negative (Sokal et al. 1983; Smouse et al. 1999). The initial intercept with the x-axis (i.e. *r*=0) is the spatial extent of positive genetic structure ('segs' hereafter), which is the geographic distances at which autocorrelation of the relatedness between individuals becomes negligible (Sokal *et al.* 1989; Peakall *et al.* 2003).

Investigations on a single spatial scale are valid at the specific spatial scale as are segs values, but the results may not be consistent with the use of different lag sizes even within the same study system (Peakall *et al.* 2003). Should lags be larger than the scale of genetic structure one would fail to detect the structure at all whereas if lags were far below the segs, there would be unnecessarily small sample sets and low statistical

power. Thus the extent of the actual (and as yet unknown) spatial structure must be less than the spatial extent of sampling (Smouse et al. 1999; Peakall et al. 2003). While SAC at a specific lag size can be useful for visualising changes in *r* with incremental lags and the position of the segs, more reliable estimates of segs are best obtained by performing a multiple distance class analysis (MDC) as outlined by Peakall et al. (2003). In an MDC analysis, the SAC is repeated at multiple lag sizes by varying the lag size between the minimum pairwise geographical distance (pgd) to ~1/3 of the maximum pgd. When there is structure present, the estimated *r* will decrease with increasing lag size and the lag at which the estimated *r* is no longer significant would approximate true extent of detectable positive spatial genetic structure i.e. the lag size exceeds the segs.

SAC between individuals was examined with a genetic similarity index estimated in the Excel macro GenAlExV6.5 (Smouse et al. 1999; Peakall et al. 2012). This genetic distance measure and the SAC method are intrinsically multivariate and do not rely on per-locus analysis (Smouse et al. 1999). A null expectation was generated for testing the significances of the *r* values by randomly shuffling the genetic data between individuals within each lag thus creating a completely random expectation of no spatial structure (Smouse et al. 1999). One thousand permutations were used and the values were sorted to retrieve the 95% confidence interval (CI). When the interest is in detecting positive autocorrelation at short distances, as it was in this case, a one-tailed test for positive spatial genetic structure is used. Additionally, bootstrap estimates of *r*-values were calculated with 1000 replications and the values were sorted to retrieve the 975th and 25th ranked values in order to determine the 95% CI interval about the *r*-values. A joint visual assessment for significance was used, as described by Peakall et al. (2003), whereby statistical significance is only declared when the 95% bootstrap CI about the *r*-value does not intercept the x=0 line and the estimated *r*-value lies outside the permutation 95% CI envelope about the null hypothesis of *r*=0. The bootstrap errors tend to be larger than the permutation errors, particularly for small sample sizes and will favour the null hypothesis more frequently than the permutation tests, making them more conservative estimates. For the MDC analysis, the single population option was used. Linear and logarithmically transformed geographical distances (in metres or km) were compared. Bootstrapping was set to 9999 for both the 95% CI values around the mean coefficient and the 95% CI values around the random distribution (*r*=0). Correlograms at specific lag intervals were used to visualise the change in *r*.

3 RESULTS

3.1 ANNUAL DISPERSAL DISTANCE ESTIMATES

3.1.1 NEIGHBOURHOOD SIZE ESTIMATES

Neighbourhood size (NS) for local-scale sample sites and the broad-scale site were estimated using a regression of the multilocus genetic distance estimator, $\hat{e}$, against the logarithm of the geographical distance. The one-tailed Mantel tests for all local-scale sites were not significant ($p>0.05$) indicating no isolation-by-distance at the local-scale (Figure 4.3). Only the correlation at broad-scale was significant ($p<0.001$) and positive, supporting an isolation-by-distance spatial pattern. However, sampling in this case was not continuous and violates an assumption of the NS estimation; possibly introducing bias.

Estimates of NS (i.e. the number of breeding individuals within the neighbourhood area) from the inverse of the regression slopes varied greatly with RNR1 at 945 individuals and HART at -79 individuals (i.e. ∞) (Table 4.2). In all local-scale data sets, lower bounds were >30 but the upper bounds were ∞. For sampling at broad-scale, estimates were more precise because the regression was significant but error margins were still wide (339 individuals, 95% CI 76 – 1787). Results indicate that local-scale sampling (<2km pgd) is too geographically restricted to detect the IBD pattern and sampling should be conducted across several kilometres at the least. Additionally, should sampling have been more comprehensive across RNR, the precision of the estimates may have been greatly improved.

3.1.2 GENETIC D_E ESTIMATES

The effective population density was estimated for each of the local-scale sampling sites using a sib-ship analysis. To account for an edge effect on the sampling area a mean maximum distance moved (MMDM) buffer was included in the calculation of the sampling area. For RNR1 the MMDM value was $30.2 \pm 4m$ and $169.5 \pm 55m$ for RNR2. Effective population densities estimated from the sib-ship N_E values ranged from 0.43×10^{-4} to $3.03 \times 10^{-4} \; individuals.m^{-2}$ (Table 4.3). Density estimates for RNR2 and HART were similar but were ~6 fold lower than RNR1.

Table 4.2: Isolation-by-distance regression for neighbourhood size (NS) estimation at populations of *Pedioplanis lineoocellata.* Indicated are the intercept and slope, p-value for one-tailed Mantel test, sample size (n) and the estimates of NS for each data set. The 95% confidence interval (CI) bounds for regression slope and NS indicated. The NS upper bound estimated based on $\hat{e}$ estimates and lower bound based on $\hat{a}$ estimates. Values in **bold** are significant at $\alpha = 0.05$.

Site	n	intercept	slope	p-value	NS (individuals)
RNR1	35	-0.0017	0.0011 (-0.0048 - 0.0222)	p=0.38	945.19 (45.11 - ∞)
RNR2	29	-0.0215	0.0046 (-0.0087 - 0.0286)	p=0.08	216.88 (34.91 - ∞)
HART	19	0.0823	-0.0126 (-0.0431 - 0.0194)	p=0.10	∞ (51.60 - ∞)
All RNR	77	-0.0220	0.0029 (0.0006 - 0.0131)	**p=0.001**	339.10 (76.39 - 1787.83)
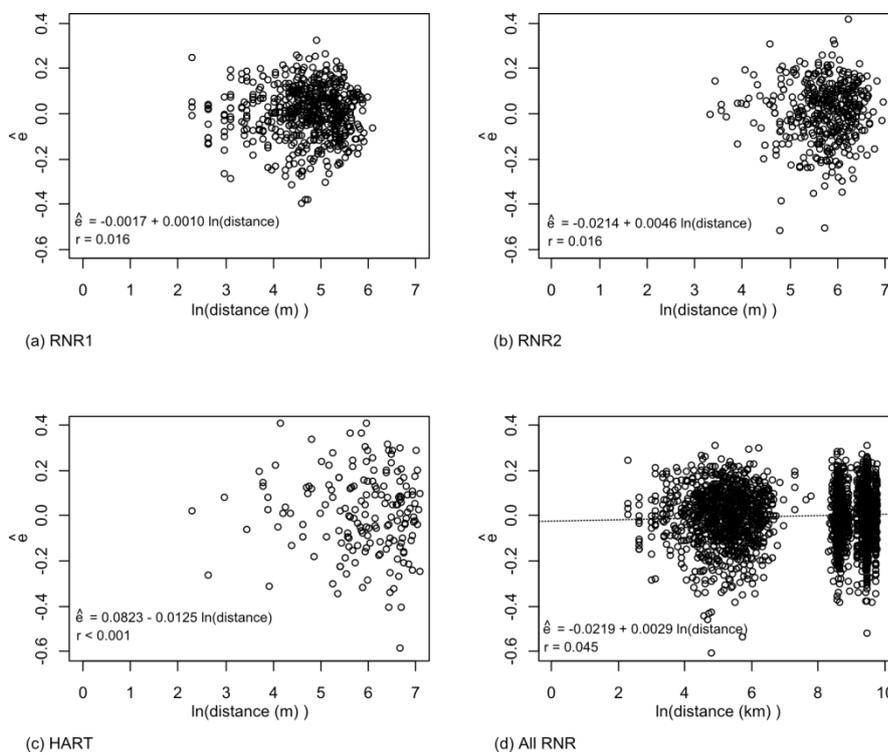
Figure 4.3: Scatterplots of genetic distance ($\hat{e}$) against the logarithm of the geographical distance for *Pedioplanis lineoocellata* at local-scale and broad-scale sampling. Trend lines indicated by lines and *r*-values indicated in plot.

Table 4.3: Annual dispersal distance estimates based on effective population densities from (a) kinship and (b) spatially explicit capture-recapture (SECR) estimates for *Pedioplanis lineoocellata*. Estimates for effective population size (N_E), effective density (D_E), second moment of parent-offspring distance (σ^2) and annual dispersal distance (σ) are indicated for each data set. The 95% confidence intervals (CI) for sib-ship D_E are based on 95% CI of N_E estimates over the sampled area (including mean maximum distance moved- MMDM- buffer). The α values are a non-random mating parameter of sib-ship analysis. SECR analysis effective sampling areas (ESA) indicated. * indicates D_E estimates for cryptic adult population in RNR1, adjusted by using juvenile density estimates and the adult: juvenile ratio estimated from Wasiolka (2007). # indicates values estimated using average D_E of RNR1 and RNR2. Maximum pairwise geographic distance (Max. pgd) at each site is indicated along with the corresponding 20σ value for evaluating possible bias (see Results). The 95% CI for all values indicated in brackets inside the table. All units indicated in brackets in column name.

(a)	Sib-ship						
Sites	Area (m^2)	n	α	N_E	D_E ($10^{-4}\ ind.m^{-2}$)	σ^2 ($10^3\ m^2.year^{-2}$)	σ ($m.year^{-1}$)
RNR1	118 732	35	0.00	36 (22 -67)	$3.03\ (1.85-5.64)$	248 (6.36 - ∞)	498 (80 - ∞)
RNR2	1 049 947	29	0.17	54 (29 -139)	$0.514\ (0.276-1.32)$	336 (21.0 - ∞)	579 (145 - ∞)
HART	714 553	19	0.00	31 (16 -74)	$0.434\ (0.224-1.04)$	∞ (39.6 - ∞)	∞ (199 - ∞)
Across RNR	-	-	-	-	$1.77\ (1.06-3.48)$#	152 (17.5 − 1336)#	$390\ (132-1156)$#

(b)	SECR				20σ (m)	Max pgd (m)
Sites	ESA (m^2)	D_E ($10^{-4}\ ind.m^{-2}$)	σ^2 ($10^3\ m^2.year^{-2}$)	σ ($m.year^{-1}$)		
RNR1	190 289	$1.69\ (1.23-2.33)$*	445 (15.4 - ∞)	667 (124 - ∞)	1730 - ∞	477
RNR2	348 041	$1.29\ (0.861-1.94)$	133 (14.3 - ∞)	365 (120 - ∞)	3486 - ∞	1123
HART	-	-	-	-	3982 - ∞	1274
Across RNR	-	-	-	-	2917 - 23977	18701

3.1.3   DEMOGRAPHIC D_E ESTIMATES

Demographic estimates of D_E were obtained using the spatially explicit capture-recapture procedure. Forty one individuals were caught during the April-May survey at RNR1, four of which were adults and the remainder were juveniles based on both SVL and sex identification (Table B 1). The mean SVL was 33.28±7.06mm (range 19.2-52.9mm, n=41). A total of 50 individuals were caught during October-November at RNR2, only one of which was not an adult (Appendix Table B 2). The mean SVL length was 51.12±2.90mm. Females were marginally larger than males (Females: 51.61±3.25mm n=30, Males: 50.67±2.51mm n=33) but the difference was not significant by a Welch two sample t-test (t = 1.26, d.f. = 54.4, p-value = 0.21). The sex ratio for the RNR2 data set was 1.12:1 (male: female), which is similar to

results (0.89-1.1) of Wasiolka (2007). A detailed summary of the sampling occasions at both sites can be found in Table 4.4 and capture occasion covariate details can be found in Appendix Table B 3.

Table 4.4: Summary of details for the survey occasions (season and survey effort) and capture data for *Pedioplanis lineoocellata* at both RNR1 and RNR2, Northern Cape. Some days were not surveyed due to unfavourable weather.

	RNR1	RNR2
Season (dates)	Autumn 2012 (11/04-03/05)	Spring 2012 (14/10-02/11)
Total days surveyed (not surveyed)	18 (5)	13 (6)
Total captures (recaptures)	41 (54)	50 (18)
Captures per day (mean±S.D.)	6.8±3.6	4.5±2.5
Daily temperature range (mean±S.D.) °C	19-32°C (25.9±4.5)	23-39°C (32.4±4.0)
Daily survey effort (mean±S.D.) hrs./day	7.5±1.0	7.4±1.9

The influence of adults in the April dataset was examined by running the spatially explicit capture recapture (SECR) analyses with and without adults for which there were no changes to the order of the top SECR models (results not shown) supporting the use of the entire dataset. Similarly, the sole sub-adult from the October dataset was included in the analysis. In both cases individuals were retained because of the sparcity of the available data, however SECR analyses have been performed on equally sparse data (Royle et al. 2008; Efford 2011) and in some cases far less data- a total of 10 recaptures for seven animals (Efford *et al.* 2009).
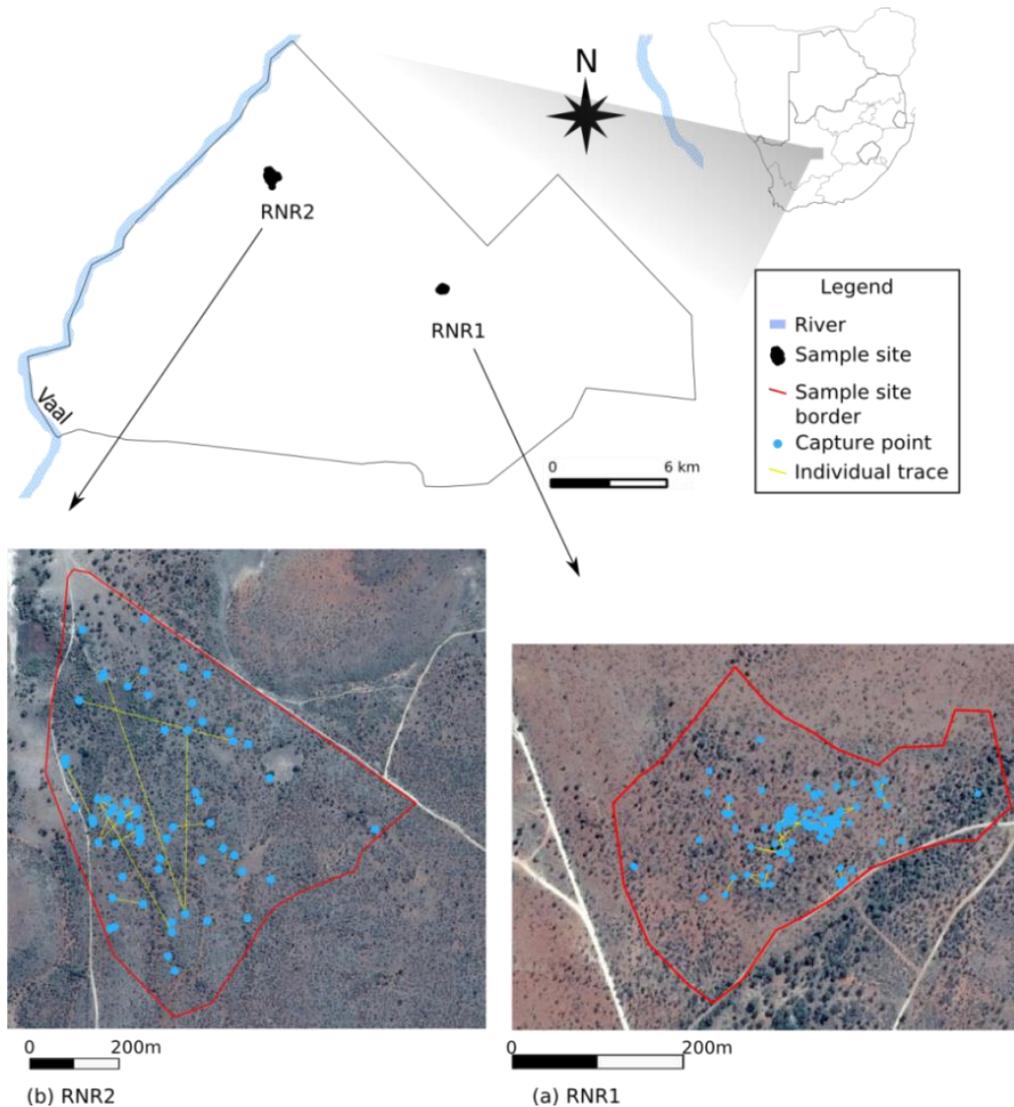
Figure 4.4: Distribution maps of the capture histories of *Pedioplanis lineoocellata* individuals from (a) RNR1 and (b) RNR2. Sample site borders (red) indicate the extent of the sampling areas and the yellow lines connecting blue dots (capture points) are the traces between sightings of a single individual. Inset map of southern Africa and RNR provided above (a) and (b).

For RNR1 the best SECR model included a learned effect on g0, but this model had similar support ($\Delta AICc < 2$) as models including covariates for rain fall events and SVL (Table 4.5). The g0 value for the best model was 0.13 (95% CI 0.08-0.21) for initial detection probabilities and 0.09 (95% CI 0.07-0.12) for subsequent captures in April (Figure 4.5). For the October data set the best model included minimum daytime temperature as a covariate. Re-capture probabilities at the activity centre declined from 0.09 (95% CI 0.05-0.16) with an increase in minimum daytime temperature (Figure 4.6). Based on the models with the greatest log-likelihood, activity area size in April ($\tau = 12.77$m, 95% CI 11.24-14.52) was smaller than that in

November (116.58m, 95% CI 104.42-129.98) and, as would be expected, the densities at the two sites differed notably. The density at site RNR1 ($2.52 \times 10^{-4}\ indiv.m^{-2}$) was estimated for the juvenile skewed dataset and needed to be adjusted to reflect adult density. The adult: juvenile ratio for April-February is 0.67 (Wasiolka, 2007), giving an adult density of $1.69 \times 10^{-4}\ indiv.m^{-2}$ for RNR1, which was higher than that of RNR2 ($1.29 \times 10^{-4}\ indiv.m^{-2}$).

Table 4.5: Top ten SECR models for *Pedioplanis lineoocellata* from (a) RNR1 and (b) RNR2. Snout-vent length (SVL) was included as an individual covariate. Trap effect (b) allowed initial and subsequent encounter probabilities to differ. Temporal covariates relating to the daily average of daytime weather conditions included the maximum, minimum and average temperature (Max, Min and Ave temp), maximum cloud cover estimated in field (Max cloud), average humidity (Ave humid), precipitation on the day (PPT), maximum wind speed (Max wind) and rainfall events on the day (Events). The number of hours (Hours) spent on each occasion surveying is indicated. ~1 is used to indicate a constant model. For each model, the number of parameters specified (npar), log-likelihood estimates (LogLik), difference in AIC between each model and the best model (dAICc), as well as the Akaike weight (AICcwt) are indicated. Best models within 2 AICc are highlighted. Point estimates for activity centre size (τ) and population density (D) are indicated but were assumed constant within each model.

(a) RNR1

Model	npar	logLik	dAICc	AICcwt	$\tau\ (m^2)$	$D\ (10^{-4} ind.m^{-1})$
g0~b	3	-412.837	0	0.403	12.77	2.52
g0~Events	5	-410.459	0.127	0.3782	12.80	2.82
g0~SVL	3	-413.448	1.221	0.2188	12.81	2.82
g0~Maxwind	3	-459.414	93.154	0	14.31	2.56
g0~Avetemp	3	-459.618	93.562	0	14.31	2.56
g0~Avehumid	3	-459.959	94.244	0	14.31	2.56
g0~Maxtemp	3	-460.644	95.614	0	14.31	2.56
g0~PPT	3	-461.086	96.498	0	14.31	2.56
g0~Hours	3	-461.235	96.794	0	14.31	2.56
g0~1	2	-463.374	98.795	0	14.31	2.56
g0~Mintemp	3	-462.376	99.078	0	14.31	2.56
g0~Maxcloud	3	-462.801	99.927	0	14.31	2.56

(b) RNR2

	npar	logLik	dAICc	AICcwt	τ (m^2)	D $(10^{-4} ind.m^{-1})$
g0~Min.temp	3	-433.99	0	0.4775	116.58	1.29
g0~Max.temp	3	-435.752	3.525	0.0819	116.53	1.29
g0~Ave.temp	3	-435.832	3.685	0.0756	116.50	1.29
g0~Max.cloud	3	-435.886	3.793	0.0717	116.50	1.16
g0~1	2	-437.069	3.858	0.0694	116.50	1.29
g0~SEX	3	-436.319	4.658	0.0465	116.49	1.36
g0~Hours	3	-436.561	5.142	0.0365	116.50	1.29
g0~Max.wind	3	-436.706	5.433	0.0316	116.50	1.29
g0~b	3	-436.924	5.868	0.0254	117.20	1.06
g0~PPT	3	-436.933	5.886	0.0252	116.81	1.29
g0~Ave.Humid	3	-437.069	6.158	0.022	116.50	1.29
g0~SVL	3	-437.069	6.158	0.022	116.50	1.29
g0~Events	5	-434.985	6.943	0.0148	116.50	1.29

Demographic estimates of population densities were comparable to that from the sib-ship analysis but with some discrepancies (Table 4.3). Sib-ship estimate of D_E for RNR1 ($3.03 \times 10^{-4}\ indiv.m^{-2}$) was almost twice the demographic estimate ($1.69 \times 10^{-4}\ indiv.m^{-2}$) though the error margins did overlap. The discrepancy for RNR2 was even larger, sib-ship estimate ($0.51 \times 10^{-4}\ indiv.m^{-2}$) was ~2.5 times smaller than demographic estimate ($1.29 \times 10^{-4}\ indiv.m^{-2}$), in this case however, the SECR estimate fell within the 95% CI for the sib-ship estimate.

3.1.1   DISPERSAL DISTANCE ESTIMATES

Estimates of annual dispersal distances (σ) from NS and D_E were largely similar between methods of estimating D_E (Table 4.3). Only the estimates of D_E changed between the two methods so differences between σ estimates are best explained by the comparisons of D_E (above). Overall, σ estimates were comparable within a 100m when using genetic D_E and lower 95% CI bounds for all local-scale sites were ~80-200m but upper bounds were ∞. Large error bars (including ∞) indicates great inaccuracy in all but the broad-scale estimates. Estimates across broad-scale, across the whole of RNR, were more precise though error bounds remained very large (95% CI $132 - 1156\ m.year^{-1}$). The point estimate of $390 m.year^{-1}$ was similar to the two local RNR estimates. Application of individual-based regression methods are less biased with pairwise geographic distances (pgd) shorter than $\sim 20\sigma$ (Rousset 1997, 2000). Only for analysis at broad-scale was the max pgd greater than 20 times the lower 95% CI value for σ, (Table 4.2) indicating possible bias only at this scale.
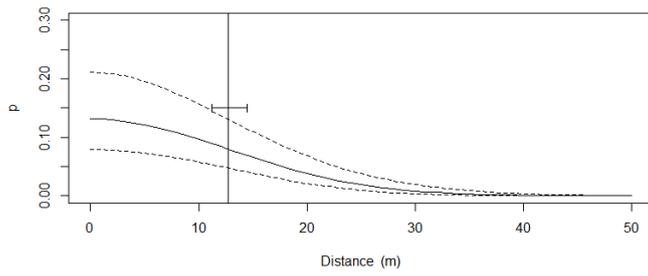
(a)
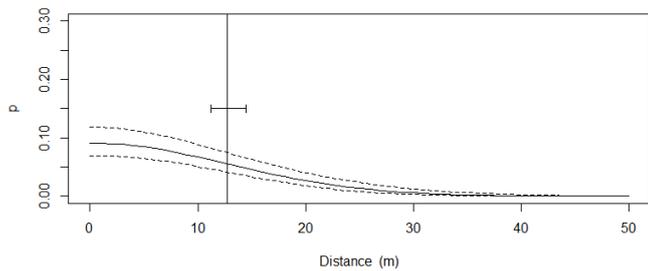
(b)

Figure 4.5: Fitted halfnormal detection function, g(d), for *Pedioplanis lineoocellata* with 95% confidence interval (CI) for RNR1 along increasing distances from the activity centre at re-capture occasion (a) t = 1 and (b) t > 1. Results based on model $g0{\sim}b$. Vertical line indicates the estimated τ, i.e. the standard deviation of the detection function, with 95% CI. τ and D assumed constant for the model.
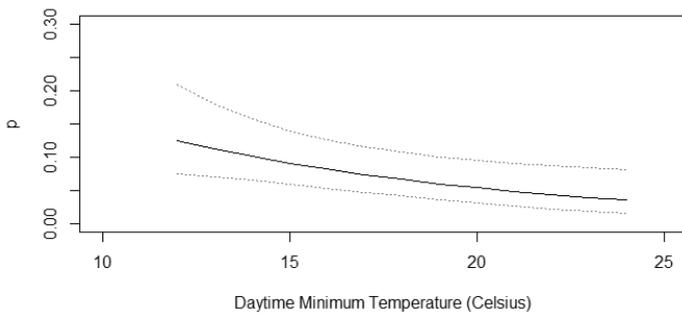
Figure 4.6: Detection probabilities at the activity centre, g0, for *Pedioplanis lineoocellata* with 95% confidence interval for RNR2 along increasing minimum daytime temperature based on best model, $g0{\sim}Min.temp.$ τ and D assumed constant for the model.

3.2 SPATIAL AUTOCORRELATION OF RELATEDNESS

Spatial autocorrelations were calculated to examine patterns of relatedness with distance between individuals. At the local-scale, within sampling sites (<2km pgd), there was no evidence from the multiple distance class (MDC) analysis for spatial structure across all lags sizes (10m-350m) for all three sites (RNR1, RNR2 and HART; $p>0.05$) and r-values were consistently low ($r\sim0$). The results suggest very high levels of gene flow at the local spatial scale (Figure 4.7) and therefore panmictic populations. In this case sites would need to be embedded within a wider distributed sample set to identify spatial structuring (Sokal et al. 1983). At the broad-scale (<15km pgd), the MDC analysis with lag sizes 20-14500m had r-values that initially were not different from 0 and the autocorrelations were not significant, likely due to small sample sizes per lag (results not shown). At lags greater than 300m the autocorrelation became significant and after the 5220m lag size the bootstrap test became non-significant (conservative estimate of the spatial extent of genetic structure- segs), followed by the permutation test after 10440m (less conservative estimate) (Figure 4.8). Correlograms were visualised at 4000m lags (Figure 4.9). The large segs estimates (~5.2-11km) but low r-values (<0.01 at 4000m lags) indicate a persistent genetic structure that spans large distances relative to the extent of sampling with a signal of low relatedness. The result tentatively supports a very gradual structure however, the lack of higher order oscillations between positive and negative r-coefficients indicates that the true extent of spatial structure exceeds the scale of sampling (Peakall et al. 2003).

At the regional scales (over 1600km pgd), there was clear evidence for spatial genetic structure (Figure 4.10) and the MDC analysis suggested that the spatial extent of sampling was adequate. The r-values were well above zero at lag sizes of 50-600km, after which they drop sharply and were no longer statistically significant. Based on the MDC analysis, the segs is 600-650km. Lags of 250km were used to visualise the correlogram to include the 600-650km range and to retain useful sample sizes per interval. The correlogram oscillated between positive and negative as was expected. The overall correlation was significant ($p<0.0001$) thus at the spatial scale over hundreds of kilometres there is clear spatial genetic structure supporting an IBD pattern for *P. lineoocellata*.
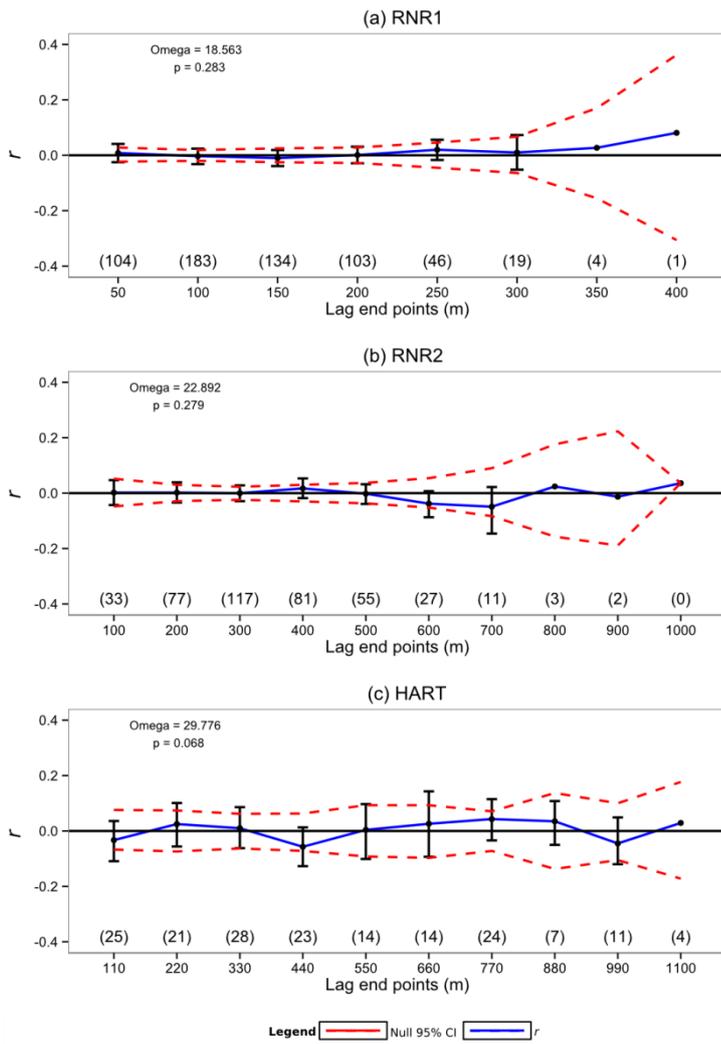
Figure 4.7: Local-scale spatial autocorrelation (SAC) of correlation coefficients, *r*, with increasing pairwise geographic distances (pgd) between individuals of *Pedioplanis lineoocellata* (<2km pgd). Sampling performed at (a) RNR1, (b) RNR2 and (c) HART in the Northern Cape, South Africa. SAC performed at specific lags. The 95% permutation confidence intervals (CI) about the null hypothesis of a random distribution and 95% bootstrap confidence error bars around *r*, are indicated. Overall significance indicated (Ω & p-value) for a one-tailed test for positive autocorrelation. Sample sizes at each lag (n) indicated above lag end point labels.
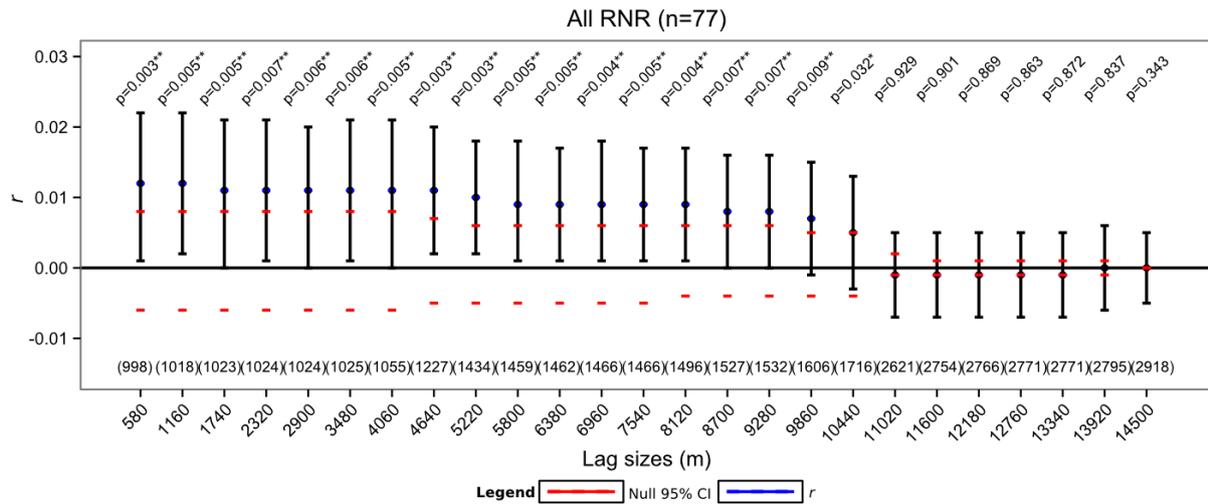
Figure 4.8: Broad-scale spatial autocorrelation for multiple distance class analysis for *Pedioplanis lineoocellata* across Rooipoort Nature Reserve (<15km pairwise geographic distances) at 580 - 14500m lags. The correlation coefficients, *r*, from the first lag for each lag size is shown. The 95% permutation confidence interval about the null hypothesis of a random distribution and 95% bootstrap confidence error bars around *r* are indicated. Sample sizes at the first lag (n) indicated above lag size labels. Overall significance indicated (p-value) at each lag size for a one-tailed test for positive autocorrelation.
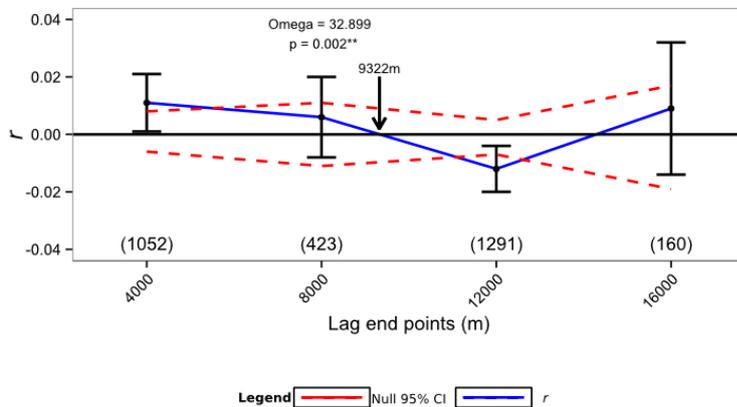
Figure 4.9: Broad-scale spatial autocorrelation coefficient, *r*, for *Pedioplanis lineoocellata* across Rooipoort Nature Reserve (pairwise geographic distances <15km) at 4000m lags. The 95% permutation confidence interval about the null hypothesis of a random distribution and 95% bootstrap confidence error bars around *r* are indicated. Overall significance (Ω & p-value) for a one-tailed test for positive autocorrelation indicated. Sample sizes at each lag (n) indicated above lag end point labels.
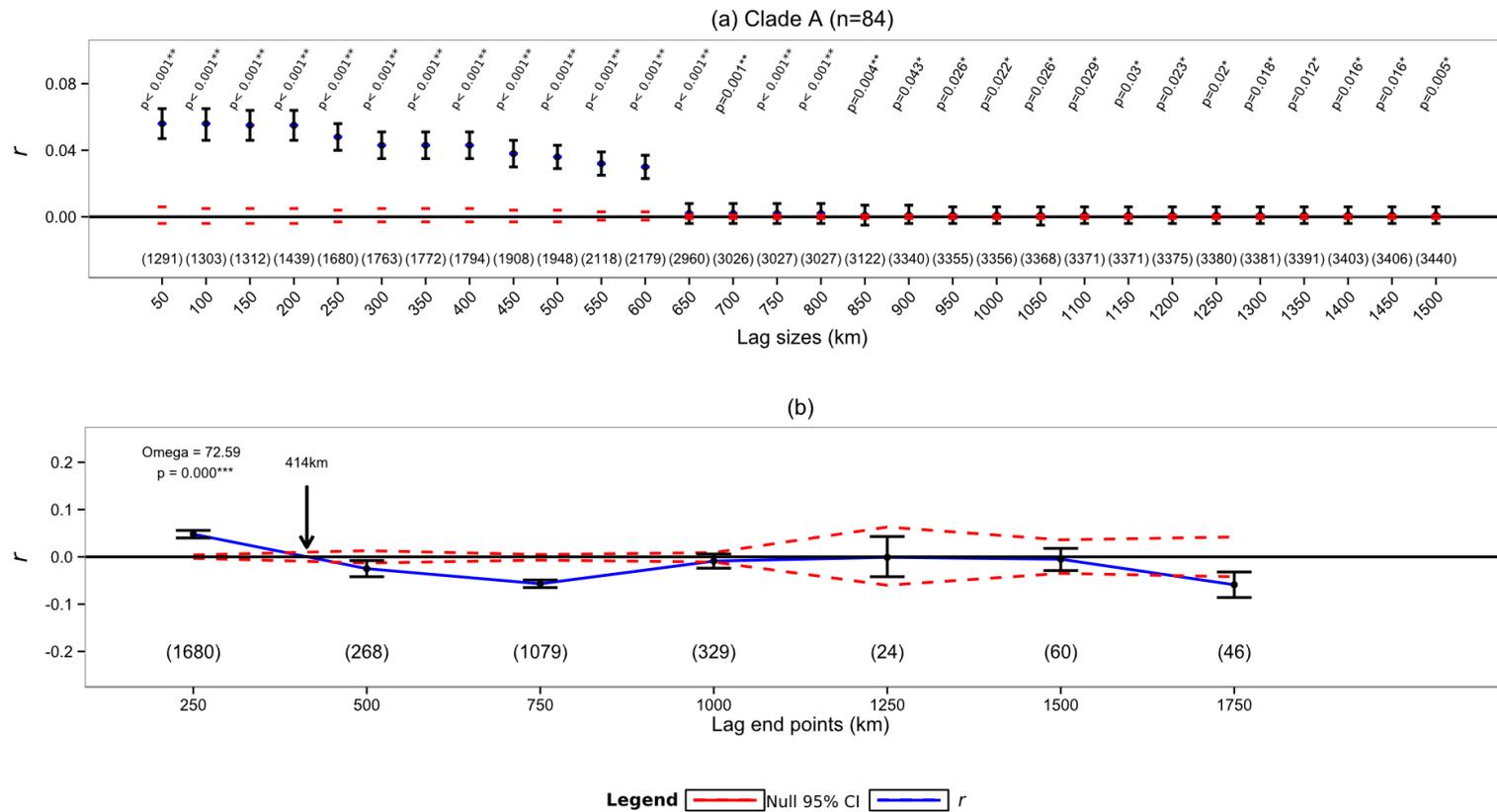
Figure 4.10: Regional-scale spatial autocorrelation of correlation coefficients, *r*, with increasing pairwise geographic distances (pgd) between individuals of *Pedioplanis lineoocellata* across the distribution of Clade A (>1600km pgd). The 95% permutation confidence intervals about the null hypothesis of a random distribution and 95% bootstrap confidence error bars around *r*, are indicated. Overall significance indicated (Ω & p-value) for a one-tailed test for positive autocorrelation. (a) Multiple distance class analysis across Clade A with increasing lag intervals. Sample sizes at each of the first lags (n) indicated above lag interval label (b) SAC visualised at 250km lag intervals. Sample sizes at each lag (n) indicated above lag end point labels.

4 DISCUSSION

4.1 DISPERSAL IN *PEDIOPLANIS LINEOOCELLATA*

Large annual dispersal distance estimates and statistically significant, but weak, spatial autocorrelation (SAC) among individual *Pedioplanis lineoocellata* across geographic scales covering several kilometres indicates high levels of dispersal and gene flow for the wide-spread southern African lacertid. Results suggest local-scale panmixia as a result of large dispersal distances and only at the broad geographic scale across Rooipoort Nature Reserve (RNR) was there evidence of positive genetic structure. Consequently, only at this scale or broader scales is it possible to estimate precise neighbourhood size values and annual dispersal distances for the species. Although tentative, the lower 95% confidence interval bounds for annual dispersal distances at broad-scale ($\sigma \sim 80m.\,year^{-1}$) were greater than the typical dispersal distances measured for other lacertids (12-60m) but not much greater than that predicted from home range estimates for *P. lineoocellata* (Wasiolka 2007)*,* assuming dispersal distance is a few (~3-5) home range diameters (i.e. ~50-140m; see Shields 1987; Clobert & Baguette 2012). The SAC result also provides support for an isolation-by-distance pattern at regional-scale, across mtDNA Clade A, which was not detected in Chapter 3 using the Mantel test. The evidence for strong dispersal provide support of gene flow as the potential cause for discordance between microsatellite clusters and mtDNA clades as found in Chapter 3.

The dispersal point estimate based on a significant Neighbourhood Size (NS) regression ($\sigma \sim 390m.\,year^{-1}$) far exceeded the usual demographic estimates for other similar sized lizards [e.g. Australian alpine skink, *Niveoscincus microlepidotus* ~12-14m, (Olsson et al. 2003); *Zootoca vivipara* , 20m (Clobert *et al.* 1994); *Lacerta agilis,* 25-57m (Olsson *et al.* 1996)]. However, larger NS-based estimations have been retrieved for other species, e.g. *Gehyra variegata* $1.6 - 6km.\,yr^{-1}$ (vs. an upper 95% CI 1156m for *P. lineoocellata*). *Gehyra variegata* and another Australian gecko, *Oedura reticulata,* also have population assignment based evidence for dispersal distances greater than 100m (Hoehn *et al.* 2007)*.*

Considering the evidence of daily movement found in this study (~30 and 170m MMDM) and $\sim 25 - 45m.\,day^{-1}$ using telemetry in another study (Wasiolka *et al.* 2009b), large dispersal estimates are not surprising. Large daily movement estimates could indicate that it is within the physiological capabilities of the species to move long distances during dispersal. Foraging mode is known to affect daily movement (Cooper et al. 1999; Verwaijen et al. 2008a, 2008c) and consequently larger home range sizes are expected for active foragers (Verwaijen et al. 2008b). While there is no review on the relationship of home range size and mean dispersal distance for lizards, in birds and mammals the two are linearly related (Bowman *et al.*

2002; Bowman 2003). However, this having been said, several Eurasian lacertids with smaller demographic dispersal estimates are more active foragers than *P. lineoocellata* with comparable home range sizes, these include *L. agilis, Z. vivipara,* and *Lacerta monticola* (Cooper et al. 1999; Verwaijen et al. 2008a, 2008b). The relationship of home range size and dispersal propensity in lizards requires further work.

Spatial autocorrelation at broad-scale produced a signal indicative of large neighbourhood sizes relative to the sampling scale, further supporting large dispersal distances (Sokal *et al.* 1989). Other species with strong dispersal propensity, such as the Cane toad, *Bufo marinus,* and the small-eyed elapid, *Rhinoplocephalus nigrescens,* also show no spatial genetic structure at local-scale sampling (Leblois *et al.* 2000; Keogh *et al.* 2007). The lower estimates of the SAC correlation coefficients (~0.05-0.15) for other studies at smaller geographic scales are comparable to the estimates for *P. lineoocellata* (Smouse et al. 1999; Peakall et al. 2003; Chapple et al. 2005; Hoehn et al. 2007; Row et al. 2010; Duckett et al. 2013). While comparisons of correlation coefficients are not easily made between studies because of the biological differences between species and the different geographic coverage, low correlations values are a general expectation under mutation-drift models (Sumner *et al.* 2001).

Results match fairly well the low, but significant, genetic differentiation found between RNR populations and evidence for between-site migrants (Chapter 3). However, as BayesAss only estimates migration within the past three generations (Wilson et al. 2003) I would expected that the actual recent gene flow between populations, which SAC measures, is underestimated by BayesAss. Dispersal across RNR may be truncated by unfavourable habitat and this restriction should generate small-scale panmictic populations but IBD over larger geographic scales (e.g. Clobert et al. 1994; Hoehn et al. 2007; Wang et al. 2009; Dubey et al. 2011; Nathan et al. 2012). While flexibility in the species' foraging behaviour, and thus the habitat occupied, could facilitate movement across unfavourable areas, the somatic costs of occupying or moving through sub-optimum habitat should reduce fitness. Fitness costs through increased risk of predation, reduced foraging or basking time and the additional energy expenditure of traversing unsuitable terrain should reduce the frequency of dispersal over very large distances and limit gene flow. This was found for two Australian arboreal geckos, *G. variegata* and *O. reticulata*, where local-scale panmixia was identified in continuous and favourable habitats, whilst in fragmented habitats, populations had higher genetic differentiation over a similar geographic scale (Hoehn *et al.* 2007).

Results presented here are the first dispersal estimates for *Pedioplanis lineoocellata*. With strong evidence for dispersal across the broad-scale, it should follow that across the regional-scale -the distribution of Clade A- there would be a fair amount of gene flow. The estimated spatial extent of positive genetic structure (segs) was large relative to the spatial extent of sampling but the correlation coefficients were low,

indicating that although extensive, the autocorrelation is weak. This might be expected from either extensive gene flow or strongly restricted gene flow (Sokal et al. 1983; Smouse et al. 1999). The former seems more likely considering the evidence of gene flow within RNR. When neighbourhood sizes are large, the size of genetically homogenous regions increases but the amplitude of the correlations decrease (Sokal *et al.* 1989). In contrast, no IBD pattern was detected across the distribution of Clade A using the Mantel test in Chapter 3. An IBD signal may not be detected when sampling is at small spatial scales with panmictic populations or when sampling is at large spatial scales where divergence is regulated more strongly by mutation and genetic drift than it is by gene flow. It is highly unlikely that the entire clade A would be panmictic and this is supported by the positive genetic structure found in this chapter. The possibility that the sampled area was too large also seems unlikely, as the DAPC analysis in Chapter 3 indicated that many of the identified clusters had very wide-spread distributions covering large parts of the species range. Based on these results, I would propose that the interplay of historic spatial patterns in nuclear DNA clusters and contemporaneous barriers to gene flow are responsible for the weak signal. Similar large segs estimates have been reported for the Australian gecko, *Gehyra variegata,* which had segs of 900m- over 50km across different habitat types and for *Oedura reticulate,* which had segs of ~500m, with corresponding large dispersal estimates (Hoehn *et al.* 2007; Duckett *et al.* 2013).

In Chapter 3 I proposed that the discordance between mtDNA clade boundaries and microsatellite clusters was the result of gene flow after Plio-Pleistocene range contraction and expansion events. These first estimates of dispersal for *P. lineoocellata* indicate high levels of dispersal and gene flow across local and broader scales as does the evidence of spatial autocorrelation. The results support the possibility of great gene flow across mtDNA clade boundaries in the recent past.

4.2 METHODOLOGICAL CONSIDERATIONS

Estimates of σ from Wright's NS are comparable to direct estimates of dispersal (Rousset 2000, 2004; Sumner *et al.* 2001; Broquet *et al.* 2006; Watts *et al.* 2007), though estimates may be of two-fold difference due to biases introduced by demographic methods and the assumptions needed for mathematical models to be developed (Rousset 2000; Sumner *et al.* 2001). Direct measures were not available for *P. lineoocellata* and thus comparisons relied on the use of effective population size data. Validating the accuracy of the genetic method, however, will require estimates of ecological dispersal for comparison.

Comparisons of D_E between sampling sites were not possible because site was confounded by season, age structure, and geographic separation, and even comparisons between methods for a single site are tentative because NS regressions were not significant at local-scale. Although the absence of an IBD pattern is in itself

an indication of large dispersal, precise and accurate estimates of σ require statistically significant regressions of $\hat{e}$ and ln(geographic distance). Sampling across RNR may have further biases as it was not continuous, which violated an assumption of the NS regression method and the maximum pgd was greater than 20σ (Rousset 1997, 2000). The NS estimates should be obtained from sampling across pair-wise geographic distances (pgd) ~400m-7.8km or 1600ha area in a continuous population (Rousset 2000). This would exclude pgd below σ and above 20σ.

There are several other possible sources of error in estimating σ that may have widened error margins. The differences between σ based on demographic and genetic D_E estimates were of similar magnitude to that of other studies however, only the measure of D_E changed between estimation methods thus direct comparisons of D_E between methods is necessary. Although D_E estimates were within a single order of magnitude of each other, there was notable variation between sites and for each site between methods. Differences would relate to data collection and the different assumption of each of the analyses. Foremost, between site comparisons are completely confounded thus discrepancies between sites may simply be for true biological reasons.

The sib-ship D_E estimates were obtained indirectly by estimating N_E from genetic data. The analysis assumes samples are from a single cohort of a species with discrete generations (Wang 2009). For the RNR1 sample this should be a minor concern as most of the individuals sampled were juveniles, and probably hatched earlier that year. The age of RNR2 individuals could not be verified because no information is available on the relationship of SVL and age for this species. As a result, the population size estimated for RNR2 might have been 'intermediate' between the actual N_E values for all the component age classes, likewise for HART (Wang 2009). Provided individual survival is independent to that of their kin, the use of multiple life stages should not alter the estimated results but may still compromise the power of the analysis (Wang 2009). There have been no studies on family structure in *P. lineoocellata* but based on the general absence among lizards (Gardner *et al.* 2001; Somma 2003), the bias seems unlikely.

Large null allele frequencies may seriously limit parentage analyses (Dakin et al. 2004) and to a lesser extent, sib-ship analyses. All three local-scale sites had support for the presence of null alleles based on the INEst models in Chapter 2 but the point estimates for null allele frequencies at all loci were at or below the 0.2 threshold proposed by Dakin & Avise (2004) for retaining loci. Below this threshold, null alleles cause an increase in false exclusion of possible parents (or sibs) without warranting great concern (Dakin et al. 2004). This should not be as influential for sib-ship analysis as for parentage analysis because fewer corresponding alleles are required to estimate sibling relationship.

Population density could not be estimated for the whole of RNR from the sib-ship analysis because sampling across the reserve was highly localised (i.e. not entirely random), which is a violation of a sib-ship analysis assumption and risks underestimating the population size. Sib dyads may be more frequent in a non-random sample compared to the actual frequency of the population (Wang 2009). Therefore, the average density from RNR1 and RNR2 was used instead but this required an assumption of constant population density across the reserve and that average density serves as a good approximate for the entire reserve. Both assumptions are not likely to be true and better estimates will require random but geographically continuous sampling.

Inaccuracies in the estimates may have also been introduced from the demographic methods used to estimate D_E. Population density estimation through spatially explicit capture recapture (SECR) are generally robust and reliable provided the data are appropriate (Efford et al. 2004) but there are still several possible means by which inaccuracy may have been generated. Area searches were used for surveys and survey effort varied across the area, which may have introduced bias. Moreover, because of differences in encounter frequencies, the size of the surveyed areas between sites differed and consequently the amount of survey effort per m^2 differed. I believe that the different age composition and season of sampling for each RNR data set was the cause of the different encounter frequencies because the relative size- and thus home range- and daily foraging behaviour of individuals from either of the two data sets should differ as well. One possible cause for the necessity of the adjustment in sample area size at RNR2 was that the mean maximum daily movement (MMDM) of ~170m for adult lizards was unexpectedly large when compared to the $19-45m.day^{-1}$ estimates of Wasiolka (2007). The latter estimates are more comparable to that estimated from juveniles at RNR1 (~30m). Furthermore, Wasiolka (2007) used radio tracking to measure movement over 24hrs whereas the estimates of MMDM were simply the average distances between recapture points, therefore MMDM should underestimate daily movement. The most appropriate study design for future studies should have a geographic extent of ~100ha to account for possible large MMDM and additionally, should include other sites within 500m of each other monitored simultaneously. This may allow one to directly estimate physical dispersal between sites.

The precision of D_E estimates from SECR depends on the data. *Pedioplanis lineoocellata* is a difficult species to recapture as evident from low recapture rates, similar to another elusive, open-habitat lizard, *Phrynosoma mcalli* (Royle et al. 2008). Low recapture rates in conjunction with small sample sizes for a CMR study will reduce the precision for parameter estimates reliant on multiple captures such as distance attribute to the g(d) function and the estimates of MMDM. Temporary emigration and immigration at a study site can affect parameter estimates (Kendall 2001; Royle et al. 2008; Borchers et al. 2008) therefore I attempted to reduce the effects by using the SECR analysis for demographic D_E estimates and by including

an MMDM buffer on sampling area in estimating the sib-ship D_E. However, the effectiveness of both methods will be reliant on the available data. More generally, there were several possible SECR models that could not be examined with the sparse data as it would result in model over-fitting. This was particularly the case for τ, the activity centre for which no covariates could be used. Sample sizes and recapture rates could be increased by employing permanent markings such as toe clippings, branding or photographic identification and extending the study over a longer period of time. This could allow for the inclusion of additional SECR models, more precise estimates of movement and activity centres and further facilitate direct estimates of dispersal.

Adult population density for RNR1 could not be measured directly, instead D_E was inferred from a measure of juvenile D adjusted according to adult: juvenile ratio from another study, which may further contribute to discrepancies. The population sizes estimated from the study by Wasiolka (2007) would be subject to their own methodological errors and biases. Furthermore, the site was in a different region from HART and RNR, thus geographic variation in population structure may also affect results. Geographic variation in life history traits is well documented (Niewiarowski 1994) therefore, it is not wise to assume consistency between populations of a species from different areas (Niewiarowski 1994; Du *et al.* 2005; Díaz *et al.* 2007). Future studies should consider simultaneous and contemporary estimates of all parameters needed to perform such density estimation (e.g. Sumner et al. 2001; Watts et al. 2007).

5   CONCLUSION

The assessment of dispersal in this study for *Pedioplanis lineoocellata* may be the first for any southern African lacertid species. The SAC results met predictions of IBD at broad- and regional-scale but the annual dispersal distances were larger than predicted, although in agreement with the SAC results. Spatial autocorrelation and dispersal estimates indicate high rates of gene flow and dispersal across large geographic scales, though interpretation was hindered by the absence of site-specific and between-site comparisons of demographic information. Direct estimates of dispersal are required to ground truth the results presented here. The use of other forms of analyses such as least-cost-path analysis and more geographically extensive CMR studies will further contribute to a clearer understanding. These two individual-based methods applied in this chapter offer a powerful means of investigating within- and between- population processes by providing insight into the possible dynamics that may have generated large scale genetic patterns. An understanding of dispersal and barriers to dispersal is particularly important for interpreting genetic patterns and inferring changes in gene flow between populations either on a contemporary time scale or historically. The discordance between microsatellite clusters and mtDNA clades might well be explained by extensive

dispersal, for which there is now some support. Scenarios of population expansions, contractions and gene flow between regions could be developed based on this contemporary evidence for dispersal tenacity. Dispersal estimates and other biological parameters such as thermal or desiccations tolerance, can be used for predicting the response of a species to environmental scenarios, for example, through inclusion in predictive distribution models (Midgley et al. 2006; Engler et al. 2009). *Pedioplanis lineoocellata* could be minimally affected by predicted climate change as a consequence of it dispersal ability. Palaeo-distributions may also be modelled by following the procedure in reverse by considering a contraction rate that incorporates the magnitude of dispersal. This approach may help identify refugial areas that were of importance during glacial periods (e.g. Barlow et al. 2013) and predict the species resiliency to contemporary threats of habitat fragmentation and degradation (Hoehn et al. 2007; Tolley et al. 2010b; Duckett et al. 2013).

Chapter 5 General Conclusion

Demographic and genetic connections between populations are immensely important in ecology yet they are unexamined for many taxa. For this thesis I described genetic spatial patterns, gene flow and estimated dispersal distances for the wide-spread southern African endemic lacertid lizard, *Pedioplanis lineoocellata*. Nine of the first microsatellite primer pairs for the species were tested and optimised to investigate nuclear genetic patterns across the range of two of the most wide-spread mtDNA clades and to estimate dispersal distances. These estimates fill a gap in the available information for the species and are the first dispersal estimates of a southern African lacertid lizard. I also incorporated capture-mark-recapture data as an independent dataset in estimating dispersal distances. From this research, I provided new information on the dispersal and gene flow biology of *P. lineoocellata* and highlight how much has yet to be elucidated- particularly with regards to the mtDNA and microsatellite discordance; information essential for understanding the connectivity between populations both current and historic.

1   A SOUTHERN AFRICAN LACERTID ON THE MOVE

Pedioplanis lineoocellata was found to have peculiarly high dispersal estimates for a lacertid species and this is reflected in the extensive but weak isolation-by-distance (IBD) pattern across the distribution of Clade A. The disruption to the IBD pattern could be the result of complex genetic spatial patterns resulting from interactions between geographic features, climatic change and historic and contemporary gene flow. Although no well delimited boundaries between microsatellite clusters were recovered when compared to the mtDNA clades, this type of discrepancy has been seen in other taxa (Schulte *et al.* 2012; Visser 2013) and is most likely due to the differences in temporal scale at which the two sets of markers best retain resolution. Lastly, there was no evidence of fairly recent genetic admixture between mtDNA lineages. Furthermore, the putative hybrid population at the Loeriesfontein area, Farm Hartebeeslaagte, may possibly have an old connection to the Pylkop population.

While the results from this thesis are interesting, they do also lead to new questions. From Chapter 3, it appears as though contact regions between mtDNA clades may be more extensive than expected. A better sampled fine-scale examination of the distribution of mtDNA lineages simultaneously with an examination of microsatellite clusters across the distribution will better clarify whether or not the lineages occur sympatrically and to what extent and where there is evidence of hybridization. An important step in addressing the disjunction between large scale patterns of genetic structure and, local-scale patterns and dispersal measures is to identify the mechanisms that likely affect successful dispersal. An examination of dispersal between populations within the range of a few tens of

kilometres through least-cost-path analysis by including environmental data could help identify barriers to gene flow. The complex arrangement of microsatellite clusters can be particularly useful for this purpose as it allows one to target specific areas where strong dispersal between some populations but not between others is expected. For example, the west coast of South Africa, between Garies and HART and the area further south, GANS and surrounding populations, may be useful for this purpose. How the results for Clade A and Clade B relate to the other two mtDNA clades, Clade C and Clade D also awaits better sampling. Similarly, because there were no populations within Clade B that had been thoroughly sampled, comparisons across clades at a population level could not be done. This is work that should be addressed in future investigations.

Lastly, anthropogenic climate change is predicted to have significant impacts on southern African climate including increased aridification, drastic changes in vegetation composition and weather predictability (Boko *et al.* 2007). Southern Africa has the third most biologically rich lizard fauna worldwide (Branch 1998; Bates *et al.* 2014) in part because of an arid adapted assemblage of lizards with a demonstrated evolutionary history strongly influenced by historic climate change (Lamb et al. 2003; Swart 2006; Tolley et al. 2006; Makokha et al. 2007; Swart et al. 2009; Edwards 2013; Edwards et al. 2013). Historic disruptions have been significant influences in the evolutionary history for many species and similar disruptions are predicted for future anthropogenic climate change (Araújo et al. 2006; Boko et al. 2007; Deutsch et al. 2008; Sinervo et al. 2010). Moreover, habitat loss and fragmentation caused by urban development, farming practices and invasion of alien species have added to the mounting pressure on indigenous biota. Organisms may respond by adaptation, behavioural plasticity or by dispersal to new environments; failing the above, populations are predicted to decline and go extinct. The ability of any species to disperse at pace with a spatially shifting climatic niche will determine the likelihood of its persistence. By modelling species range shifts according to shifts in suitable habitat, species dispersal capabilities and, behavioural and physiological responses, we can predict their persistence potential. The dispersal estimates generated in this thesis, are necessary data for performing such predictive modelling.

Reciprocally, we can attempt to reconstruct palaeo-distributions using climate reconstructions in an attempt to determine refugial areas as well as test hypotheses regarding the formation of the identified mtDNA lineages and their secondary contact zone (e.g. Barlow et al. 2013). The identification of refugia that correspond geographically with areas of low climate velocity (Tolley *et al.* 2014a) may be areas of noteworthy conservation value if shared by several taxa. For flora in the Cape Floristic Region, climatic stability has been important for conserving species and preventing extinctions during Plio-Pleistocene climate change (Tolley *et al.* 2009, 2014a; Dupont *et al.* 2011; Pio

et al. 2014). Bio-climatic models using anthropogenic climate change scenarios indicate that assuming similar stability in the future may be an over-simplification because of species-specific responses and community level changes (Pearson et al. 2003; Chefaoui et al. 2005; Tolley et al. 2009). However, refugial areas offer insights into what environmental attributes facilitate environmental stability and the persistence of populations during climate change events. This can be useful for predicting which present day environments may act as refugia for species under predicted climate change.

This study also prompted several new questions. Savanna and grassland biomes burn frequently (c.f. every 5-30 years; Govender et al. 2006) and fire has been demonstrated to be important for dispersal ecology for other lizard species (Chapple et al. 2005; Ujvari et al. 2008; Schrey et al. 2011; Templeton et al. 2011). With an extensive burn event at RNR in 2012, a follow up study may similarly provide insight into the effects on population bottlenecks, re-colonisation and dispersal following fire events. Secondly, during the sample collection in 2012, I observed extreme age-skews between sample periods (as discussed in Chapter 4). There appears to be seasonal variation in population structure. Similar observations have been reported for *Pedioplanis lineoocellata* (Wasiolka 2007) and for another species (Goldberg 2008), *Meroles squamulosus*. Moreover at RNR, *Meroles squamulosus*, was found active only subsequent to the hatching of juvenile *P. lineoocellata*- the time when adult *P. lineoocellata* were least frequently encountered. There may be some form of temporal niche partitioning between species in addition to temporal partitioning of activities within species. This has been proposed for other lacertids, including between *Ichnotropis capensis* and *M. squamulosus* (Jacobsen 1987; Mayer et al. 1999; Goldberg 2008). This may be one of the mechanisms that maintain high herpetofaunal diversity in southern Africa.

2   NOTES ON DATA COLLECTION AND ANALYSES

While many of the conclusions in this thesis may hold, it is important to remember that several possible sources of inaccuracy could not be accounted for with the present dataset. This was mostly due to low sample sizes, the presence of null alleles and the lack of basic information about demographic parameters for the species at the sampled populations. Parameters such as sex ratios, population densities and dispersal were only partially or indirectly estimated and often without account for geographic or temporal variation. Though these issues do not seriously detract from the conclusions drawn, the estimation of such parameters should be considered in future project design.

Several taxonomic groups, including lacertid lizards, have been found to have an 'unusual' prevalence of null alleles (Dakin et al. 2004; Chapuis et al. 2007; Lemer et al. 2011; Tolley et al. 2014b) and analyses are routinely performed despite deviations from HWE. A more effective way of accounting for null alleles is needed because the common practise of replacing affected loci or ignoring possible side-effects is detrimental to research on non-charismatic and non-model organisms where available loci are often limited. In this study the original nine microsatellite markers were retained because of the time and financial costs involved in developing and testing new primer pairs. To mitigate the effects of null alleles, simultaneous estimation of null alleles and inbreeding coefficients was used to account for possible inflation of inbreeding coefficients. Furthermore, analyses that made assumptions about HWE or mutation models were avoided. Unfortunately several analyses such as sib-ship assignments, migration estimates and Bayesian population assignment, did not include means of accounting for null alleles. Only few population genetic statistics adequately account for influence of null alleles despite the concurrent effects from null alleles, inbreeding, population bottlenecks, selection, migration and genotyping errors (Chybicki et al. 2009). More work is required to develop means of including null allele frequencies in estimation of other important parameters. At the least, simulation studies should include predictions of how estimators will behave when influenced by null alleles or genotyping errors (e.g. Cornuet & Luikart 1996; Luikart et al. 1998b; Piry et al. 1999; Dakin & Avise 2004).

Individual based methods, such as spatial autocorrelation and estimating Neighbourhood Size from a genetic regression are methods largely overlooked in many animal studies (Rousset 2000; Sumner *et al.* 2001; Peakall *et al.* 2003) but have been shown to produce accurate results for several species. In this study I have demonstrated how such analyses may be performed when background information for the species is patchy and I pointed out work that needs to be done before attempting such analyses on other taxa. Specifically, the estimation of demographic parameters for population density and dispersal, as well as associated parameters like sex ratios, age-structure and reproductive variance are needed. Such information can be attained during DNA collection or CMR data collection.

Compared to other southern African taxa such as indigenous fish, insects and terrestrial molluscs, a fair amount of work has been done on lacertid lizards and other squamate reptiles. The representation is not comparable to the work that has been conducted on Eurasian taxa and many researchers have acknowledged this and attempts have been made to redress the issue (Branch 1999; Branch *et al.* 2006; Wasiolka 2007; Barlow *et al.* 2013). Because of the amount of research that has been conducted on *Pedioplanis lineoocellata* this species may quickly become a better understood representative for southern African lacertids and lizards in general. Well-studied species can be useful for performing

investigations with novel methods or testing more complex hypotheses that require extensive 'background' information.

REFERENCES

ALBERTO, F. 2009. MsatAllele_1.0: An R package to visualize the binning of microsatellite alleles. *The Journal of Heredity*. 100: 394–397.

ALJANABI, S. M. & I. MARTINEZ. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*. 25: 4692–4693.

ANDREASSEN, H. P., N. C. STENSETH & R. A. IMS. 2002. Dispersal Behaviour and Population Dynmics of Vertebrates. Pp. 237–256. In: J. M. Bullock, R. E. Kenward & R. S. Hails (Eds.), Dispersal ecology : the 42nd symposium of the British Ecological Society held at the University of Reading, 2-5 April 2001. Blackwell Publishing Ltd, UK.

APLET, G. H. 2011. Alteration of earthworm community biomass by the alien *Myrica faya* in Hawai'i. *Internation Association of Ecology*. 82: 414–416.

ARAÚJO, M. B., W. THUILLER & R. G. PEARSON. 2006. Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography*. 33: 1712–1728.

ARCTANDER, P., C. JOHANSEN & M. A COUTELLEC-VRETO. 1999. Phylogeography of three closely related African bovids (tribe Alcelaphini). *Molecular biology and evolution*. 16: 1724–1739.

AVISE, J. C. 1989. A role for molecular genetics in the recognition and conservation of endangered species. *Trends in Ecology & Evolution*. 4: 279–281.

BARLOW, A., K. BAKER, C. R. HENDRY, L. PEPPIN, T. PHELPS, K. A. TOLLEY, C. E. WÜSTER & W. WÜSTER. 2013. Phylogeography of the widespread African puff adder (*Bitis arietans*) reveals multiple Pleistocene refugia in southern Africa. *Molecular ecology*. 22: 1134–1157.

BATES, M. F., W. R. BRANCH, A. M. BAUER, M. BURGER, J. MARAIS, G. J. ALEXANDER & M. S. DE VILLIERS (Eds). 2014. Atlas and Red List of the Reptiles of South Africa, Lesotho and Swaziland. Suricata 1. P. 487. South African National Biodiversity Institute, Pretoria, South Africa.

BAUWENS, D. & RAMÓN DIAZ-URÍARTE. 1997. Covariation of life-history traits in lacertid lizards : a comparative study. *The American Naturalist*. 149: 91–111.

BEERLI, P. & J. FELSENSTEIN. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*. 152: 763–773.

BEERLI, P. & M. PALCZEWSKI. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics*. 185: 313–326.

BENNETT, A. F., R. B. HUEY & H. JOHN-ALDER. 1984. Physiological correlates of natural activity and locomotor capacity in two species of lacertid lizards. *Journal of Comparative Physiology B*. 154: 113–118.

BERTRAM, B. P. & H. G. COGGER. 1971. A noosing gun for live captures of small lizards. *Copeia*. 1971: 371–373.

BEZUIDENHOUT, H. 2009. The classification, mapping and description of the vegetation of the Rooipoort Nature Reserve, Northern Cape, South Africa. *Koedoe*. 51: 69–79.

BJÖRKLUND, M. 2005. A method for adjusting allele frequencies in the case of microsatellite allele drop-out. *Molecular Ecology Notes*. 5: 676–679.

BLOOR, P., I. H.-B. DE LAGUNA & S. J. KEMP. 2006. Highly polymorphic tetranucleotide microsatellite loci for the eastern Canary Island lizard, *Gallotia atlantica*. *Molecular Ecology Notes*. 6: 737–739.

BLOOR, P., V. RODRÍGUEZ, B. TERRASA, R. P. BROWN, V. PÉREZ-MELLADO, J. A. CASTRO, A. PICORNELL & M. M. RAMON. 2010. Polymorphic microsatellite loci for the Balearic Island Lizard *Podarcis lilfordi* (Squamata: Lacertidae). *Conservation Genetics Resources*. 3: 323–325.

BLUMROEDER, J., J. A. ECCARD & N. BLAUM. 2012. Behavioural flexibility in foraging mode of the spotted sand lizard (*Pedioplanis l. lineoocellata*) seems to buffer negative impacts of savanna degradation. *Journal of Arid Environments*. 77: 149–152.

BÖHME, M. U. & J. MORAVEC. 2011. Conservation genetics of *Lacerta viridis* populations in the Czech Republic (Reptilia: Lacertidae). *Acta Societatis Zoologicae Bohemicae*. 75: 7–21.

BOHONAK, A. J. 1999. Dispersal, gene flow, and population structure. *Quarterly Review of Biology*. 74: 21–45.

BOKO, M., I. NIANG, A. NYONG, C. VOGEL, A. GITHEKO, M. MEDANY, B. OSMAN-ELASHA, R. TABO & P. YANDA. 2007. Africa. Pp. 433–467. In: M. Parry, O. Canziani, J. Palutikof, P. Van der Linden & C. Hanson (Eds.), Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

BORCHERS, D. L. & M. G. EFFORD. 2008. Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics*. 64: 377–385.

BOSSART, J. L. & D. P. PROWELL. 1998. Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends in Ecology & Evolution*. 13: 202–206.

BOWIE, R. C. K., J. FJELDSÅ, S. J. HACKETT, J. M. BATES & T. M. CROWE. 2006. Coalescent models reveal the relative roles of ancestral polymorphism, vicariance, and dispersal in shaping phylogeographical structure of an

African montane forest robin. *Molecular phylogenetics and evolution*. 38: 171–188.

BOWMAN, J. 2003. Is dispersal distance of birds proportional to territory size? *Canadian Journal of Zoology*. 81: 195–202.

BOWMAN, J., J. A. G. JAEGER & L. FAHRIG. 2002. Dispersal Distance of Mammals Is Proportional To Home Range Size. *Ecology*. 83: 2049–2055.

BRANCH, W. R. 1998. Field Guide to the Snakes and Other Reptiles of Southern Africa. P. 399. Struik Publishers, Cape Town, South Africa.

BRANCH, W. R. 1999. Reptile Systematic Studies in Southern Africa: a Brief History and Overview. *Transactions of the Royal Society of South Africa*. 54: 137–156.

BRANCH, W. R., K. A. TOLLEY, M. CUNNINGHAM, A. M. BAUER, G. ALEXANDER, J. A. HARRISON, A. A. TURNER & M. F. BATES (Eds). 2006. A plan for phylogenetic studies of southern African reptiles: proceedings of a workshoop held at Kirstenbosch, February 2006. Biodiversity Series 5. P. 48. South African National Biodiversity Institute, Pretoria, South Africa.

BROQUET, T., C. A JOHNSON, E. PETIT, I. THOMPSON, F. BUREL & J. M. FRYXELL. 2006. Dispersal and genetic structure in the American marten, *Martes americana*. *Molecular ecology*. 15: 1689–1697.

BULLOCK, J. M., R. KENWARD & R. HAILS (Eds). 2002. Dispersal ecology: the 42nd symposium of the British Ecological Society held at the

University of Reading, 2-5 April 2001. P. 458. Blackwell Publishing Ltd, MA, USA.

CAMPAGNE, P., P. E. SMOUSE, G. VAROUCHAS, J.-F. SILVAIN & B. LERU. 2012. Comparing the van Oosterhout and Chybicki-Burczyk methods of estimating null allele frequencies for inbred populations. *Molecular Ecology Resources*. 12: 975–982.

CHAPPLE, D. G. & J. S. KEOGH. 2005. Complex mating system and dispersal patterns in a social lizard, *Egernia whitii*. *Molecular ecology*. 14: 1215–1227.

CHAPUIS, M.-P. & A. ESTOUP. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular biology and evolution*. 24: 621–31.

CHEFAOUI, R. M., J. HORTAL & J. M. LOBO. 2005. Potential distribution modelling, niche characterization and conservation status assessment using GIS tools: a case study of Iberian *Copris* species. *Biological Conservation*. 122: 327–338.

CHO, S., A. MITCHELL, J. C. REGIER, C. MITTER, R. W. POOLE, T. P. FRIEDLANDER & S. ZHAO. 1995. A highly conserved nuclear gene for low-level phylogenetics: elongation factor-1 alpha recovers morphology-based tree for heliothine moths. *Molecular biology and evolution*. 12: 650–656.

CHYBICKI, I. J. & J. BURCZYK. 2009. Simultaneous estimation of null alleles and inbreeding

coefficients. *The Journal of Heredity*. 100: 106–113.

CHYBICKI, I. J., A. OLEKSA & J. BURCZYK. 2011. Increased inbreeding and strong kinship structure in *Taxus baccata* estimated from both AFLP and SSR data. *Heredity*. 107: 589–600.

CLARK, R. W., W. S. BROWN, R. STECHERT & K. R. ZAMUDIO. 2008. Integrating individual behaviour and landscape genetics : the population structure of timber rattlesnake hibernacula. *Molecular ecology*. 17: 719–730.

CLOBERT, J. 2012. The common lizard (*Zootoca vivipara*, anciently *Lacerta vivipara*)- a model system for the study of the causes, mechanisms, and consequences of dispersal. In: J. Clobert (Ed.), Dispersal: Ecology and Evolution. Oxford University Press, Oxford, UK.

CLOBERT, J. & M. BAGUETTE (Eds). 2012. Dispersal Ecology and Evolution. Oxford University Press, USA.

CLOBERT, J., M. MASSOT, J. LECOMTE, G. SORCI, M. DE FRAIPONT & R. BARBAULT. 1994. Determinants of Dispersal Behavior: The Common Lizard as a Case Study. Pp. 183–206. In: L. J. Vitt & E. R. Pianka (Eds.), Lizard Ecology: Historical and Experimental Perspectives. Princeton University Press, New Jersey, USA. http://ecologie.snv.jussieu.fr/mmassot/Publicati ons/Clobert et al Lizard Ecology 1994.pdf (26 November 2013).

COCKERHAM, C. C. & B. S. WEIR. 1993. Estimation of Gene Flow from F-Statistics. *Evolution*. 47: 855–863.

COLLINGHAM, Y. C. & B. HUNTLEY. 2000. Impacts of habitat fragmentation and patch size upon migration rates. *Ecological Applications*. 10: 131–144.

CONRADIE, W., G. J. MEASEY, W. R. BRANCH & K. A. TOLLEY. 2012. Revised phylogeny of African sand lizards (*Pedioplanis*), with the description of two new species from south-western Angola. *African Journal of Herpetology*. 61: 1–22.

COOPER, W. E., V. PREZ-MELLADO & L. J. VITT. 2004. Ease and effectiveness of costly autotomy vary with predation intensity among lizard populations. *Journal of Zoology*. 262: 243–255.

COOPER, W. E. & M. J. WHITING. 1999. Foraging modes in lacertid lizards from southern Africa. *Amphibia-Reptilia*. 20: 299–311.

CORNUET, J. M. & G. LUIKART. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*. 144: 2001–2014.

COTE, J. & J. CLOBERT. 2007. Social personalities influence natal dispersal in a lizard. *Proceedings of The Royal Society B, Biological sciences*. 274: 383–390.

COULON, A, J. F. COSSON, J. M. ANGIBAULT, B. CARGNELUTTI, M. GALAN, N. MORELLET, E. PETIT, S. AULAGNIER & A J. M. HEWISON. 2004. Landscape connectivity influences gene

flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Molecular ecology*. 13: 2841–2850.

DAKIN, E. E. & J. C. AVISE. 2004. Microsatellite null alleles in parentage analysis. *Heredity*. 93: 504–509.

DAMME, R. & B. VANHOOYDONCK. 2002. Speed versus manoeuvrability: association between vertebral number and habitat structure in lacertid lizards. *Journal of Zoology*. 258: 327–334.

DANIELS, S. R., M. D. HOFMEYR, B. T. HENEN & K. A CRANDALL. 2007. Living with the genetic signature of Miocene induced change: evidence from the phylogeographic structure of the endemic angulate tortoise *Chersina angulata*. *Molecular Phylogenetics and Evolution*. 45: 915–26.

DESTRO-BISOL, G., F. DONATI, V. COIA, I. BOSCHI, F. VERGINELLI, A. CAGLIÀ, S. TOFANELLI, G. SPEDINI & C. CAPELLI. 2004. Variation of female and male lineages in sub-Saharan populations: the importance of sociocultural factors. *Molecular biology and evolution*. 21: 1673–1682.

DEUTSCH, C. A., J. J. TEWKSBURY, R. B. HUEY, K. S. SHELDON, C. K. GHALAMBOR, D. C. HAAK & P. R. MARTIN. 2008. Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences of the United States of America*. 105: 6668–6672.

DÍAZ, J. A., J. PÉREZ-TRIS, D. BAUWENS, D. PÉREZ-ARANDA, R. CARBONELL, T. SANTOS & J. L. TELLERÍA. 2007. Reproductive performance of a lacertid lizard at the core and the periphery of the species' range. *Biological Journal of the Linnean Society*. 92: 87–96.

DICICCIO, T. J. & B. EFRON. 1996. Bootstrap Confidence Intervals 11: 189–228.

DOLMAN, G. & C. MORITZ. 2006. A multilocus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). *Evolution*. 60: 573–582.

DOUGHTY, P. & B. SINERVO. 1994. The effects of habitat, time of hatching, and body size on the dispersal of hatchling *Uta stansburiana*. *Journal of Herpetology*. 28: 485–490.

DRISCOLL, D. A. 2004. Extinction and outbreaks accompany fragmentation of a reptile community. *Ecological Applications*. 14: 220–240.

DRISCOLL, D. A. & C. M. HARDY. 2005. Dispersal and phylogeography of the agamid lizard *Amphibolurus nobbi* in fragmented and continuous habitat. *Molecular ecology*. 14: 1613–1629.

DU, W., X. JI, Y. ZHANG, X. XU & R. SHINE. 2005. Identifying sources of variation in reproductive and life-history traits among five populations of a Chinese lizard (*Takydromus septentrionalis*, Lacertidae). *Biological Journal of the Linnean Society*. 85: 443–453.

DUBEY, S., J. SUMNER, D. A. PIKE, J. S. KEOGH, J. K. WEBB & R. SHINE. 2011. Genetic connectivity among populations of an endangered snake species from southeastern Australia (*Hoplocephalus bungaroides*, Elapidae). *Ecology and Evolution*. 1: 218–227.

DUCKETT, P. E., P. D. WILSON & A. J. STOW. 2013. Keeping up with the neighbours: using a genetic measurement of dispersal and species distribution modelling to assess the impact of climate change on an Australian arid zone gecko (*Gehyra variegata*). *Diversity and Distributions*. 19: 964–976.

DUCKWORTH, G. D., R. ALTWEGG & D. GUO. 2010. Soil moisture limits foraging: a possible mechanism for the range dynamics of the hadeda ibis in southern Africa. *Diversity and Distributions*. 16: 765–772.

DUFTNER, N., K. M. SEFC, S. KOBLMÜLLER, B. NEVADO, E. VERHEYEN, H. PHIRI & C. STURMBAUER. 2006. Distinct population structure in a phenotypically homogeneous rock-dwelling cichlid fish from Lake Tanganyika. *Molecular ecology*. 15: 2381–2395.

DUPONT, L. M., H. P. LINDER, F. ROMMERSKIRCHEN & E. SCHEFUß. 2011. Climate-driven rampant speciation of the Cape flora. *Journal of Biogeography*. 38: 1059–1068.

EDWARDS, S. 2013. Patterns and processes of adaptation in lacertid lizards to environments in southern Africa. Unpubl. Ph.D. Dissertation, Stellenbosch University, Stellenbosch, South Africa.

EDWARDS, S., W. R. BRANCH, B. VANHOOYDONCK, A. HERREL, G. J. MEASEY & K. A. TOLLEY. 2013. Taxonomic adjustments in the systematics of the southern African lacertid lizards (Sauria: Lacertidae). *Zootaxa*. 3669: 101–114.

EDWARDS, S., J. CLAUDE, B. J. VAN VUUREN & C. A. MATTHEE. 2011. Evolutionary history of the Karoo bush rat, *Myotomys unisulcatus* (Rodentia: Muridae): disconcordance between morphology and genetics. *Biological Journal of the Linnean Society*. 102: 510–526.

EDWARDS, S., B. VANHOOYDONCK, A. HERREL, G. J. MEASEY & K. A. TOLLEY. 2012. Convergent evolution associated with habitat decouples phenotype from phylogeny in a clade of lizards. *PloS one*. 7: e51636.

EFFORD, M. G. 2011. Estimation of population density by spatially explicit capture-recapture analysis of data from area searches. *Ecology*. 92: 2202–2207.

EFFORD, M. G., D. L. BORCHERS & A. E. BYROM. 2009. Density Estimation by Spatially Explicit Capture–Recapture: Likelihood-Based Methods. Pp. 255–269. In: D. L. Thomson, E. G. Cooch & M. J. Conroy (Eds.), Modeling Demographic Processes in Marked Populations: Environmental and Ecological Statistics, Vol. 3. Springer. DOI: 10.1007/978-0-387-78151-8_11.

EFFORD, M. G., D. K. DAWSON & C. S. ROBBINS. 2004. DENSITY: software for analysing

capture-recapture data from passive detector arrays. *Animal Biodiversity and Conservation*. 1: 217–228.

ELLEGREN, H. 2004. Microsatellites: simple sequences with complex evolution. *Nature reviews. Genetics*. 5: 435–445.

ENGLER, R. & A. GUISAN. 2009. MigClim: Predicting plant distribution and dispersal in a changing climate. *Diversity and Distributions*. 15: 590–601.

EXCOFFIER, L. & H. E. L. LISCHER. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*. 10: 564–567.

FAHRIG, L. 2003. Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics*. 34: 487–515.

FAUBET, P., R. WAPLES & O. GAGGIOTTI. 2007. Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Molecular Ecology*. 16: 1149–1166.

FELDHEIM, K. A., L. F. CHAUKE, K. P. HOPKINS & K. A. TOLLEY. 2010. Characterization of microsatellite loci from a South African endemic, the Cape Dwarf Chameleon (*Bradypodion pumilum*). *Conservation Genetics Resources*. 2: 165–168.

FITZSIMONS, V. F. 1943. The Lizards of South Africa. Transvaal Museum, Pretoria, South Africa.

FONTAINE, M. C., K. A. TOLLEY, J. R. MICHAUX, A. BIRKUN, M. FERREIRA, T. JAUNIAUX, A. LLAVONA, B. OZTÜRK, A. A OZTÜRK, V. RIDOUX, E. ROGAN, M. SEQUEIRA, J.-M. BOUQUEGNEAU & S. J. E. BAIRD. 2010. Genetic and historic evidence for climate-driven population fragmentation in a top cetacean predator: the harbour porpoises in European water. *Proceedings of The Royal Society B, Biological sciences*. 277: 2829–2837.

FRASER, D. J. & L. BERNATCHEZ. 2001. Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology*. 10: 2741–2752.

FRITZ, U., D. AYAZ, J. BUSCHBOM, H. G. KAMI, L. F. MAZANAEVA, A A ALOUFI, M. AUER, L. RIFAI, T. SILIĆ & A K. HUNDSDÖRFER. 2008. Go east: phylogeographies of *Mauremys caspica* and *M. rivulata*- discordance of morphology, mitochondrial and nuclear genomic markers and rare hybridization. *Journal of evolutionary biology*. 21: 527–540.

FUCHS, J., T. M. CROWE & R. C. K. BOWIE. 2011. Phylogeography of the fiscal shrike (*Lanius collaris*): a novel pattern of genetic structure across the arid zones and savannas of Africa. *Journal of Biogeography*. 38: 2210–2222.

GALÁN, P. 2004. Structure of a population of the lizard *Podarcis bocagei* in northwest Spain: variations in age distribution, size distribution and sex ratio. *Animal Biology*. 54: 57–75.

GARDNER, M. G., C. M. BULL, S. J. B. COOPER & G. A. DUFFIELD. 2001. Genetic evidence for a

family structure in stable social aggregations of the Australian lizard *Egernia stokesii*. *Molecular Ecology*. 10: 175–183.

GARDNER, M. G., M. C. BULL, S. J. B. COOPER & G. A. DUFFIELD. 2000. Microsatellite mutations in litters of the Australian lizard *Egernia stokesii*. *Journal of Evolutionary Biology*. 13: 551–560.

GARZA, J. & E. WILLIAMSON. 2001. Detection of reduction in population size using data from microsatellite loci. *Molecular ecology*. 10: 305–318.

GLENN, T. C. & N. A. SCHABLE. 2005. Isolating microsatellite DNA loci. *Methods in enzymology*. 395: 202–222.

GOLD, J. R., C. P. BURRIDGE & T. F. TURNER. 2001. A modified stepping-stone model of population structure in red drum, *Sciaenops ocellatus* (Sciaenidae), from the northern Gulf of Mexico. *Genetica*. 111: 305–317.

GOLDBERG, S. R. 2008. Reproductive cycle of the common rough-scaled lizard, *Ichnotropis squamulosa* (Squamata: Lacertidae) from southern Africa. *Texas Journal of Science*. : 1–6.

GOLDBERG, S. S. R. 2006. Reproductive cycle of the spotted sand lizard, *Pedioplanis lineoocellata* (Squamata: Lacertidae) from southern Africa. *Texas Journal of Science*. 58: 1–6.

GOVENDER, N., W. S. W. TROLLOPE & B. W. VAN WILGEN. 2006. The effect of fire season, fire frequency, rainfall and management on fire intensity in savanna vegetation in South Africa. *Journal of Applied Ecology*. 43: 748–758.

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

GUICHOUX, E., L. LAGACHE, S. WAGNER, P. CHAUMEIL, P. LÉGER, O. LEPAIS, C. LEPOITTEVIN, T. MALAUSA, E. REVARDEL, F. SALIN & R. J. PETIT. 2011. Current trends in microsatellite genotyping. *Molecular Ecology Resources*. 11: 591–611.

GUILLOT, G., F. MORTIER & A. ESTOUP. 2005. Geneland: a computer package for landscape genetics. *Molecular Ecology Notes*. 5: 712–715.

GULLBERG, A, H. TEGELSTRÖM & M. OLSSON. 1997. Microsatellites in the sand lizard (*Lacerta agilis*): description, variation, inheritance, and applicability. *Biochemical genetics*. 35: 281–295.

GUO, S. W. & E. A. THOMPSON. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*. 48: 361–372.

HALPIN, Z. T. 1987. Natal dispersal and the formation of new social groups in a newly established town of Black-tailed prairie dogs (*Cynomys ludovicianus*). Pp. 104–118. In: B. D. Chepko-Sade & Z. T. Halpin (Eds.), Mammalian Dispersal Patterns: The Effects of Social Structure on Population Genetics. University of Chicago Press, Chicago, USA.

HAMILTON, W. D. & R. M. MAY. 1977. Dispersal in stable habitats. *Nature*. 269: 578–581.

HANSKI, I. 1998. Metapopulation dynamics. *Nature*. 396: 41–49.

HANSKI, I. 1999. Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes. *Oikos*. 87: 209–219.

HARRIS, D. J., E. N. ARNOLD & R. H. THOMAS. 1998. Relationships of lacertid lizards (Reptilia: Lacertidae) estimated from mitochondrial DNA sequences and morphology. *Proceedings of The Royal Society B, Biological sciences*. 265: 1939–1948.

HARTL, D. L. & A. G. CLARK. 1997. Principles of Population Genetics. Pp. 1–542. Sinauer Associates, Inc., Sunderland, USA.

HELLBORG, L., C. W. WALKER, E. K. RUENESS, J. E. STACY, I. KOJOLA, H. VALDMANN & C. VIL. 2002. Differentiation and levels of genetic variation in northern European lynx (*Lynx lynx*) populations revealed by microsatellites and mitochondrial DNA analysis. *Conservation Genetics*. 3: 97–111.

HEY, J. 2010. Isolation with migration models for more than two populations. *Molecular biology and evolution*. 27: 905–920.

HIPSLEY, C. A., L. HIMMELMANN, D. METZLER & J. MÜLLER. 2009. Integration of Bayesian molecular clock methods and fossil-based soft bounds reveals early Cenozoic origin of African lacertid lizards. *BMC Evolutionary Biology*. 9: 151–164.

HOEHN, M., S. D. SARRE & K. HENLE. 2007. The tales of two geckos: does dispersal prevent extinction in recently fragmented populations? *Molecular ecology*. 16: 3299–3312.

HOLT, R. D. & M. BARFIELD. 2001. On the relationship between the ideal free distribution and the evolution of dispersal. Pp. 83–95. In: J. Clobert, E. Danchin, A. Dhondt & J. Nichols (Eds.), Dispersal. Oxford University Press, New York, USA. http://people.biology.ufl.edu/rdholt/holtpublications/107.pdf.

HOLZHAUER, S. I. J., K. EKSCHMITT, A.-C. SANDER, J. DAUBER & V. WOLTERS. 2006. Effect of historic landscape change on the genetic structure of the bush-cricket *Metrioptera roeseli*. *Landscape Ecology*. 21: 891–899.

HOLZMAN, J. P., A J. BOHONAK, L. R. KIRKENDALL, D. GOTTLIEB, A R. HARARI & S. T. KELLEY. 2009. Inbreeding variability and population structure in the invasive haplodiploid palm-seed borer (*Coccotrypes dactyliperda*). *Journal of evolutionary biology*. 22: 1076–87.

HOUSE, S. M., P. J. TAYLOR & I. F. SPELLERBERG. 1980. Patterns of daily behaviour in two lizard species *Lacerta agilis* L . and *Lacerta vivipara* Jacquin. *Oecologia*. 44: 396–402.

HUEY, R. B., A. F. BENNETT, H. JOHN-ALDER & K. A. NAGY. 1984. Locomotor capacity and foraging behaviour of Kalahari lacertid lizards. *Animal Behaviour*. 32: 41–50.

HUEY, R. & M. SLATKIN. 1976. Cost and benefits of lizard thermoregulation. *Quarterly Review of Biology*. 51: 363–384.

JACOBSEN, N. 1987. Notes on reproduction in *Ichnotropis squamulosa* and interspecific competition with *I. capensis* (Reptilia, Lacertidae) in the Transvaal. *The Journal of the Herpetological Association of Africa*. 1: 37–41.

JENSEN, J. L., A. J. BOHONAK & S. T. KELLEY. 2005. Isolation by distance, web service. *BMC genetics*. 6: 13–19.

JOHNSON, M. L. & M. S. GAINES. 1990. Evolution of Dispersal: Theoretical Models and Empirical Tests Using Birds and Mammals. *Annual Review of Ecology and Systematics*. 21: 449–480.

JOMBART, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 24: 1403–1405.

JOMBART, T. 2013. A tutorial for Discriminant Analysis of Principal Components (DAPC) using adegenet 1.4-0.

JOMBART, T., S. DEVILLARD & F. BALLOUX. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics*. 11: 94–109.

KARSTEN, M., B. J. VAN VUUREN, A. BARNAUD & J. S. TERBLANCHE. 2013. Population genetics of *Ceratitis capitata* in South Africa: implications for dispersal and pest management. *PloS one*. 8: e54281.

KENDALL, W. L. 2001. The robust design for capture-recapture studies: analysis using program MARK. Pp. 357–360. In: R. Field, R. J. Warren, H. Okarma & P. R. Sievert (Eds.), Wildlife, land, and people: priorities for the 21st century Proceedings of the Second. International Wildlife Management Congress, held June 28-July2, 1999. The Wildlife Society, Bethesda, Maryland, USA. http://medcontent.metapress.com/index/A65RM03P4874243N.pdf (15 August 2013).

KEOGH, J. S., J. K. WEBB & R. SHINE. 2007. Spatial genetic analysis and long-term mark-recapture data demonstrate male-biased dispersal in a snake. *Biology Letters*. 3: 33–35.

KIMURA, M. & G. H. WEISS. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*. 49: 561–576.

KRYGER, U., T. J. ROBINSON & P. BLOOMER. 2004. Population structure and history of southern African scrub hares, *Lepus saxatilis*. *Journal of Zoology*. 263: 121–133.

LALOI, D., M. RICHARD, J. LECOMTE, M. MASSOT & J. CLOBERT. 2004. Multiple paternity in clutches of common lizard *Lacerta vivipara*: data from microsatellite markers. *Molecular Ecology*. 13: 719–723.

LAMB, T. & A. M. BAUER. 2003. *Meroles* revisited: complementary systematic inference from additional mitochondrial genes and complete taxon sampling of southern Africa's desert lizards. *Molecular Phylogenetics and Evolution*. 29: 360–364.

LAWSON, E. & A. RODGERS. 1997. Differences in home-range size computed in commonly used software programs. *Wildlife Society Bulletin*. 25: 721–729.

LEBLOIS, R., F. ROUSSET, D. TIKEL, C. MORITZ & A. ESTOUP. 2000. Absence of evidence for isolation by distance in an expanding cane toad (*Bufo marinus*) population: an individual-based analysis of microsatellite genotypes. *Molecular ecology*. 9: 1905–1909.

LEBRETON, J. & K. BURNHAM. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs*. 62: 67–118.

LEMER, S., E. ROCHEL & S. PLANES. 2011. Correction method for null alleles in species with variable microsatellite flanking regions, a case study of the black-lipped pearl oyster *Pinctada margaritifera*. *The Journal of Heredity*. 102: 243–246.

LÉNA, J. P., J. CLOBERT, M. DE FRAIPONT, J. LECOMTE & G. GUYOT. 1998a. The relative influence of density and kinship on dispersal in the common lizard. *Behavioral Ecology*. 9: 500–507.

LÉNA, J. P. & M. DE FRAIPONT. 1998b. Kin recognition in the common lizard. *Behavioral Ecology and Sociobiology*. 42: 341–347.

LESCANO, N. V. 2010. Population bottleneck and range expansion in *Podocris muralis*, a wall lizard introduced from Italy. Unpubl. MSc Thesis, University of Cincinnati, Cincinnati, USA.

LI, H., Z.-S. ZHOU, J. GUO & L. H. LIN. 2012. Polymorphic microsatellite loci in the rapid racerunner *Eremias velox* (Squamata: Lacertidae). *Genetics and Molecular Research*. 11: 4707–4710.

LI, M. & J. MERILÄ. 2010. Sex-specific population structure, natural selection, and linkage disequilibrium in a wild bird population as revealed by genome-wide microsatellite analyses. *BMC evolutionary biology*. 10: 66.

LORENZEN, E. D., R. HELLER & H. R. SIEGISMUND. 2012. Comparative phylogeography of African savannah ungulates. *Molecular ecology*. 21: 3656–3670.

LORENZEN, E. D., C. MASEMBE, P. ARCTANDER & H. R. SIEGISMUND. 2010. A long-standing Pleistocene refugium in southern Africa and a mosaic of refugia in East Africa: insights from mtDNA and the common eland antelope. *Journal of Biogeography*. 37: 571–581.

LUIKART, G., F. W. ALLENDORF, J. M. CORNUET & W. B. SHERWIN. 1998a. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *The Journal of Heredity*. 89: 238–247.

LUIKART, G., W. B. SHERWIN, B. M. STEELE & F. W. ALLENDORF. 1998b. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular ecology*. 7: 963–374.

MAKOKHA, J. S., A. M. BAUER, W. MAYER & C. A MATTHEE. 2007. Nuclear and mtDNA-based phylogeny of southern African sand lizards, *Pedioplanis* (Sauria: Lacertidae). *Molecular Phylogenetics and Evolution*. 44: 622–633.

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

MARTINS, E. P. 1994. Phylogenetic perspectives on the evolution of lizard territoriality. Pp. 117–144. In: L. J. Vitt & E. R. Pianka (Eds.), Lizard Ecology: Historical and Experimental Perspectives. Princeton University Press, New Jersey, USA.

MAYER, W. & P. BEYERLEIN. 1999. Ecological niche segregation of seven sympatric lacertid lizards in the Peloponnese highlands. *Nat. Croat.* 8: 339–344.

MCBRAYER, L. D. & J. E. WYLIE. 2009. Concordance between locomotor morphology and foraging mode in lacertid lizards. *Zoology*. 112: 370–378.

MCPEEK, M. A. & R. D. HOLT. 1992. The evolution of dispersal in spatially and temporally varying environments. *The American naturalist*. 140: 1010–1027.

MEYLAN, S., J. BELLIURE, J. CLOBERT & M. DE FRAIPONT. 2002. Stress and body condition as prenatal and postnatal determinants of dispersal in the common lizard (*Lacerta vivipara*). *Hormones and Behavior*. 42: 319–326.

MIDGLEY, G. F., G. O. HUGHES, W. THUILLER & A. G. REBELO. 2006. Migration rate limitations on climate change-induced range shifts in Cape Proteaceae. *Diversity & Distributions*. 12: 555–562.

MILNE, T., M. C. BULL & M. N. HUTCHINSON. 2002. Characteristics of litters and juvenile dispersal in the endangered Australian skink *Tiliqua adelaidensis*. *Journal of Herpetology*. 36: 110–112.

MIRALDO, A., C. FARIA, G. M. HEWITT, O. S. PAULO & B. C. EMERSON. 2013. Genetic analysis of a contact zone between two lineages of the ocellated lizard (*Lacerta lepida* Daudin 1802) in south-eastern Iberia reveal a steep and narrow hybrid zone. *Journal of Zoological Systematics and Evolutionary Research*. 51: 45–54.

MIRALDO, A., G. M. HEWITT, O. S. PAULO & B. C. EMERSON. 2011. Phylogeography and demographic history of *Lacerta lepida* in the Iberian Peninsula: multiple refugia, range expansions and secondary contact zones. *BMC evolutionary biology*. 11: 170.

NATHAN, R., E. KLEIN, J. J. ROBLEDO-ARNUNCIO & E. REVILLA. 2012. Dispersal kernels: review. Pp. 187–210. In: J. Clobert & M. Baguette (Eds.), Dispersal: Ecology and Evolution. Oxford University Press, USA. http://www.myilibrary.com?ID=395195.

NIEWIAROWSKI, P. H. 1994. Understanding geographic life-history variation in lizards. Pp. 31–50. In: L. J. Vitt & E. R. Pianka (Eds.), Lizard Ecology: Historical and Experimental

Perspectives. Princeton University Press, New Jersey, USA.

NUZZO, V. A, J. C. MAERZ & B. BLOSSEY. 2009. Earthworm invasion as the driving force behind plant invasion and community change in northeastern North American forests. *Conservation Biology*. 23: 966–974.

OLSSON, M., A. GULLBERG & H. TEGELSTRÖM. 1996. Malformed offspring, sibling matings, and selection against inbreeding in the sand lizard (*Lacerta agilis*). *Journal of Evolutionary Biology*. 9: 229–242.

OLSSON, M. & R. SHINE. 2003. Female-biased natal and breeding dispersal in an alpine lizard, *Niveoscincus microlepidotus*. *Biological Journal of the Linnean Society*. 79: 277–283.

VAN OOSTERHOUT, C., W. F. HUTCHINSON, D. P. M. WILLS & P. SHIPLEY. 2004. Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Molecular Ecology Notes*. 4: 535–538.

PAETKAU, D., R. SLADE, M. BURDEN & A. ESTOUP. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular ecology*. 13: 55–65.

PEAKALL, R., M. RUIBAL & D. B. LINDENMAYER. 2003. Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. *Evolution*. 57: 1182–1195.

PEAKALL, R. & P. E. SMOUSE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics*. 28: 2537–2539.

PEARSON, R. G. & T. P. DAWSON. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*. 12: 361–371.

PERRY, G. & T. GARLAND, JR. 2002. Lizard home ranges revisited: effects of sex, body size, diet, habitat, and phylogeny. *Ecology*. 83: 1870–1885.

PFENNINGER, M., D. POSADA & F. MAGNIN. 2003. Evidence for survival of Pleistocene climatic changes in Northern refugia by the land snail *Trochoidea geyeri* (Soós 1926) (Helicellinae, Stylommatophora). *BMC evolutionary biology*. 3: 8.

PINCEEL, J., K. JORDAENS & T. BACKELJAU. 2005. Extreme mtDNA divergences in a terrestrial slug (Gastropoda, Pulmonata, Arionidae): accelerated evolution, allopatric divergence and secondary contact. *Journal of evolutionary biology*. 18: 1264–1280.

PIO, D. V., R. ENGLER, H. P. LINDER, A. MONADJEM, F. P. D. COTTERILL, P. J. TAYLOR, M. C. SCHOEMAN, B. W. PRICE, M. H. VILLET, G. EICK, N. SALAMIN & A. GUISAN. 2014. Climate change effects on animal and plant phylogenetic diversity in southern Africa. *Global Change Biology*. 20: 1538–1549.

PIRY, S., G. LUIKART & J. M. CORNUET. 1999. BOTTLENECK: a computer program for detecting recent reducutions in the effective population size using allele frequency data. *Journal of Heredity*. 90: 502–503.

PLESSIS, I. & P. MOUTON. 2011. Habitat preferences of three sympatric lacertid lizards in the arid Tankwa Karoo Basin of South Africa. *African Zoology*. 46: 88–94.

PRINSLOO, P. & T. J. ROBINSON. 1992. Geographic mitochondrial DNA variation in the rock hyrax, *Procavia capensis*. *Molecular Biology and Evolution*. 9: 447–456.

PRITCHARD, J. K., M. STEPHENS & P. DONNELLY. 2000. Inference of population structure using multilocus genotype data. *Genetics*. 155: 945–959.

QGIS DEVELOPMENT TEAM. 2013. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org.

R CORE TEAM. 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org (9 July 2014).

RANNALA, B. 2013, June 5. BayesAss Edition 3.0 User's Manual. University of California Davis, California, USA. http://rannala.org/docs/BayesAss.1.3.pdf (5 June 2013).

RAYBOULD, A. F., R. T. CLARKE, J. M. BOND, R. E. WELTERS & C. J. GLIDDON. 2002. Inferring patterns of dispersal from allele frequency data.

Pp. 89–112. In: J. M. Bullock, R. E. Kenward & R. S. Hails (Eds.), Dispersal Ecology. Blackwell Publishing Ltd.

RAYMOND, M. & F. ROUSSET. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*. 86: 248–249.

REMÓN, N., P. GALÁN, M. VILA, O. ARRIBAS & H. NAVEIRA. 2013. Causes and evolutionary consequences of population subdivision of an Iberian mountain lizard, *Iberolacerta monticola*. *PloS one*. 8: e66034.

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

RICHARD, M. & R. S. THORPE. 2000. Highly polymorphic microsatellites in the lacertid *Gallotia galloti* from the western Canary Islands. *Molecular Ecology*. 9: 1919–1952.

RONCE, O., J. CLOBERT & M. MASSOT. 1998. Natal dispersal and senescence. *Proceedings of the National Academy of Sciences of the United States of America*. 95: 600–605.

ROONEY, A P., R. L. HONEYCUTT, S. K. DAVIS & J. N. DERR. 1999. Evaluating a putative bottleneck in a population of bowhead whales from patterns of microsatellite diversity and genetic disequilibria. *Journal of Molecular Evolution*. 49: 682–690.

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

ROUSSET, F. 2000. Genetic differentiation between individuals. *Journal of Evolutionary Biology*. 13: 58–62.

ROUSSET, F. 2004. Genetic structure and selection in subdivided populations. P. 264. Princeton University Press, Princeton and Oxford. http://books.google.com/books?id=rB7W9PE1-DEC&pgis=1.

ROUSSET, F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*. 8: 103–106.

ROW, J. R., G. BLOUIN-DEMERS & S. C. LOUGHEED. 2010. Habitat distribution influences dispersal and fine-scale genetic population structure of eastern foxsnakes (*Mintonius gloydi*) across a fragmented landscape. *Molecular ecology*. 19: 5157–5171.

ROYLE, J. A. & K. V YOUNG. 2008. A hierarchical model for spatial capture-recapture data. *Ecology*. 89: 2281–2289.

SCHOWALTER, T. D. 2006. Insect ecology: an ecosystem approach. Elsevier Inc., London. http://books.google.com/books?hl=en&lr=&id=LQqHWCtj0F0C&oi=fnd&pg=PP2&dq=Insect+Ecology-An+Ecosystem+Approach&ots=CqLUxQlOjx&sig=FL96W4K0nPCje4swor5Dn2ayhGA (8 May 2012).

SCHREGEL, J., A. KOPATZ, S. B. HAGEN, H. BRØSETH, M. E. SMITH, S. WIKAN, I. WARTIAINEN, P. E. ASPHOLM, J. ASPI, J. E. SWENSON, O. MAKAROVA, N. POLIKARPOVA, M. SCHNEIDER, P. M. KNAPPSKOG, M. RUOKONEN, I. KOJOLA, K. F. TIRRONEN, P. I. DANILOV & H. G. EIKEN. 2012. Limited gene flow among brown bear populations in far Northern Europe? Genetic analysis of the east-west border population in the Pasvik Valley. *Molecular ecology*. 21: 3474–3488.

SCHREY, A. W., K. G. ASHTON, S. HEATH, E. D. MCCOY & H. R. MUSHINSKY. 2011. Fire alters patterns of genetic diversity among 3 lizard species in Florida Scrub habitat. *The Journal of Heredity*. 102: 399–408.

SCHULTE, U., M. VEITH & A. HOCHKIRCH. 2012. Rapid genetic assimilation of native wall lizard populations (*Podarcis muralis*) through extensive hybridization with introduced lineages. *Molecular ecology*. 21: 4313–4326.

SEIELSTAD, M. T., E. MINCH & L. L. CAVALLI-SFORZA. 1998. Genetic evidence for a higher female migration rate in humans. *Nature genetics*. 20: 278–280.

SELKOE, K. A & R. J. TOONEN. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*. 9: 615–629.

SHIELDS, W. M. 1987. Dispersal and mating systems: investigating their causal connections. Pp. 3–24. In: B. D. Chepko-Sade & Z. T. Halpin (Eds.), Mammalian Dispersal Patterns: The Effects of Social Structure on Population Genetics. The University of Chicago Press, Chicago.

DA SILVA, J. M. 2013. An investigation of the evolutionary diversification of a recent radiation of dwarf chameleons (*Bradypodion*) from KwaZulu-Natal Province, South Africa. Unpubl. Ph.D. Dissertation, Stellenbosch University, Stellenbosch, South Africa.

SINERVO, B., F. MÉNDEZ-DE-LA-CRUZ, D. B. MILES, B. HEULIN, E. BASTIAANS, M. VILLAGRÁN-SANTA CRUZ, R. LARA-RESENDIZ, N. MARTÍNEZ-MÉNDEZ, M. L. CALDERÓN-ESPINOSA, R. N. MEZA-LÁZARO, H. GADSDEN, L. J. AVILA, M. MORANDO, I. J. DE LA RIVA, P. VICTORIANO SEPULVEDA, C. F. D. ROCHA, N. IBARGÜENGOYTÍA, C. AGUILAR PUNTRIANO, M. MASSOT, V. LEPETZ, T. A OKSANEN, D. G. CHAPPLE, A. M. BAUER, W. R. BRANCH, J. CLOBERT & J. W. SITES. 2010. Erosion of lizard diversity by climate change and altered thermal niches. *Science (New York, N.Y.)*. 328: 894–899.

SLATKIN, M. 1991. Inbreeding coefficients and coalescence times. *Genetical Research*. 89: 479–487.

SLATKIN, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*. 139: 457–462.

SMIT, H. A., B. JANSEN VAN VUUREN, P. C. M. O'BRIEN, M. FERGUSON-SMITH, F. YANG & T. J. ROBINSON. 2011. Phylogenetic relationships of elephant-shrews (Afrotheria, Macroscelididae). *Journal of Zoology*. 284: 133–143.

SMIT, H. A., T. J. ROBINSON & B. J. VAN VUUREN. 2007. Coalescence methods reveal the impact of vicariance on the spatial genetic structure of *Elephantulus edwardii* (Afrotheria, Macroscelidea). *Molecular ecology*. 16: 2680–2692.

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

SOKAL, R. R., G. M. JACQUEZ & M. C. WOOTEN. 1989. Spatial autocorrelation analysis of migration and selection. *Genetics*. 121: 845–855.

SOKAL, R. & D. WARTENBERG. 1983. A test of spatial autocorrelation analysis using an isolation-by-distance model. *Genetics*. 105: 219–237.

SOMMA, L. A. 2003. Reptilian parental behavior. *The Linnean*. 19: 42–46.

SORCI, G., M. MASSOT & J. CLOBERT. 1994. Maternal parasite load increases sprint speed and philopatry in female offspring of the common lizard. *The American Naturalist*. 144: 153–164.

STOW, A. J. & P. SUNNUCKS. 2004. High mate and site fidelity in Cunningham's skinks (*Egernia cunninghami*) in natural and fragmented habitat. *Molecular Ecology*. 13: 419–430.

STOW, A. J., P. SUNNUCKS, D. A. BRISCOE & M. G. GARDNER. 2001. The impact of habitat fragmentation on dispersal of Cunningham's skink (*Egernia cunninghami*): evidence from allelic and genotypic analyses of microsatellites. *Molecular ecology*. 10: 867–878.

Page | 114

SUMNER, J., F. ROUSSET, A. ESTOUP & C. MORITZ. 2001. "Neighbourhood"size, dispersal and density estimates in the prickly forest skink (*Gnypetoscincus queenslandiae*) using individual genetic and demographic methods. *Molecular Ecology*. 10: 1917–1927.

SUTHERLAND, W. J., S. BARDSLEY, L. BENNUN, M. CLOUT, I. M. CÔTÉ, M. H. DEPLEDGE, L. V DICKS, A. P. DOBSON, L. FELLMAN, E. FLEISHMAN, D. W. GIBBONS, A. J. IMPEY, J. H. LAWTON, F. LICKORISH, D. B. LINDENMAYER, T. E. LOVEJOY, R. MAC NALLY, J. MADGWICK, L. S. PECK, J. PRETTY, S. V PRIOR, K. H. REDFORD, J. P. W. SCHARLEMANN, M. SPALDING & A. R. WATKINSON. 2011. Horizon scan of global conservation issues for 2011. *Trends in Ecology & Evolution*. 26: 10–16.

SVENNING, J.-C., S. NORMAND & M. KAGEYAMA. 2008. Glacial refugia of temperate trees in Europe: insights from species distribution modelling. *Journal of Ecology*. 96: 1117–1127.

SWART, B. L. 2006. The phylogeography of the southern rock agama (*Agama atra*) in the Cape Fold Mountains, South Africa. Unpubl. M.Sc. Thesis, Stellenbosch University, Stellenbosch, South Africa.

SWART, B. L., K. A. TOLLEY & C. A. MATTHEE. 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.

TEMPLETON, A. R., H. BRAZEAL & J. L. NEUWALD. 2011. The transition from isolated patches to a metapopulation in the eastern collared lizard in response to prescribed fires. *Ecology*. 92: 1736–1747.

TISHKOFF, S. A, M. K. GONDER, B. M. HENN, H. MORTENSEN, A. KNIGHT, C. GIGNOUX, N. FERNANDOPULLE, G. LEMA, T. B. NYAMBO, U. RAMAKRISHNAN, F. A REED & J. L. MOUNTAIN. 2007. History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. *Molecular biology and evolution*. 24: 2180–2195.

TOKARSKA, M., T. MARSHALL, R. KOWALCZYK, J. M. WÓJCIK, C. PERTOLDI, T. N. KRISTENSEN, V. LOESCHCKE, V. R. GREGERSEN & C. BENDIXEN. 2009. Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: the case of European bison. *Heredity*. 103: 326–332.

TOLLEY, K. A., R. C. K. BOWIE, B. W. PRICE, G. J. MEASEY & F. FOREST. 2014a. The shifting landscape of genes since the Pliocene: terrestrial phylogeography in the Greater Cape Floristic Region. Pp. 142–163. In: N. Allsopp, J. F. Colville & A. Verboom (Eds.), Ecology, Evolution and Conservation of a Megadiverse Region. Oxford University Press, Cape Town, South Africa.

TOLLEY, K. A., A. BRAAE & M. CUNNINGHAM. 2010a. Phylogeography of the Clicking Stream Frog *Strongylopus grayii* (Anura, Pyxicephalidae) reveals cryptic divergence across climatic zones in an abundant and

widespread taxon. *African Journal of Herpetology*. 59: 17–32.

TOLLEY, K. A., M. BURGER, A. A TURNER & C. A MATTHEE. 2006. Biogeographic patterns and phylogeography of dwarf chameleons (*Bradypodion*) in an African biodiversity hotspot. *Molecular ecology*. 15: 781–793.

TOLLEY, K. A., R. J. DANIELS & K. A. FELDHEIM. 2014b. Characterisation of microsatellite markers in the Spotted Sand Lizard (*Pedioplanis lineoocellata*) shows low levels of inbreeding and moderate genetic diversity on a small spatial scale. *African Journal of Herpetology*. 63: 1–11.

TOLLEY, K. A., J. S. MAKOKHA, D. T. HOUNIET, B. L. SWART & C. A MATTHEE. 2009. The potential for predicted climate shifts to impact genetic landscapes of lizards in the South African Cape Floristic Region. *Molecular Phylogenetics and Evolution*. 51: 120–130.

TOLLEY, K. A., R. RAW & R. ALTWEGG. 2010b. Chameleons on the move: survival and movement of the Cape Dwarf Chameleon, *Bradypodion pumilum*, within a fragmented urban habitat. *African Zoology*. 45: 99–106.

UJVARI, B., M. DOWNTON & T. MADSEN. 2008. Population genetic structure, gene flow and sex-biased dispersal in frillneck lizards (*Chlamydosaurus kingii*). *Molecular Ecology*. 17: 3557–3564.

ULLER, T. & M. OLSSON. 2008. Multiple paternity in reptiles: patterns and processes. *Molecular ecology*. 17: 2566–2580.

URBAN, M. C., B. L. PHILLIPS, D. K. SKELLY & R. SHINE. 2008. A toad more traveled: the heterogeneous invasion dynamics of cane toads in Australia. *The American naturalist*. 171: E134–148.

VANHOOYDONCK, B., R. S. JAMES, J. TALLIS, P. AERTS, Z. TADIC, K. A. TOLLEY, G. J. MEASEY & A. HERREL. 2014. Is the whole more than the sum of its parts? Evolutionary trade-offs between burst and sustained locomotion in lacertid lizards. *Proceedings of The Royal Society B, Biological sciences*. 281: 20132677.

VERCKEN, E., B. SINERVO & J. CLOBERT. 2012. The importance of a good neighborhood: dispersal decisions in juvenile common lizards are based on social environment. *Behavioral Ecology*. 23: 1059–1067.

VERWAIJEN, D. & R. VAN DAMME. 2008a. Foraging mode and its flexibility in lacertid lizards from Europe. *Journal of Herpetology*. 42: 124–133.

VERWAIJEN, D. & R. VAN DAMME. 2008b. Wide home ranges for widely foraging lizards. *Zoology (Jena, Germany)*. 111: 37–47.

VERWAIJEN, D. & R. VAN DAMME. 2008c. Foraging mode and locomotor capacities in Lacertidae. *Amphibia-Reptilia*. 29: 197–206.

VERWAIJEN, D., R. VAN DAMME & A. HERREL. 2002. Relationships between head size, bite force, prey handling efficiency and diet in two

sympatric lacertid lizards. *Functional Ecology*. 16: 842–850.

VISSER, J. H. 2013. Gene-flow in the rock hyrax (*Procavia capensis*) at different spatial scales. Unpubl. M.Sc. Thesis, Stellenbosch University, Stellenbosch, South Africa.

WAGNER, A P., S. CREEL & S. T. KALINOWSKI. 2006. Estimating relatedness and relationships using microsatellite loci with null alleles. *Heredity*. 97: 336–345.

WANG, I. J., W. K. SAVAGE & H. BRADLEY SHAFFER. 2009. Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (*Ambystoma californiense*). *Molecular ecology*. 18: 1365–1374.

WANG, J. 2008. COLONY. Wang, Jinliang, London, UK. http://www.zsl.org/science/software/colony.

WANG, J. 2009. A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular ecology*. 18: 2148–2164.

WAPLES, R. S. & O. GAGGIOTTI. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular ecology*. 15: 1419–39.

WASIOLKA, B. 2007. The impact of overgrazing on reptile diversity and population dynamics of *Pedioplanis l. lineoocellata* in the southern Kalahari. Unpubl. PhD Dissertation, Universität Potsdam, Germany. http://opus.kobv.de/ubp/volltexte/2008/1661/.

WASIOLKA, B., N. BLAUM, F. JELTSCH & J. HENSCHEL. 2009a. Behavioural responses of the lizard *Pedioplanis l. lineoocellata* to overgrazing. *Acta Oecologica*. 35: 157–162.

WASIOLKA, B., F. JELTSCH, J. HENSCHEL & N. BLAUM. 2009b. Space use of the spotted sand lizard (*Pedioplanis l. lineoocellata*) under different degradation states. *African Journal of Ecology*. 48: 96–104.

WATTS, P. C., F. ROUSSET, I. J. SACCHERI, R. LEBLOIS, S. J. KEMP & D. J. THOMPSON. 2007. Compatible genetic and ecological estimates of dispersal rates in insect (*Coenagrion mercuriale*: Odonata: Zygoptera) populations: analysis of "neighbourhood size" using a more precise estimator. *Molecular ecology*. 16: 737–751.

WHITLOCK, M. C. & D. E. MCCAULEY. 1999. Indirect measures of gene flow and migration : F ST ⊢1 /(4Nm+1). *Heredity*. 82: 117–125.

WIENS, J. J. 1997. Metapopulation dynamics and landscape ecology. Pp. 43–62. In: I. Hanski & M. Gilpin (Eds.), Metapopulation biology: Ecology, Genetics and Evolution. Academic Press, London, UK.

WILSON, D., R. L. MULVEY & R. D. CLARK. 2007. Sampling skinks and geckos in artificial cover objects in a dry mixed grassland-shrubland with mammalian predator control. *New Zealand Journal of Ecology*. 31: 169–185.

WILSON, G. A. & B. RANNALA. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*. 163: 1177–1191.

WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics*. 16: 97–159.

WRIGHT, S. 1935. Evolution in populations in approximate equilibrium. *Journal of Genetics*. 30: 257–266.

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

WRIGHT, S. 1946. Isolation by distance under diverse systems of mating. *Genetics*. 31: 39–59.

WRIGHT, S. 1978. Evolution and genetics of populations vol. 4: Variability within and among natural populations. University of Chicago Press, Chicago, USA.

ZANI, P. A., T. D. JONES, R. A. NEUHAUS & J. E. MILGROM. 2009. Effect of refuge distance on escape behavior of side-blotched lizards (*Uta stansburiana*). *Canadian Journal of Zoology*. 87: 407–414.

ZHAO, Q., H.-X. LIU, L.-G. LUO & X. JI. 2011. Comparative population genetics and phylogeography of two lacertid lizards (*Eremias argus* and *E. brenchleyi*) from China. *Molecular Phylogenetics and Evolution*. 58: 478–491.

Appendix A COMPLEX SPATIAL GENETIC PATTERNS AND EXTENSIVE SECONDARY CONTACT IN THE WIDESPREAD SPOTTED SAND LIZARD (*PEDIOPLANIS LINEOOCELLATA*)

Table A 1: List of all *Pedioplanis lineoocellata* genetic samples used in this thesis. HerpBank tissue index (INDEX), sample code (SAMPLE), former sub-species status (SUB-SPP), COUNTRY (SA=South Africa, NM=Namibia), province (PROV; NC=Northern Cape, WC= Western Cape, LP=Limpopo province), the primary sample site to which the sample belongs (SITE) and the general sampling area(AREA) are indicated. Also indicated is the mtDNA clade to which samples belong (CLADE) based on geographic position and, if genotyped (GENOTYPED=Y), mitochondrial DNA sequences from 16S and ND4 regions (Edwards 2013).

INDEX	SAMPLE	SUB-SPP	COUNTRY	PROV	SITE	AREA	CLADE	GENOTYPED
1199	AMB6214	lineoocellata	SA	NC		Askham	A	Y
1233	AMB7656	lineoocellata	NM			Aus	A	Y
1234	AMB7657	lineoocellata	NM			Aus	A	Y
1235	AMB7658	lineoocellata	NM			Aus	A	Y
3953	QQ0337	pulchella	SA	NC		Calvinia	A	Y
3689	MH0336	lineoocellata	SA	WC		Cederberg	A	Y
3607	CF028	lineoocellata	SA	WC		Cederberg	A	N
1183	CF134	lineoocellata	SA	WC		Cederberg	A	N
3688	MH0313	lineoocellata	SA	WC		Cederberg	A	N
1465	MB20613	pulchella	SA	NC		Farm Eselkopvlakte	A	N
1582	MB20765	lineoocellata	SA	NC	ESLF	Farm Eselsfontein	A	N
1600	MB20786	lineoocellata	SA	NC	ESLF	Farm Eselsfontein	A	N
1607	MB20797	lineoocellata	SA	NC	ESLF	Farm Eselsfontein	A	N
1608	MB20798	lineoocellata	SA	NC	ESLF	Farm Eselsfontein	A	N
1615	MB20810	lineoocellata	SA	NC	ESLF	Farm Eselsfontein	A	N
2009	H1685	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	Y
2008	H1688	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	Y

INDEX	SAMPLE	SUB-SPP	COUNTRY	PROV	SITE	AREA	CLADE	GENOTYPED
2013	H6158	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	Y
2011	H6176	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	Y
2022	H1686	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2019	H1689	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2020	H1690	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2010	H1691	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2016	H6156	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2021	H6157	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2017	H6159	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2018	H6171	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2015	H6172	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2023	H6173	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2012	H6174	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2014	H6175	lineoocellata	SA	WC	GANS	Farm Gansfontein	A	N
2223	JM03547	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	A	Y
2224	JM03557	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	A	Y
2234	JM03565	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	A	Y
2235	JM03566	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	A	Y
3568	JM03538	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
3569	JM03540	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
2222	JM03541	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
3571	JM03545	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
3763	JM03548	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
3574	JM03552	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
3576	JM03554	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N

INDEX	SAMPLE	SUB-SPP	COUNTRY	PROV	SITE	AREA	CLADE	GENOTYPED
2236	JM03567	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
2237	JM03570	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
2238	JM03571	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
2239	JM03573	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
2240	JM03575	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
5749	MBUR00004	lineoocellata	SA	NC	HART	Farm Hartebeeslaagte	AB	N
5750	MBUR00043	pulchella	SA	NC	HART	Farm Hartebeeslaagte	AB	N
5751	MBUR00045	pulchella	SA	NC	HART	Farm Hartebeeslaagte	AB	N
1491	MB20642	pulchella	SA	NC		Farm Kaboes	AB	N
1544	MB20707	pulchella	SA	NC		Farm Langberg	A	N
1505	MB20662	pulchella	SA	NC		Farm Loerkop	AB	N
5754	MBUR00580	pulchella	SA	EC		Farm Matjiesfontein	B	N
5374	MBUR00629	pulchella	SA	EC		Farm Matjiesfontein	B	Y
1660	MB20903	pulchella	SA	NC		Farm Rooidam	A	Y
1657	MB20905	lineoocellata	SA	NC		Farm Rooidam	A	Y
1678	MB20906	lineoocellata	SA	NC		Farm Rooidam	A	Y
1661	MB20904	pulchella	SA	NC		Farm Rooidam	A	N
4698	SVN059	pulchella	SA	NC		Garies	A	Y
4699	SVN060	pulchella	SA	NC		Garies	A	Y
4721	SVN082	pulchella	SA	NC		Garies	A	N
4722	SVN083	pulchella	SA	NC		Garies	A	N
3436	MH0989	lineoocellata	SA	WC		Geelboslaagte	B	N
4588	AJC641	pulchella	SA	NC		Geelvloer dam	A	N
712	KTH518	lineoocellata	SA	WC		Goedemoed	B	N
4786	SVN147	pulchella	SA	NC		Goraas se leegte	B	Y

INDEX	SAMPLE	SUB-SPP	COUNTRY	PROV	SITE	AREA	CLADE	GENOTYPED
4787	SVN148	pulchella	SA	NC		Goraas se leegte	B	Y
5482	MBUR01004	lineoocellata	SA	LP		Greater Kuduland	A	Y
908	EL030	pulchella	SA	WC		Groot Swartberg	C	Y
1343	MCZA38342	lineoocellata	NM			Grunau	A	Y
1344	MCZA38364	lineoocellata	NM			Grunau	A	Y
1210	AMB6862	lineoocellata	NM			Helmeringhausen	A	Y
1337	MCZA38271	lineoocellata	NM			Helmeringhausen	A	Y
4597	AJC720	lineoocellata	SA	NC		Kakamas	A	Y
3695	MH0200	lineoocellata	SA	WC		Keeromsberg climb	B	N
1237	AMB8393	lineoocellata	SA	LP		Kgama	D	Y
4576	AJC595	lineoocellata	SA	LP		Komonande Nature Reserve	D	Y
1960	ATKMPL01	lineoocellata	SA	EC		Kouga Mnts	C	Y
4286	WRB106	pulchella	SA	NC		Kwaggasfonteinlaagte	A	Y
2887	HB050	lineoocellata	SA	EC		Kwandwe	Unknown	N
222	SVN337	lineoocellata	SA	LP		Lapalala Game Reserve	D	Y
226	SVN341	lineoocellata	SA	LP		Lapalala Game Reserve	D	N
4827	SVN189	pulchella	SA	NC		Loxton	B	Y
709	KTH512	lineoocellata	SA	WC		Naudesberg	B	Y
729	KTH556	lineoocellata	SA	WC		Naudesberg	B	N
4845	SVN208	pulchella	SA	WC		Prince Albert	B	N
4846	SVN209	pulchella	SA	WC		Prince Albert	B	N
3079	MCZ38797	pulchella	SA	LP	PYLK	Pylkop 593MS	A	Y
3080	MCZ38798	pulchella	SA	LP	PYLK	Pylkop 593MS	A	Y
3082	MCZ38800	pulchella	SA	LP	PYLK	Pylkop 593MS	A	N
3083	MCZ38801	pulchella	SA	LP	PYLK	Pylkop 593MS	A	N

INDEX	SAMPLE	SUB-SPP	COUNTRY	PROV	SITE	AREA	CLADE	GENOTYPED
3084	MCZ38802	pulchella	SA	LP	PYLK	Pylkop 593MS	A	N
3085	MCZ38803	pulchella	SA	LP	PYLK	Pylkop 593MS	A	N
1190	AMB4579	lineoocellata	SA	NC		Richtersveld National Park	A	N
1973	ATDGPL01	lineoocellata	SA	WC		Riversonderende	B	N
3694	MH0706	lineoocellata	SA	WC		Riversonderende	B	N
4272	RSP003	lineoocellata	SA	NC	HRBK	Rooipoort Nature Reserve	A	N
4276	RSP026	lineoocellata	SA	NC	HRBK	Rooipoort Nature Reserve	A	N
-	RJD001	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD002	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD003	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD004	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD005	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD006	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD007	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD009	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD010	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD011	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD012	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD013	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD014	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD016	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD018	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD020	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD022	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD023	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N

INDEX	SAMPLE	SUB-SPP	COUNTRY	PROV	SITE	AREA	CLADE	GENOTYPED
-	RJD025	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD026	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD027	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD029	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD030	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD031	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD033	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD035	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD036	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD039	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD040	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD045	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD046	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD047	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD050	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD052	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD053	lineoocellata	SA	NC	RNR1	Rooipoort Nature Reserve	A	N
-	RJD062	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD063	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD064	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD077	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD080	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD082	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD086	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD095	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N

INDEX	SAMPLE	SUB-SPP	COUNTRY	PROV	SITE	AREA	CLADE	GENOTYPED
-	RJD096	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD100	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD101	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD102	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD103	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD104	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD105	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD107	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD108	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD112	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD118	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD119	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD126	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD127	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD130	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD131	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD132	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD134	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD139	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD142	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD144	lineoocellata	SA	NC	RNR2	Rooipoort Nature Reserve	A	N
-	RJD067	lineoocellata	SA	NC	RNR3	Rooipoort Nature Reserve	A	N
-	RJD089	lineoocellata	SA	NC	RNR3	Rooipoort Nature Reserve	A	N
4265	RSP129	lineoocellata	SA	NC	RNR3	Rooipoort Nature Reserve	A	N
4266	RSP130	lineoocellata	SA	NC	RNR3	Rooipoort Nature Reserve	A	N

INDEX	SAMPLE	SUB-SPP	COUNTRY	PROV	SITE	AREA	CLADE	GENOTYPED
4267	RSP131	lineoocellata	SA	NC	RNR3	Rooipoort Nature Reserve	A	N
-	RJD145	lineoocellata	SA	NC	RNR4	Rooipoort Nature Reserve	A	N
-	RJD146	lineoocellata	SA	NC	RNR4	Rooipoort Nature Reserve	A	N
-	RJD152	lineoocellata	SA	NC	RNR4	Rooipoort Nature Reserve	A	N
4277	RSP046	lineoocellata	SA	NC	WATK	Rooipoort Nature Reserve	A	Y
4278	RSP108	lineoocellata	SA	NC	WATK	Rooipoort Nature Reserve	A	Y
4279	RSP109	lineoocellata	SA	NC	WATK	Rooipoort Nature Reserve	A	Y
4280	RSP110	lineoocellata	SA	NC	WATK	Rooipoort Nature Reserve	A	Y
4281	RSP111	lineoocellata	SA	NC	WATK	Rooipoort Nature Reserve	A	Y
5752	MBUR00531	pulchella	SA	EC		Sterkstroom	Unknown	N
1298	ABA26mu	pulchella	NM			Suidkruis	A	Y
4764	SVN125	pulchella	SA	NC		Sutherland	B	Y
4765	SVN126	pulchella	SA	NC		Sutherland	B	Y
5414	MBUR00728	pulchella	SA	WC		Tierberg	B	Y
5756	MBUR00702	pulchella	SA	WC		Tierberg	B	Y
5921	RSP236	lineoocellata	SA	NC		Tswalu Nature Reserve	A	Y
5998	RSP291	lineoocellata	SA	NC		Tswalu Nature Reserve	A	Y
6048	RSP344	lineoocellata	SA	NC		Tswalu Nature Reserve	A	Y
6049	RSP345	lineoocellata	SA	NC		Tswalu Nature Reserve	A	Y
5949	RSP263	lineoocellata	SA	NC		Tswalu Nature Reserve	A	N
6342	RSP473	lineoocellata	SA	LP		Venetia Reserve	A	Y
6330	RSP478	lineoocellata	SA	LP		Venetia Reserve	A	Y
6471	RSP371	lineoocellata	SA	LP		Venetia Reserve	A	N
6476	RSP479	lineoocellata	SA	LP		Venetia Reserve	A	N

Information A 1: First page of the published paper describing the eleven microsatellite loci developed for *Pedioplanis lineoocellata*. Tolley, KA, RJ Daniels, KA Feldheim (2014). Characterisation of microsatellite markers in the Spotted Sand Lizard (*Pedioplanis lineoocellata*) shows low levels of inbreeding and moderate genetic diversity on small spatial scale. Afri J Herpetol. 63(2): 1-11.

UNISA | Taylor & Francis
Taylor & Francis Group

Characterisation of microsatellite markers in the Spotted Sand Lizard (*Pedioplanis lineoocellata*) shows low levels of inbreeding and moderate genetic diversity on a small spatial scale

Krystal A. Tolley[1,2]*, Ryan J. Daniels[1,2] &
Kevin A. Feldheim[3]

[1]*Applied Biodiversity Research Division, South African National Biodiversity Institute, Claremont, South Africa;*
[2]*Department of Botany and Zoology, University of Stellenbosch, Matieland, South Africa;* [3]*Pritzker Laboratory for Molecular Systematics and Evolution, The Field Museum, Chicago, IL, USA*

Abstract.—Population genetic methods can be useful for understanding spatial genetic patterns, gene flow and diversity. While genetic markers such as gene sequences are useful for understanding broad scale phylogeographic patterns, microsatellite markers allow for inferences within species and on smaller spatial scales. Such data can then be used to glean information on biological processes at the landscape level. Because portions of southern Africa are predicted to be heavily impacted by climatic changes forecast for the next century, a better understanding of the biota at the species level would be beneficial for advancing knowledge on the southern African system. Therefore, we characterised 11 microsatellite markers for an arid adapted lacertid lizard, *Pedioplanis lineoocellata*, in order to obtain a microsatellite library for future studies. Five of 11 loci were out of Hardy-Weinberg (H-W) equilibrium, showing a homozygote excess. Frequency-based analysis suggested this is due to the presence of null alleles. We further investigated this using a model-based Bayesian approach to simultaneously estimate inbreeding, null alleles and random amplification failures. In contrast to the frequency-based approach, this method indicated that both inbreeding and random amplification failures, not null alleles, are the most significant components of the model for nine of 11 loci. Overall, the inbreeding coefficient was low, but not zero, suggesting that low levels of inbreeding contribute to H-W deviations. Finally, we examined whether population bottlenecks could be a factor in deviations from H-W using multiple methods, but found no evidence to suggest that bottlenecks have played a role.

Key words.—Africa, inbreeding, Lacertidae, null alleles, population bottlenecks, population genetics, reptiles

INTRODUCTION

Southern Africa, the geographic region south of the Cunene and Zambezi Rivers, has the third most biologically rich lizard fauna worldwide, in part because of an arid adapted assemblage of lizards (Branch 1998). This region is predicted to experience large climatic shifts over the next century, and arid adapted lizards may be vulnerable to the forecasted environmental changes (Boko *et al.* 2007; Deutsch *et al.* 2008; Tolley *et al.* 2009). In the face of climatic shifts, species must either alter their distribution or adapt to new
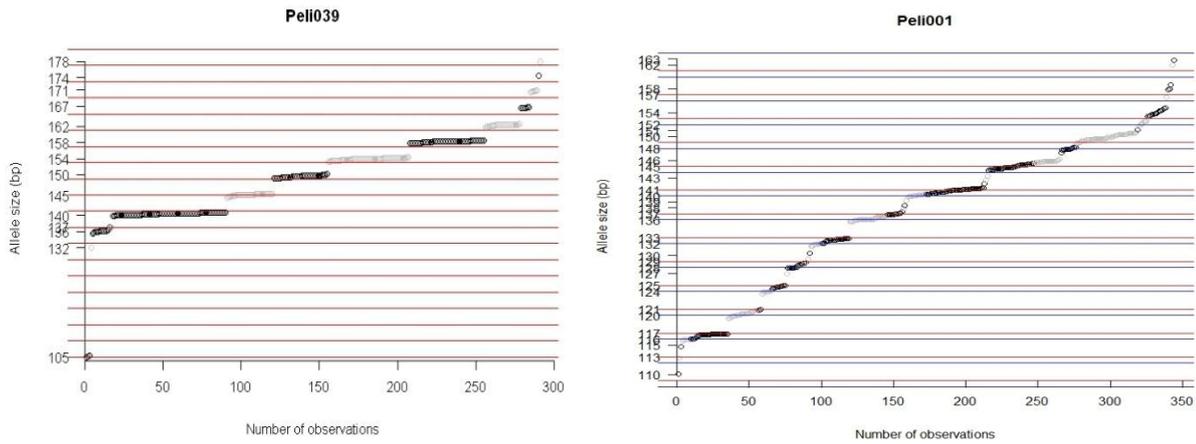
Figure A 1: Examples of loci amplified for *Pedioplanis lineoocellata* with (a) no allele call size inconsistencies from the expected repeat motif (Peli039) and (b) a large amount of allele size inconsistencies (Peli001). In both examples the red line indicated the repeat motif size sequence used. In (b) the blue line indicates another possible allele size sequence, though with fewer alleles conforming.

Table A 2: Error rate statistics for the amplification of microsatellite loci for all *Pedioplanis lineoocellata* samples. Loci specific number of errors (No. Errors) and proportion of erroneous allele calls (Prop. Error) shown according to the number of PCR repeated (No. runs), estimated from *ad hoc* repeat genotyping of random samples for all eleven microsatellite loci. Mean of the standard deviations (Mean S.D.) in fragment size per locus is shown.

Loci	No. Errors	No. runs	Prop. Error	Mean S.D.
Peli001	0	4	0.00	0.63
Peli005	2	12	0.17	0.27
Peli018	1	13	0.08	0.33
Peli020	4	14	0.29	0.26
Peli021	1	5	0.20	0.67
Peli022	3	33	0.09	0.09
Peli030	2	4	0.50	0.16
Peli034	1	22	0.05	0.08
Peli039	0	4	0.00	0.13
Peli048	0	12	0.00	0.07
Peli050	0	4	0.00	0.08
all	14	127	0.11	0.25

Table A 3: Tables of summary statistics of microsatellite data for *Pedioplanis lineoocellata* by sample site for the eight primary sites used in Chapter 2. Sample sizes (n) indicated beside the sample site name. The number of alleles (Na), difference between the lowest and highest allele size (Range), observed and expected heterozygosity (H_O and H_E) are indicated. M-ratio/ G-W statistic (G-W; values in **bold**<0.68) and Hardy-Weinberg equilibrium statistics (HW p-value & S.D.; p values in **bold**<0.05) shown. Mean and standard deviation (Mean & S.D.) were only calculated for polymorphic loci.

ELSF	5							GANS	16						
Locus	Na	Range	H_O	H_E	G-W	HW p-value	S.D.	Locus	Na	Range	H_O	H_E	G-W	HW p-value	S.D.
Peli001	6	5	0.60	0.89	1.00	0.150	0.0002	Peli001	6	11	0.88	0.82	**0.50**	0.890	0.0002
Peli005	5	18	0.40	0.84	**0.26**	**0.030**	0.0001	Peli005	7	7	0.63	0.82	0.88	**0.010**	0.0001
Peli018	4	3	0.60	0.78	1.00	0.690	0.0003	Peli018	5	5	0.50	0.68	0.83	0.260	0.0003
Peli020	2	1	0.20	0.56	1.00	0.360	0.0003	Peli020	3	2	0.50	0.48	1.00	1.000	0.0000
Peli022	6	7	0.80	0.89	0.75	0.620	0.0003	Peli022	8	10	0.94	0.87	0.73	0.980	0.0001
Peli034	5	9	0.75	0.86	**0.50**	0.660	0.0002	Peli034	8	12	0.56	0.81	**0.62**	**0.020**	0.0001
Peli039	3	3	0.40	0.38	0.75	1.000	0.0000	Peli039	6	13	0.56	0.74	**0.43**	**0.000**	0.0000
Peli048	6	17	0.80	0.89	**0.33**	0.620	0.0003	Peli048	10	19	0.73	0.80	**0.50**	**0.040**	0.0001
Peli050	3	4	1.00	0.83	**0.60**	1.000	0.0000	Peli050	4	25	0.56	0.54	**0.15**	**0.010**	0.0001
Mean	4.44	7.44	0.62	0.77	0.69			Mean	6.33	11.56	0.65	0.73	**0.63**		
S.D.	1.51	6.17	0.25	0.18	0.28			S.D.	2.18	7.04	0.16	0.14	0.26		

HART	19							PYLK	6						
Locus	Na	Range	H_O	H_E	G-W	HW p-value	S.D.	Locus	Na	Range	H_O	H_E	G-W	HW p-value	S.D.
Peli001	5	8	0.32	0.33	**0.56**	0.370	0.0002	Peli001	3	3	0.40	0.51	0.75	0.330	0.0003
Peli005	8	10	0.28	0.72	0.73	**0.000**	0.0000	Peli005	3	3	0.60	0.69	0.75	0.300	0.0003
Peli018	6	9	0.28	0.73	**0.60**	**0.000**	0.0000	Peli018	2	1	0.40	0.36	1.00	1.000	0.0000
Peli020	3	7	0.32	0.28	**0.38**	1.000	0.0000	Peli020	3	8	0.67	0.55	**0.33**	1.000	0.0000
Peli022	7	6	0.83	0.80	1.00	0.870	0.0002	Peli022	8	11	1.00	0.93	**0.67**	1.000	0.0000
Peli034	8	18	0.68	0.78	**0.42**	0.060	0.0002	Peli034	5	13	0.83	0.79	**0.36**	1.000	0.0000
Peli039	7	14	0.37	0.78	**0.47**	**0.000**	0.0000	Peli039	5	5	0.60	0.76	0.83	0.240	0.0003
Peli048	4	3	0.61	0.67	1.00	0.350	0.0003	Peli048	4	4	0.83	0.68	0.80	0.080	0.0002
Peli050	1	0			1.00			Peli050	1	0			1.00		
Mean	6.00	9.38	0.46	0.64	**0.64**			Mean	4.13	6.00	0.67	0.66	0.69		
S.D.	1.85	4.72	0.22	0.21	0.25			S.D.	1.89	4.24	0.21	0.18	0.23		

RNR1	35							RNR2	29						
	Locus	Na	Range	H_O	H_E	G-W	HW p-value	S.D.	Locus	Na	Range	H_O	H_E	G-W	HW p-value	S.D.
	Peli001	8	7	0.83	0.82	1.00	0.990	0.0001	Peli001	9	8	0.83	0.79	1.00	0.770	0.0002
	Peli005	7	11	0.65	0.65	**0.58**	0.590	0.0003	Peli005	12	18	0.62	0.63	**0.63**	0.320	0.0002
	Peli018	11	12	0.92	0.87	0.85	0.930	0.0001	Peli018	8	10	0.80	0.85	0.73	0.780	0.0002
	Peli020	8	9	0.53	0.74	0.80	**0.010**	0.0000	Peli020	6	5	0.38	0.77	1.00	**0.000**	0.0000
	Peli022	12	16	0.94	0.85	0.71	0.950	0.0001	Peli022	12	15	0.77	0.90	0.75	**0.040**	0.0001
	Peli034	14	25	0.80	0.88	**0.54**	0.060	0.0001	Peli034	12	17	0.52	0.84	**0.67**	**0.000**	0.0000
	Peli039	6	5	0.45	0.73	1.00	**0.000**	0.0000	Peli039	9	9	0.50	0.81	0.90	**0.000**	0.0000
	Peli048	8	8	0.66	0.82	0.89	0.150	0.0002	Peli048	10	21	0.63	0.81	**0.45**	**0.040**	0.0001
	Peli050	5	6	0.15	0.15	0.71	1.000	0.0000	Peli050	5	5	0.39	0.35	0.83	1.000	0.0000
	Mean	8.78	11.00	0.66	0.72	0.79			Mean	9.22	12.00	0.60	0.75	0.77		
	S.D.	2.95	6.25	0.25	0.23	0.17			S.D.	2.59	5.90	0.17	0.17	0.18		

RNR3	5							WATK	5						
Locus	Na	Range	H_O	H_E	G-W	HW p-value	S.D.	Locus	Na	Range	H_O	H_E	G-W	HW p-value	S.D.
Peli001	4	4	0.80	0.78	0.80	0.700	0.0003	Peli001	5	4	0.80	0.87	1.00	0.690	0.0003
Peli005	4	6	0.80	0.71	**0.57**	1.000	0.0000	Peli005	4	5	0.40	0.64	**0.67**	0.240	0.0002
Peli018	4	3	0.75	0.79	1.00	1.000	0.0000	Peli018	6	10	0.60	0.89	**0.55**	0.150	0.0002
Peli020	3	2	0.80	0.64	1.00	0.620	0.0003	Peli020	3	6	0.75	0.61	**0.43**	1.000	0.0000
Peli022	5	11	0.60	0.87	**0.42**	0.290	0.0003	Peli022	7	7	0.60	0.93	0.88	0.060	0.0001
Peli034	8	14	0.80	0.96	**0.53**	0.240	0.0002	Peli034	7	13	0.80	0.93	**0.50**	0.370	0.0003
Peli039	2	4	0.25	0.25	**0.40**	1.000	0.0000	Peli039	5	5	0.40	0.87	0.83	0.060	0.0001
Peli048	6	19	0.60	0.89	**0.30**	0.150	0.0002	Peli048	4	3	0.40	0.78	1.00	**0.040**	0.0001
Peli050	4	3	0.60	0.53	1.00	1.000	0.0000	Peli050	1	0				1.00	
Mean	4.44	7.33	0.67	0.71	**0.67**			Mean	5.13	6.63	0.59	0.81	0.73		
S.D.	1.74	5.96	0.18	0.22	0.28			S.D.	1.46	3.34	0.18	0.13	0.23		

Table A 4: Table of Heterozygote deficit and null alleles presence as indicated by Van Oosterhout frequencies for the eight *Pedioplanis lineoocellata* sample sites used in Chapter 2. Values in **bold** indicate significant H_E deficits (H_E D) or the identified presence of null alleles (Null). Note that significant heterozygote deficit did not necessarily mean that null allele frequencies were significant. There were insufficient samples to calculate null allele frequencies at ESLF and PYLK. "-"indicates calculation not possible due to monomorphic locus.

	ESLF	GANS		HART		PYLK	RNR1		RNR2		RNR3		WATK			
Locus	H_E D	H_E D	Null	H_E D	Null	H_E D	H_E D	Null	H_E D	Null	H_E D	Null	H_E D	Null	Mean frq.	S.D.
Peli001	no	no	-0.05	no	0.04	no	no	-0.01	no	-0.03	no	-0.11	no	-0.03		
Peli005	**p<0.05**	**p<0.05**	0.10	**p<0.001**	**0.29**	no	no	0.01	no	0.01	no	-0.13	no	0.13	**0.29**	
Peli018	no	no	0.10	**p<0.001**	**0.29**	no	no	-0.04	no	0.02	no	-0.06	no	0.13	**0.29**	
Peli020	no	no	-0.06	no	-0.17	no	**p<0.001**	**0.13**	**p<0.001**	**0.24**	no	-0.21	no	-0.43	**0.18**	**0.08**
Peli022	no	no	-0.05	no	-0.04	no	no	-0.07	**p<0.05**	0.06	no	0.11	no	0.14		
Peli034	no	**p<0.05**	**0.14**	no	0.06	no	no	0.04	**p<0.001**	**0.19**	no	0.04	no	0.02	**0.16**	**0.03**
Peli039	no	**p<0.01**	0.11	**p<0.001**	**0.25**	no	**p<0.001**	**0.19**	**p<0.001**	**0.18**	-	-0.13	**no**	**0.23**	0.21	0.04
Peli048	no	**p<0.05**	0.03	no	0.02	no	no	**0.09**	**p<0.05**	**0.10**	no	0.13	**p<0.05**	0.22	**0.10**	**0.01**
Peli050	no	**p<0.05**	-0.05	-	0.00	-	no	-0.08	no	-0.20	no	-0.32	-	0.00		
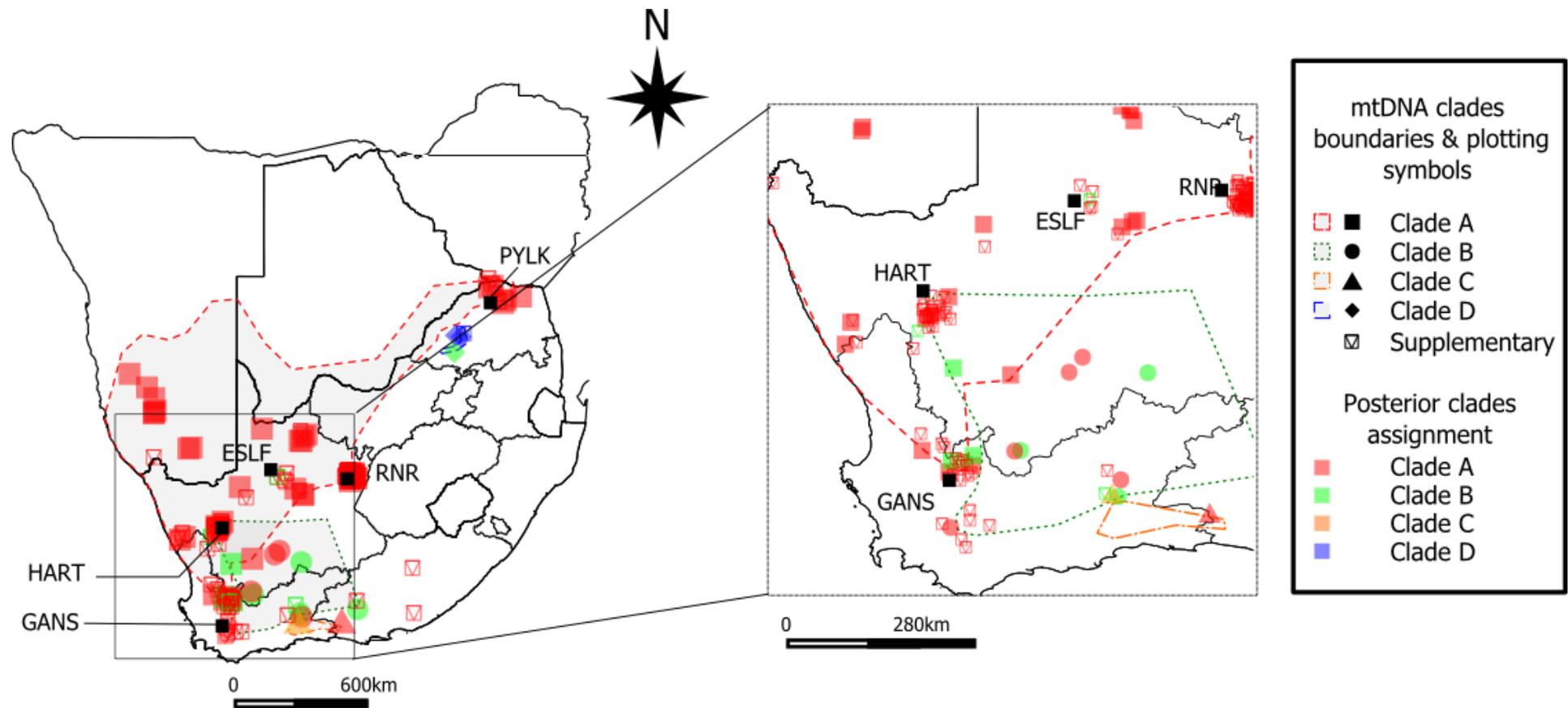
Figure A 2: Clade boundaries comparison between mtDNA delimited by Edwards (2013) and nuclear microsatellite data (this study) for *Pedioplanis lineoocellata*. Posterior clade assignments determined using DAPC analysis based on training data with *a priori* groups and including samples from Clade C and Clade D.
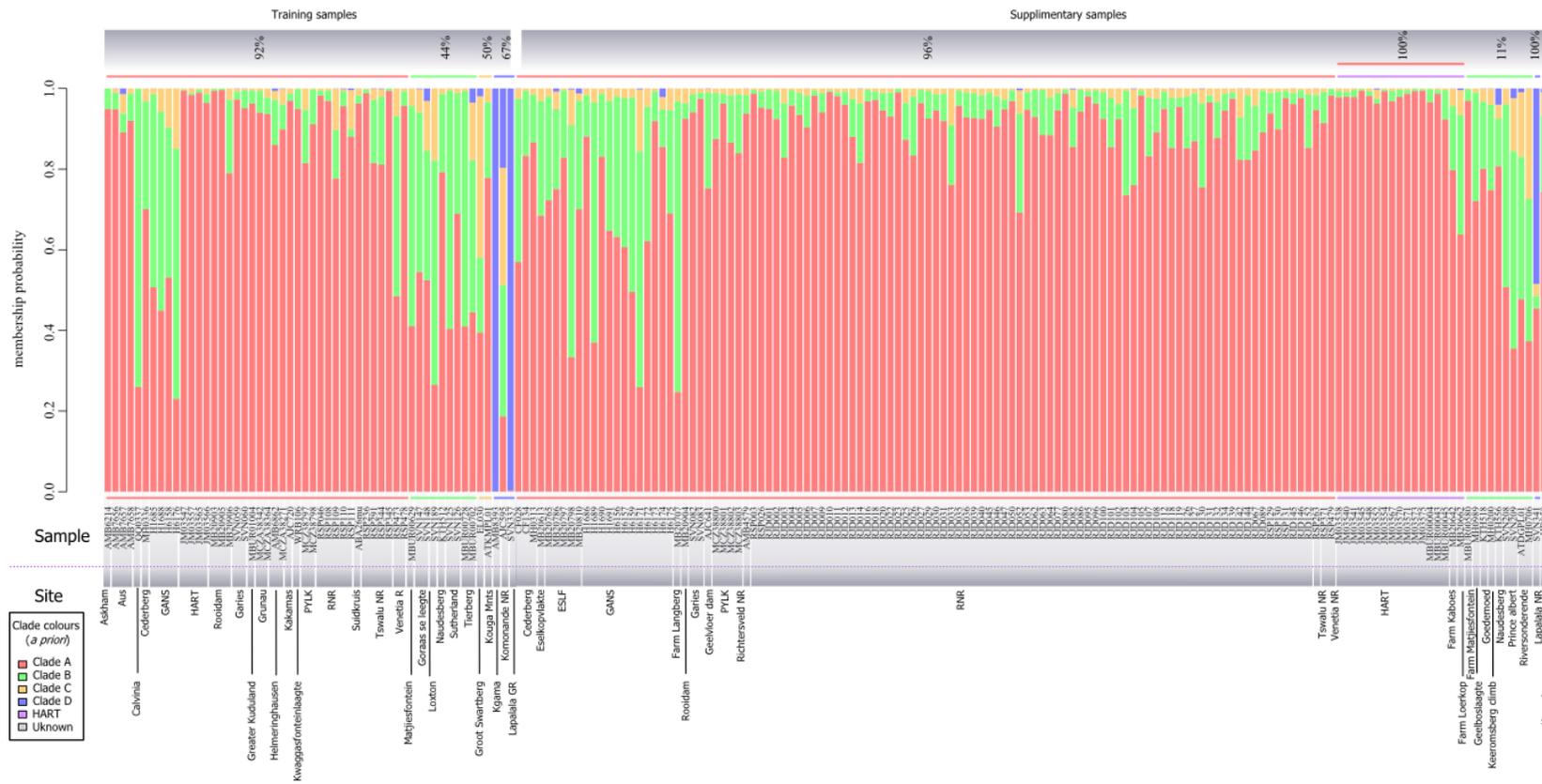
Figure A 3: Membership probabilities for individual *Pedioplanis lineoocellata,* in either training or supplementary data, to one of the four *a priori* clades. Analysis included all four mtDNA clades. Proportion of individuals successfully re-assigned to the *a priori* group indicated for each group above each group as a percentage. For supplementary data, *a priori* groups were based on geographical position with the exception of two samples that did not fall within any of the known clades- "Unknown". Coloured bars above the figure indicate the *a priori* clade to which samples belong. Sample site labels indicated below the sample names and the clade colours match those in the legend.

Information A 2: Results of the Discriminant analysis of principle components performed on *Pedioplanis lineoocellata* samples from within the distribution of mtDNA Clade A using prior group (sample site) assignments.

The Discriminant analysis of principle components (DAPC) analysis for sample sites within the distribution of mtDNA Clade A was performed with geographically defined sample sites used as prior groups to test the appropriateness of using sampling sites as training data to optimise PCs and DAs. Twelve PCs were retained for the nine training sites and eight discriminant functions were kept, including 53.4% of the total variation. Membership probabilities of training data indicate very strong support for sample site integrity with >90% correct re-assignment of all samples to their *a priori* sample site. However, sites within RNR were an exception because they had high admixture with other RNR sites (Figure A 5). Samples that were not from one of the eight *a priori* sample sites were used as supplementary and assigned to one of posterior clusters- GANS, RNR, ESLF, PYLK. The majority of the supplementary samples were assigned to the RNR cluster and covered the largest part of the distribution of Clade A, indicating that individuals in this cluster are widespread.

The high proportions of successful re-assignment of training data to prior groups supports the use of optimised DAs and PCs. However, groupings from the k-means analysis did not support RNR as a single population as it consisted of individuals from several clusters. The discrepancy may be the result of over fitting the data when performing the DAPC analysis without prior group assignments. Over fitting creates groups based on minor genetic variation. However, when repeating the k-means analysis with only 10 PCs (retaining 43% of the variation) the number of individuals with mixed assignment probabilities increased, especially within RNR, but RNR still consisted of individuals with different cluster assignments (results not shown). To account for the possible influence of sample size differences between RNR and the other sites, the DAPC analysis was re-run 20 times using only a subset of 15 randomly chosen individuals from RNR. Only minor differences were observed and all individuals with changes in posterior group assignment were from RNR and, to a lesser extent, PYLK (results not shown).

The results suggest that RNR may consist of several genetic clusters and the use of RNR as training data in the DAPC has led to the formation of an 'umbrella' group. The umbrella group forces several other potential genetic clusters under one identity and may hide useful detail.

Information A 3: Post-hoc analysis of population assignment probabilities performed on *Pedioplanis lineoocellata* samples from within the distribution of mtDNA Clade A.

Population structure was inferred across the distribution of mtDNA Clade A using the Bayesian model-based clustering method implemented in STRUCTURE v.2.3.3 (Pritchard *et al.* 2000). No prior assumptions regarding the best fit model were made, instead both "no admixture" and "admixture" models were run between K=1 and K=15. Four independent Markov chain Monte Carlo (MCMC) runs of 1 000 000 iterations and 10 000 burn-in each were used.

The 5-10 clusters best fit the data irrespective of which model was used however only the results of K=5 are presented below as it is the most parsimonious description of the data and furthermore it facilitates comparisons with the DAPC results. The composition of the five clusters formed match that of the DAPC analysis well. While results do not indicate that RNR consists of individuals from several different genetic clusters, but rather a single genetic cluster, it does show relatively greater genetic admixture than the remaining clusters and this may explain the results of the DAPC analysis. The overall conclusions drawn from the DAPC and STRUCTURE results are very similar. Neither of the results provide support for genetic admixture at HART (cluster 1) while both indicate that RNR may be an area with high genetic diversity and/or genetic admixture between divergent lineages, if not consisting of individuals from different lineages.
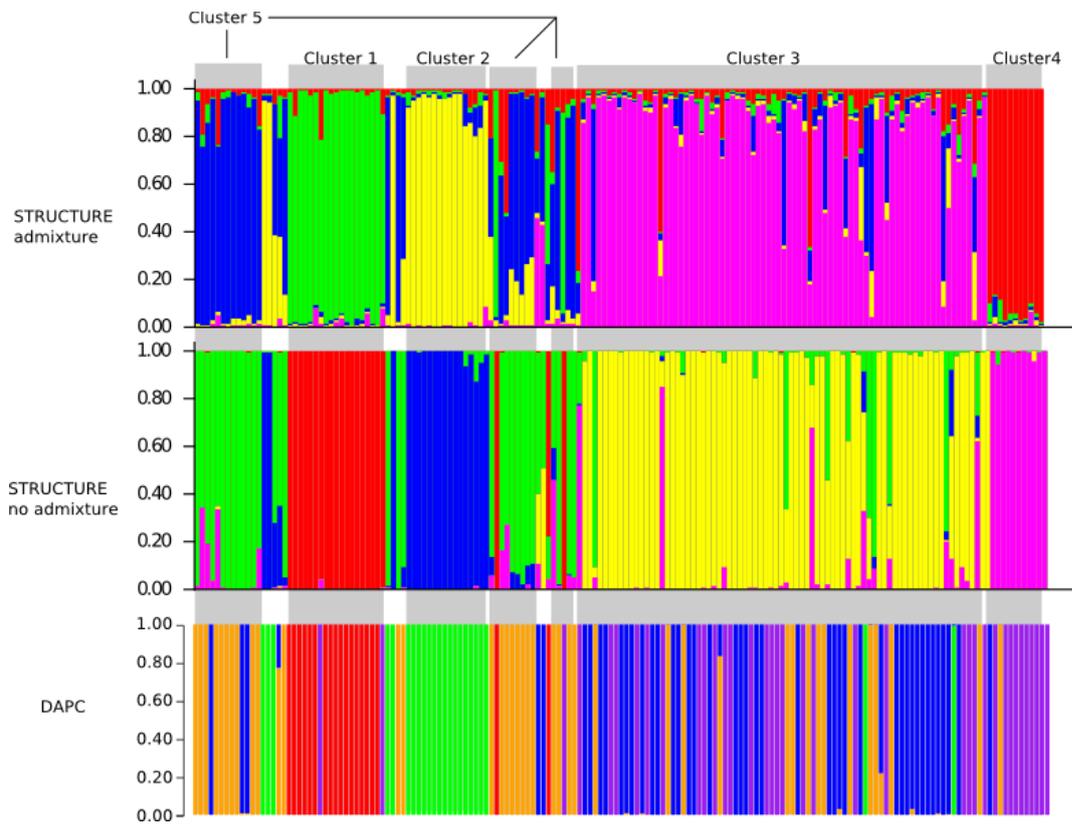
Figure A 4: Comparison of the cluster membership assignment of *Pedioplanis lineoocellata* individuals between the Discriminant analysis of principal components without prior group assignments and the two models available in STRUCTURE, "admixture" and "no admixture". Cluster numbering match that of the DAPC analysis for Population assignment within Clade A
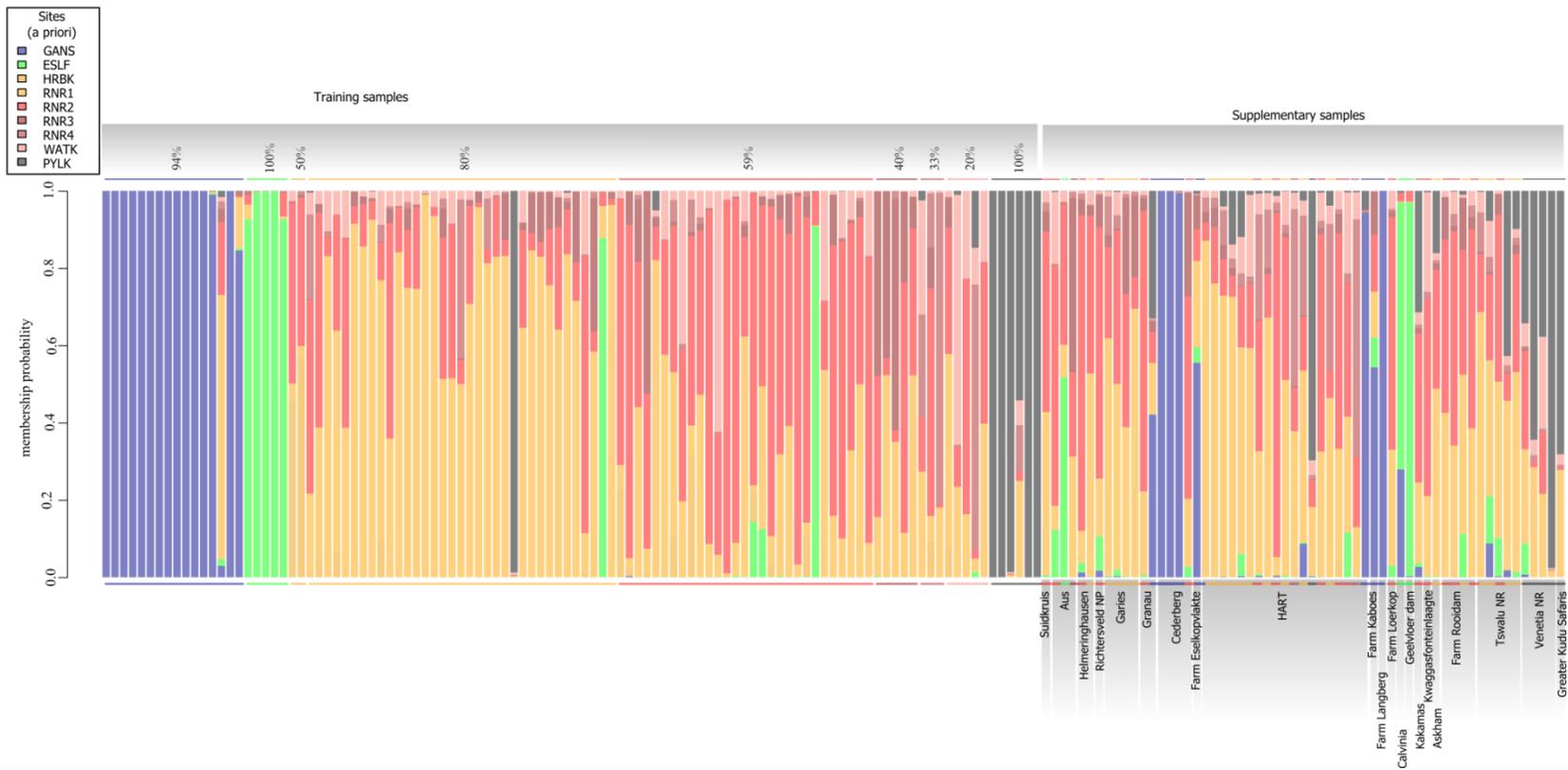
Figure A 5: Membership probabilities for individual *Pedioplanis lineoocellata* within Clade A, from either training or supplementary data, to one of the *a priori* sample sites. The proportion of individuals successfully re-assigned to the *a priori* group indicated for each group as a percentage above each group. Final sample site assignment indicated above each of the supplementary data. Sample site names provided below supplementary data. Sample site colours match those in the legend.

Appendix B DISPERSAL ESTIMATES AND SPATIAL AUTOCORRELATION INDICATE STRONG DISPERSAL IN THE SPOTTED SAND LIZARD (*PEDIOPLANIS LINEOOCELLATA*)

Table B 1: Capture-mark-recapture data for *Pedioplanis lineoocellata* from the April-May field session (RNR1) with each individuals unique tag number, Age, Snout-Vent length (SVL) and binary capture history.

Tag	Age	SVL (mm)	Capture History
C00	Juvenile	35.1	100000000000000000
C01	Juvenile	25.2	100000000000000000
C02	Juvenile	25.2	100000000000000000
C03	Juvenile	29.49	100000000000000000
C04	Juvenile	31.41	100000000000000000
C05	Juvenile	19.2	100000000000000000
C06	Juvenile	28.99	010001010000010011
C08	Juvenile	28.64	010000000001100100
C09	Juvenile	31.46	011111100000000000
C10	Juvenile	31.71	011010000110000001
C11	Juvenile	28.8	011110010000000000
C13	Juvenile	34.27	001100000000100000
C14	Juvenile	28.22	001000000000000000
C15	Juvenile	22.16	001000000000000000
C17	Juvenile	24.53	001000100100000000
C18	Juvenile	28.95	001000000010000000
C19	Juvenile	35.16	000100000100010000
C21	Juvenile	36.08	000100000000000000
C22	Juvenile	29.38	000100100010000001
C23	Juvenile	34.24	000100001100100000
C28	Juvenile	35.64	000010000010110011
C29	Juvenile	32.84	000010000000000000
C30	Juvenile	35.47	000010100000000001
C32	Juvenile	35.53	000000100000000001

Tag	Age	SVL (mm)	Capture History
C33	Juvenile	30.07	000000010000000000
C34	Juvenile	37.05	000000010100000000
C35	Juvenile	26.51	000000011000000001
C36	Juvenile	30.99	000000010010000001
C38	Juvenile	27.23	000000001001101100
C39	Juvenile	33.9	000000001000000001
C42	Juvenile	35.05	000000000010101001
C43	Juvenile	37.37	000000000001000000
C44	Juvenile	37.72	000000000001000000
C45	Juvenile	27.35	000000000000100000
C46	Juvenile	34.17	000000000000110111
C48	Juvenile	34.34	000000000000010000
C24	Juvenile	30.27	000100000000000100
C07	adult	52.22	010100000000000000
C12	adult	50.49	010110010000001000
C25	adult	52.96	000100010000000000
C26	adult	49.98	000100000000000000

Table B 2: Capture-mark-recapture data for *Pedioplanis lineoocellata* from the October-November field session (RNR2) with each individuals unique tag number, Age, Snout-Vent length (SVL) and binary capture history. Single juvenile individuals indicated with an *.

Tag	Sex	SVL (mm)	Capture History
D00	male	49.51	10000000000000
D02	male	54.48	10000000000000
D04	male	50.45	10001000100100
D07	male	49.18	01000000000000
D08	male	49.79	01000000000000
D10	male	53.11	01000000000000
D13	male	53.45	00100000000000
D16	male	51.55	00110000110000
D17	male	51.95	00100000000000

Tag	Sex	SVL (mm)	Capture History
D19	male	50.62	00010000010000
D20	male	46.72	00010000000000
D21	male	50.21	00010000000000
D23	male	55.23	00001000000000
D25	male	51.14	00001000000000
D26	male	49.08	00001000000000
D28	male	50.64	00000100000000
D29	male	51.77	00000100000000
D31*	male	44.33	00000100000000
D33	male	51.24	00000100000000
D37	male	47.15	00000010000000
D39	male	51.75	00000001000000
D42	male	54.89	00000001000000
D44	male	51.08	00000000100000
D45	male	51.54	00000000100000
D46	male	48.94	00000000100000
D48	male	52.32	00000000001000
D52	male	48.90	00000000000110
D53	male	48.05	00000000000100
D54	male	46.22	00000000000100
D01	female	51.58	10000000000000
D03	female	49.22	10000000000000
D06	female	49.24	01000000110010
D09	female	51.57	01000000000000
D11	female	51.37	01010000000000
D12	female	50.40	00100000001000
D14	female	48.70	00100100000000
D15	female	54.24	00100000000000
D18	female	50.12	00100000000000
D24	female	54.34	00001000000000

Tag	Sex	SVL (mm)	Capture History
D27	female	54.30	00000100000000
D30	female	60.40	00000111000000

Information B 1: Temporal and individual covariate details for *Pedioplanis lineoocellata* from capture-mark-recapture field work that were included in spatially explicit capture recapture analysis at RNR1 and RNR2.

Re-sighting probabilities may potentially vary with age and sex of the individual. Sex was included as an individual covariate. Age could not be used as a covariate because of the skew in the sampling (see Demographic D_E estimates), instead snout-vent length was included as a continuous individual covariate because SVL can be a good proxy for age in lizards (Sumner *et al.* 2001; Galán 2004). Re-sighting probabilities can also vary with temperature and weather on the day of survey. To account for temporal environmental variation several weather variables were included as covariates on capture probabilities g(0) (Appendix Table B 3). Daily temperature readings were taken from half hourly in-field ibutton readings (iButton©.com Wire® Pharmochron, 1108 AC2 DS1921 G#F50). The ibuttons were kept in transparent plastic tubes (3.5cm dia.) with ventilation holes in the sides. The tubes were placed in an open area at ground level but out of direct sunlight. Daily minimum, maximum and average daytime temperatures were determined from readings between 8am-7pm, which began an hour before and ended an hour after apparent periods of activity for *P. lineoocellata,* respectively. This allowed the recapture rates to be estimated as a function of daytime temperatures. A gross estimation of cloud cover at the study site at midday was included as a covariate. Rainfall events were considered in two ways; rainfall on the day was described as either "none", "drizzle", "thunderstorms" or "the day after" and was included as a discrete data covariate while precipitation was included as a continuous data covariate. The category "the day after" was included to account for possible increases in activity the day after a day of unfavourable weather. Average humidity, precipitation on the day and maximum wind speed data were retrieved from http://www.wunderground.com/history/airport/FAKM/2012/4/09/DailyHistory.html, which is data collected from the airport in the Kimberley area. To account for variation in survey effort, the number of hours spent surveying each day was included as a covariate on g(0).

Table B 3: Temporal covariate details from the (a) April/RNR1 and (b) October/RNR2 capture-mark-recapture field work that were included in the spatially explicit capture-recapture analysis. Occasion indicates the sampling occasion number for each sampling day. Explanation of abbreviations follow: Max temp, Min temp, Ave temp are maximum, minimum and average daytime temperatures during survey periods, respectively, recorded through ibuttons placed in the field in the open (see Information B 1 above). Max cloud was maximum % cloud cover estimated in field. Ave humid is the Average humidity. PPT, the precipitation on the day and Max wind, the maximum wind speed. Events is a simple summary of the weather events. [1] indicates in field measurements and [2] indicates data retrieved from http://www.wunderground.com/history/airport/FAKM/2012/4/11/DailyHistory.html for Kimberley. Minimum, maximum and mean values of each variable only calculated over days surveyed (Survey=1).

(a) April-May field session

Date	Occasion	Survey	Number of hours	Max temp (°C)[1]	Min temp (°C)[1]	Ave temp (°C)[1]	Max cloud (%)[1]	Ave humid (%)[2]	PPT (mm)[2]	Max wind (Km/hr.)[2]	Events[1]
11/04	1	1	8	29.50	19.00	21.80	0	45	0.0	32	none
12/04	2	1	8	43.00	13.50	31.23	0	29	0.0	13	none
13/04	3	1	8	44.50	16.50	32.88	20	39	0.0	7	none
14/04	4	1	8	43.00	16.50	31.00	40	40	0.0	33	none
15/04	5	1	5	43.50	17.00	32.13	40	36	0.0	18	none
16/04	6	1	5.5	45.50	16.00	28.06	70	51	0.2	22	rain
17/04	7	0	0	32.00	12.50	21.75	100	62	0.4	26	thunderstorm
18/04	8	1	8	32.50	13.50	22.76	40	50	0.0	24	day after
19/04	9	1	8	35.00	15.00	24.64	40	46	0.0	13	none
20/04	10	1	8	36.50	15.50	25.94	60	26	0.0	18	none
21/04	11	0	8	23.00	16.00	19.49	100	63	1.0	18	rain

Date	Occasion	Survey	Number of hours	Max temp (°C)[1]	Min temp (°C)[1]	Ave temp (°C)[1]	Max cloud (%)[1]	Ave humid (%)[2]	PPT (mm)[2]	Max wind (Km/hr.)[2]	Events[1]
22/04	12	1	6.25	36.00	12.00	22.13	100	61	0.0	26	thunderstorm
23/04	13	1	8	28.00	11.00	19.36	0	53	0.0	17	day after
24/04	14	1	8	30.50	9.50	21.06	80	52	0.0	18	none
25/04	15	0	0	22.00	14.00	17.31	100	71	4.0	15	thunderstorm
26/04	16	0	0	22.50	12.50	15.90	100	84	4.0	33	thunderstorm
27/04	17	1	8	29.50	11.50	19.74	60	67	0.0	13	day after
28/04	18	1	8	30.50	10.50	20.49	70	58	0.0	17	none
29/04	19	1	8	33.00	11.00	22.36	10	47	0.0	11	none
30/04	20	0	8	37.50	14.00	26.46	10	25	0.0	30	none
01/05	21	1	6	31.00	15.50	21.84	70	41	0.0	17	none
02/05	22	1	8	35.00	12.50	24.98	80	39	0.0	30	none
03/05	23	1	8	42.00	18.00	29.13	1	28	0.0	35	none
Minimum			5.0	28.00	9.50	19.36	0.00	26.0	0.0	7.00	
Maximum			8.0	45.50	19.00	32.88	100	67.0	0.20	35.00	
Mean±S.D.			7.49±1.0	36.03±6.0	14.11±2.8	25.09±4.5	43.39±32.3	44.89±11.3	0.10±0.1	20.22±8.2	

(b) October-November field session

Date	Occasion	Survey	Number of hours	Max temp (°C)[1]	Min temp (°C)[1]	Ave temp (°C)[1]	Max cloud (%)[1]	Ave Humid (%)[2]	PPT (mm)[2]	Max wind (km/hr.)[2]	Events[1]
14/10	1	1	7.5	32	14	23	60	26	0	43	none
15/10	2	0	0	22	12	17	100	72	11	70	rain
16/10	3	0	6	28	12	20	80	58	0	28	rain
17/10	4	1	8.5	41.0	20.5	33.2	40	40	0	26	day after
18/10	5	1	7	38.5	18.5	31.0	0	43	0	24	none
19/10	6	1	10	41.5	13.0	29.9	20	31	0	39	none
20/10	7	0	3.5	37.5	14.0	26.2	0	21	0	30	none
21/10	8	1	6	43.0	16.5	30.3	0	28	0	44	none
22/10	9	0	0	37.5	17.5	31.0	-	31	0	28	rain
23/10	10	1	7.75	46.0	17.5	33.5	0	36	0	32	day after
24/10	11	1	12	47.0	21.0	35.4	40	47	0	35	none
25/10	12	0	0	32.0	15.0	22.0	100	75	12	41	thunderstorm
26/10	13	1	8.5	47.5	19.0	33.1	0	40	0.2	32	day after
27/10	14	1	5.5	46.0	19.0	34.2	0	44	0	30	none
28/10	15	1	7	49.5	19.5	37.3	5	41	0	30	none

Date	Occasion	Survey	Number of hours	Max temp (°C)[1]	Min temp (°C)[1]	Ave temp (°C)[1]	Max cloud (%)[1]	Ave Humid (%)[2]	PPT (mm)[2]	Max wind (km/hr.)[2]	Events[1]
29/10	16	1	4.5	52.0	24.5	39.9	0	26	0	39	none
30/10	17	1	6.5	39.5	19.5	28.6	0	35	0	44	none
31/10	18	1	8	40.0	15.5	29.4	0	34	0	41	none
01/11	19	0	0	41.0	13.5	29.4	0	36	0	28	none
02/11	20	1	5.5	43.5	24.0	34.9	60	32	0	33	none
Minimum			4.5	32.0	13.0	23.0	0.0	26.0	0.0	24.0	
Maximum			12.0	52.0	24.5	39.9	60.0	47.0	0.2	44.0	
Mean±S.D.			7.4±1.9	43.4±5.0	18.7±3.2	32.4±4.0	16.1±22.7	35.9±6.6	0.0±0.1	35.1±6.4	