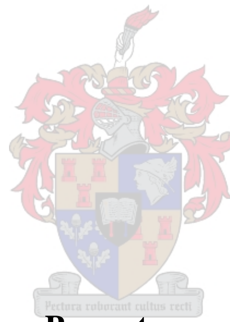
Exploring mechanisms that shape Siphonaptera composition and distribution patterns on small mammals across South Africa

Luther van der Mescht

Dissertation presented for the degree of Doctor of Science in Conservation Ecology in the Faculty of AgriSciences at Stellenbosch University



Promoters:

Dr Sonja Matthee

Prof Conrad A. Matthee

December 2015

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Stellenbosch University https://scholar.sun.ac.za

Declaration

By submitting this thesis/dissertation 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.

Some of the contents contained in this thesis (Chapters 2-4) are taken directly from manuscripts submitted or drafted for publication in the primary scientific literature. This resulted in some overlap in content between the chapters.

December 2015

Copyright © 2015 Stellenbosch University

All rights reserved

Stellenbosch University https://scholar.sun.ac.za

Abstract

Fleas (Siphonaptera) are obligate ectoparasites of mammals and birds. Due to their economic importance as disease vectors, most contemporary studies on macroparasites now also consider the spatial variation of parasite communities and the underlying mechanisms involved in shaping current distribution patterns. Fleas differ in life history traits, such as the level of host specificity and microhabitat preferences, which can result in differential evolutionary responses to similar abiotic events. The main objectives of this study are to investigate: (1) the influence of vicariance and host association on the genetic structure of two generalist flea species, *Listropsylla agrippinae*, and *Chiastopsylla rossi*. The taxa differ in the time spent on the host (fur vs. nest) and the level of host specificity; (2) the taxonomy of *Dinopsyllus ellobius* to determine whether more than one species exist and if so to then elucidate the mechanisms of speciation; and (3) the influence of flea life history on species distribution model performance and see if the relative importance of predictor variables differ between species with different life histories. A total of 1423 small mammals were brushed to collect 2906 flea individuals originating from 31 geographically distinct localities throughout South Africa (SA). Phylogeographic structure of *L. agrippinae* and *C. rossi* were determined by making use of 315 mitochondrial COII (mtDNA) and 174 nuclear EF1-α (nDNA) sequences. The more host specific fur flea, *L. agrippinae*, displayed pronounced spatial phylogeographic structure, based on mtDNA, which was congruent with host vicariance in the region. In contrast, the more generalist nest flea, *C. rossi*, showed a higher level of inter-populational divergence, based on mtDNA and nDNA, and this may be attributed to comparatively higher restrictions to dispersal when compared to the more specific fur flea. In an attempt to resolve the taxonomy of *D. ellobius,* 151 mtDNA and 68 nDNA alleles were generated from individuals meeting the morphological description of *D. ellobius*. Two distinct *D. ellobius* lineages that corresponded to previously described species (*D. ellobius* and *D. abaris*) were recorded. The results indicate that the two species indeed differ morphologically and based on the distribution of the two species it was concluded that the diversification could be a result of climate driven vicariance and subsequent ecological segregation according to habitat use. Locality records from Segerman (1995) were digitized and used as background data in species distribution modelling. Sufficient information was obtained for 21 flea species. A total of 68 climatic and landscape feature predictor variables were obtained and through a process of elimination, 19 variables were ultimately used. Model performance was good to excellent on average and the contribution of

Stellenbosch University https://scholar.sun.ac.za

climate and landscape feature variables differed between fleas with different life histories. Historical and contemporary climate has the most prominent effect on flea distribution at the regional scale, but the level of host association influences the phylogeographic pattern of fleas. This study provides the first evidence of congruent phylogeographic patterns between a generalist temporary parasite and its hosts. Our findings provide further support for the notion that more than one species exist within the *D. ellobius* complex and that speciation is a result of complex interactions. The study also provides novel data on the role of environmental variables in shaping the geographic distribution of flea species with different life histories. With the anticipated rise in flea-borne diseases worldwide, due to changes in vector distribution, the study further emphasize the need for studying the mechanisms involved in shaping flea distribution patterns.

Stellenbosch University https://scholar.sun.ac.za

Opsomming

Vlooie (Siphonaptera) is verpligte ektoparasiete van soogdiere en voëls. As gevolg van hul ekonomiese belangrikheid as vektore van oordraagbare siektes, neem meeste kontemporêre studies oor makroparasiete nou die ruimtelike variasie (vikariansie) van parasiet gemeenskappe en die onderliggende meganismes betrokke by die vorming van die huidige verspreiding patrone in ag. Vlooie verskil in lewensgeskiedenis eienskappe, soos die graad van gasheer spesifisiteit en mikrohabitat voorkeur, wat kan lei tot differensiële evolusionêre patrone tot soortgelyke abiotiese gebeure. Die belangrikste doelwitte van hierdie studie is om ondersoek in te stel na: (1) die invloed van vikariansie en gasheer assosiasie op die genetiese struktuur van twee algemene vlooi spesies, *Listropsylla agrippinae* en *Chiastopsylla rossi*. Die taxa verskil in die hoeveelheid tyd wat gespandeer word op die gasheer (pels vs. nes) en die graad van gasheer spesifisiteit; (2) die taksonomie van *Dinopsyllus ellobius* om te bepaal of daar meer as een spesie bestaan, en indien wel, die meganismes van spesiasie voortbring; en (3) die invloed van vlooi lewensgeskiedenis op die prestasie van die spesies verspreidings model en kyk of die relatiewe belangrikheid van omgewingsvoorspellers verskil tussen spesies met verskillende lewensgeskiedenisse. 'n Totaal van 1423 klein soogdiere is geborsel om 2906 vlooi individue afkomstig van 31 geografies verskillende plekke in Suid-Afrika in te samel. Filogeografiese struktuur van *L. agrippinae* en *C. rossi* is bepaal deur gebruik te maak van 315 mitochondriale COII (mtDNA) en 174 kern EF1-α (nDNA) volgorde bepalings. Die meer gasheer spesifieke pels vlooi, *L. agrippinae*, het ruimtelike filogeografiese struktuur aangetoon, gebaseer op mtDNA, wat ooreenstem met gasheer vikariansie in die streek. In teenstelling hiermee het die minder gasheer spesifieke nes vlooi, *C. rossi*, 'n hoër vlak van inter-populasie divergensie getoon, gebaseer op mtDNA en nDNA, wat toegeskryf kan word aan relatief hoër beperkings op verspreiding teenoor die meer gasheer spesifieke pels vlooi. In 'n poging om die taksonomie van *D. ellobius* op te los is 151 mtDNA en 68 nDNA allele gegenereer uit individue wat voldoen aan die morfologiese beskrywing van *D. ellobius*. Twee afsonderlike *D. ellobius* afstammelinge wat ooreengestem het met voorheen beskryfde spesies (*D. ellobius* en *D. abaris*) is aangeteken. Die resultate dui daarop dat die twee spesies inderdaad morfologies verskil en gebaseer op die verspreiding van die twee spesies is daar tot die gevolgtrekking gekom dat die diversifikasie moontlik toegeskryf kan word aan klimaat gedryfde vikariansie en daaropvolgende ekologiese segregasie volgens habitat gebruik. Geografiese verspreidings rekords van Segerman (1995) was gedigitaliseerd en gebruik as agtergrond data in die spesie verspreidings modelle. 'n

Stellenbosch University https://scholar.sun.ac.za

Totaal van 68 klimaat en landskap kenmerk voorspeller veranderlikes is verkry, en deur 'n proses van eliminasie, is 19 veranderlikes gebruik. Model prestasie was oor die algemeen goed tot uitstekend en die bydrae van klimaat en landskaps veranderlikes het verskil tussen vlooie met verskille in lewensgeskiedenis. Historiese en kontemporêre klimaat het die mees prominente effek op vlooi verspreiding by die plaaslike skaal, maar die vlak van die gasheer assosiasie beïnvloed die filogeografiese patroon van vlooie. Hierdie studie bied die eerste getuienis van ooreenstemmende filogeografiese patrone tussen 'n minder gasheer spesifieke, tydelike parasiet en sy gashere. Ons bevindinge bied verdere ondersteuning vir die idee dat meer as een spesie bestaan binne die *D. ellobius* kompleks en dat spesiasie 'n gevolg is van komplekse interaksies. Die studie bied ook nuwe data oor die bydrae van die omgewingsveranderlikes in die vorming van die geografiese verspreiding van vlooi spesies met verskillende lewensgeskiedenisse. Met die verwagte styging in vlooi-oordraagbare siektes wêreldwyd, as gevolg van veranderinge in vektor verspreiding, beklemtoon die studie verder die noodsaaklikheid vir die bestudering van die meganismes betrokke in die vorming van vlooi verspreiding patrone.

Stellenbosch University https://scholar.sun.ac.za

Acknowledgements

First, I would like to thank my supervisors for guiding, supporting and mentoring me throughout my PhD studies.

Private landowners and provincial nature conservation agencies are thanked for granting permission to trap on their property or in their reserves with the following permit numbers (Western Cape, 0035-AAA007-00423; Northern Cape, FAUNA 1076/2011; Eastern Cape, CRO37/11CR; KZN wildlife, OP4990/2010 and Gauteng, CPF 6-0153). I am grateful to Vernon Steyn, Nina Du Toit and Andrea Spickett for providing additional samples. Adriaan Engelbrecht, Karlien Malan and Jannie Groenewald are thanked for assistance in the field. I would also like to thank Michael W. Hastriter for useful discussions on flea morphology and Jean-Bernard Duchemin for providing us with some reference literature.

Funding was provided by Stellenbosch University, the National Research Foundation (NRF) South Africa (fellowship to myself) and a Department of Science and Technology (DST) Grant to Prof. Conrad A. Matthee (IBIP-BS13100752708) and NRF Grants to Dr. Sonja Matthee (GUN85718 and GUN74463). The Grant holder acknowledges that opinions, findings and conclusions or recommendations expressed in any publication generated by the NRF-supported research are those of the authors, and that the NRF accepts no liability whatsoever in this regard.

Finally, I would like to thank my family and friends for supporting me during my studies.

Stellenbosch University https://scholar.sun.ac.za

Table of Contents

Declaration……………………………………..…………………………………..………ii

Abstract…………………………………………..……………………………………..iii

Opsomming………………………………………………..……………………………v

Acknowledgements.……………………………………………...…………..……...vii

Table of Contents….…………………………………..………………………….....viii

List of Figures……………………...………………………………………......…..…..ix

List of Tables…………………………………………………………………..…..…xiii

Chapter 1: General Introduction…………………………………………………………..1

Chapter 2: Comparative phylogeography between two generalist flea species reveal a complex interaction between parasite life history and host vicariance: Parasite-host association matters……………...…………………………………………………………............14

Chapter 3: A genetic perspective on the taxonomy and evolution of the medically important flea, *Dinopsyllus ellobius* (Siphonaptera: Dinopsyllinae), and the resurrection of *Dinopsyllus abaris*.……………..…………….……………………………………………………….39

Chapter 4: The influence of life history characteristics on flea (Siphonaptera) species distribution models…………...………...63

Conclusions……………………………………………………………….……..…..81

References………………………………………………………………………......84

Appendix A: Supplementary Figures……………………………………………..…..…113

Appendix B: Supplementary Tables…………………….……………...……………..…..118

Stellenbosch University https://scholar.sun.ac.za

List of figures

Figure 2.1 Map indicating **a)** localities sampled with their respective codes indicated as in Table 3.1. The vicariant breaks in South Africa (1) winter and aseasonal rainfall break and (2) xeric and mesic biomes are indicated by white dashed lines, **b)** the mapped distribution of *L. agrippinae* and **c)** *C. rossi* are taken directly from Segerman (1995) **17**

Figure 2.2 The number of individuals sequenced and prevalence of **a)** *L. agrippinae* and **b)** *C. rossi* sampled from different small mammal host species at 20 localities throughout South Africa **23**

Figure 2.3 a) A map of geographical distribution of sampling localities in South Africa for *L. agrippinae*. Bayesian analysis of population structure (BAPS) subclusters are indicated by green, blue and red circles. **b)** Statistical parsimony mitochondrial COII haplotype network colour coded according to sampling locality. Circle size depicts frequency, branches depict single mutational steps, small black circles display intersections and cross hatching indicates missing haplotypes/mutational steps. Clades are bound by black (L1) and deep yellow (L2) line boxes. **c)** Neighbour-Net phylogenetic network for *L. agrippinae* labelled according to haplotype groupings from statistical parsimony results and are bound by black (L1) and deep yellow (L2) line boxes. Bayesian analysis of population structure (BAPS) subclusters are indicated by green, blue and red lines **25**

Figure 2.4 a) A map of geographical distribution of sampling localities in South Africa for *L. agrippinae*. **b)** Statistical parsimony nuclear EF1-α haplotype network colour coded according to sampling locality. Circle size depicts frequency, branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes/mutational steps. **c)** Neighbour-Net phylogenetic network for *L. agrippinae* **27**

Figure 2.5 Genetic landscape interpolation plot for total distribution of **a**) *L. agrippinae*, and **b**) *C. rossi* **28**

Stellenbosch University https://scholar.sun.ac.za

Figure 2.6 a) A map of geographical distribution of sampling localities in South Africa for *C. rossi*. Bayesian analysis of population structure (BAPS) revealed six subclusters indicated by different colours. **b)** Statistical parsimony mitochondrial COII haplotype network colour coded according to sampling locality. Circle size depicts frequency, branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes/mutational steps. Clusters are bound by black (C1) and deep yellow (C2) line boxes. **c)** Neighbour-Net phylogenetic network for *C. rossi* labelled according to haplotype groupings from statistical parsimony results and are bound by black (C1) and deep yellow (C2) line boxes **31**

Figure 2.7 a) A map of geographical distribution of sampling localities in South Africa for *C. rossi*. **b)** Statistical parsimony nuclear EF1-α haplotype network colour coded according to sampling locality. Circle size depicts frequency, branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes/mutational steps. **c)** Neighbour-Net phylogenetic network for *C. rossi* **33**

Figure 3.1 Localities from which small mammal hosts were sampled with locality codes as in Table B.3. (Insert: *D. ellobius* distribution from Segerman 1995) **43**

Figure 3.2 a) Geographical distribution of sampling localities in South Africa for *D. ellobius* plotted on a map of mean rainfall. Colour spectrum of the map vary from low values in red to high values in blue. (Insert: Map indicating where two lineages were found) **b)** Neighbour-Net phylogenetic network of mitochondrial COII for *D. ellobius* bound by red (D1 (*D. abaris*)) and deep yellow (D2 (*D. ellobius*)) line boxes. **c)** Statistical parsimony mitochondrial COII haplotype network colour coded according to sampling locality to represent relative frequencies. Circle size depicts frequency; branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes. The two lineages are bound by red (D1 (*D. abaris*)) and deep yellow (D2 (*D. ellobius*)) line boxes **50**

Stellenbosch University https://scholar.sun.ac.za

Figure 3.3 a) Geographical distribution of sampling localities in South Africa for *D. ellobius* plotted on a map of mean rainfall. Colour spectrum of the map vary from low values in red to high values in blue. (Insert: Map indicating where two lineages were found) **b)** Neighbour-Net phylogenetic network of subset nuclear EF1-α for *D. ellobius* with nodes coloured according to lineage. **c)** Statistical parsimony subset nuclear EF1-α haplotype network colour coded according to sampling locality to represent relative frequencies. Circle size depicts frequency; branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes **53**

Figure 3.4 Small mammal host species and proportion of host species from which *D. ellobius* lineages, **a)** D1 (*D. abaris*) and **b)** D2 (*D. ellobius*) were sequenced **54**

Figure 3.5 Boxplots for male individuals of mean log transformed length (μm) of **a)** head, **b)** coxa, **c)** femur, and **d)** tibia. ANOVA results are indicated in the bottom right corner of each boxplot. ***significant values **55**

Figure 3.6 Principal component analysis of morphological measurements recorded for **a)** male *D. ellobius* lineages. D1 (*D. abaris*, red) (N = 14) and D2 (*D. ellobius*, deep yellow) (N = 38) and **b)** female *D. ellobius* lineages. D1, *D. abaris*, red) (N = 24) and D2, *D. ellobius*, deep yellow) (N = 54) **56**

Figure 3.7 Mean femur length (μm) of male individuals at each locality. The two lineages are indicated by different colours (D1, *D. abaris*, red; D2, *D. ellobius*, deep yellow). *No males individuals were found at FB for D2 (*D. ellobius*, deep yellow) **57**

Figure 4.1 NMDS ordination plot showing the relationship between microhabitat preference of flea species and the variable importance of predictors included in each species distribution model. The best linear fit of all variables that had a significant ($p \leq 0.05$) influence are indicated. Variable codes: DT6, 6th harmonic component of daytime land surface temperature; DTMIN, minimum daytime land surface temperature; RFMIN, minimum rainfall; RFMAX, maximum rainfall; WV6, 6th harmonic regression component of water vapour **72**

Stellenbosch University https://scholar.sun.ac.za

Figure 4.2 Variable importance (i.e. percent relative predictor variable type importance) in MaxEnt models, averaged across flea species for each variable type based on microhabitat preference (see Table 4.2 for variable category information). Significant differences in the contribution of predictor variable types between the two categories of species are indicated by asterisks: *** 0.001, ** 0.01, * 0.05 **73**

Figure 4.3 NMDS ordination plot showing the relationship between host specificity of flea species and the variable importance of predictors included in each species distribution model. The best linear fit of all variables that had a significant ($p \leq 0.05$) influence are indicated. Variable codes: DT6, 6th harmonic component of daytime land surface temperature; DT7, 7th harmonic component of daytime land surface temperature; DTMIN, minimum daytime land surface temperature; NDVI7, 7th harmonic component of NDVI; RFMIN, minimum rainfall; RFMAX, maximum rainfall; WV6, 6th harmonic regression component of water vapour; WVMIN, minimum water vapour; SPS, soil percentage sand **75**

Figure 4.4 Variable importance (i.e. percent relative predictor variable type importance) in MaxEnt models, averaged across flea species for each variable type based on host specificity (see Table 4.2 for variable category information). Significant differences in the contribution of predictor variable types between the two categories of species are indicated by asterisks: *** 0.001, ** 0.01, * 0.05 **76**

Stellenbosch University https://scholar.sun.ac.za

List of tables

Table 2.1 Locality information from where specimens were obtained for each of the two flea species, with codes corresponding to Fig. 2.1 **20**

Table 2.2 Summary statistics for the mitochondrial DNA (COII) and nuclear intron (EF1-α) markers sequenced from *L. agrippinae* and *C. rossi.* The number of sequences/alleles, total number of haplotypes, total number of singleton haplotypes, fragment length (bp), number of polymorphic sites (*P*), nucleotide diversity (π) and haplotype diversity (*h*) is indicated **24**

Table 2.3 Hierarchical analysis of molecular variance (AMOVA) of mitochondrial DNA (COII) and nuclear intron (EF1-α) in two flea species for BAPS subclusters. Fixation index values are given for the hierarchical levels examined: F_{ST}, within localities; F_{CT}, among subclusters; and $F_{SC,}$ among localities within subclusters. Significant (p < 0.05) tests are indicated in bold **29**

Table 2.4 Pairwise Φst values among *L. agrippinae* sampled localities for mitochondrial DNA (COII) below and nuclear intron (EF1-α) above the diagonal. Significant values (p < 0.05) are highlighted in bold. (Locality codes as in Table 2.1) **29**

Table 2.5 Pairwise Φst values among *C. rossi* sampled localities for mitochondrial DNA (COII) below and nuclear intron (EF1-α) above the diagonal. Significant values (p < 0.05) are highlighted in bold. (Locality codes as in Table 2.1) **34**

Table 3.1 Geo-referenced localities with codes and number of fleas sequenced for each gene from which *D. ellobius* were collected **48**

Table 3.2 Summary statistics for the mitochondrial COII and nuclear EF1-α markers sequenced from two lineages in South Africa, including the number of sequences/alleles, total number of singleton haplotypes, fragment length (bp), number of polymorphic sites (*P*), nucleotide diversity (π) and haplotype diversity (*h*) **49**

Stellenbosch University https://scholar.sun.ac.za

Table 3.3 Pairwise genetic divergence values for mitochondrial COII and nuclear EF1-α within and between D1 (*D. abaris*) and D2 (*D. ellobius*) mitochondrial lineages **51**

Table 3.4 Hierarchical analysis of molecular variance (AMOVA) of mitochondrial COII and nuclear EF1-α of two mitochondrial lineages. Fixation index values are given for the hierarchical levels examined: F_{ST}, within populations; F_{CT}, among lineages; and $F_{SC,}$ among localities within lineages. Significant ($p < 0.05$) tests are indicated in bold **51**

Table 3.5 Pairwise Φst values among sampled localities for mitochondrial COII below and nuclear EF1-α above the diagonal. Significant values ($p < 0.05$) are highlighted in bold. (Locality codes as in Table 3.1) **52**

Table 3.6 The morphological characters assessed for individuals of the two lineages with resulting character scores (CS). (also see Table B.6) **58**

Table 4.1 Flea species analysed in the study. Microhabitat preference, host specificity, number of unique occurrence records and two measures of model performance (AUC and TSS) are indicated for each species **67**

Table 4.2 Final list of variable type and individual predictor variables used for modelling **70**

Stellenbosch University https://scholar.sun.ac.za

Chapter 1

General introduction

Stellenbosch University https://scholar.sun.ac.za

1. Siphonaptera

1.1 Taxonomy and systematics

Fleas (Siphonaptera) are obligate arthropod ectoparasites of mammals and birds. This highly specialized holometabolous insect order has a nearly global distribution and can be divided into 246 genera, comprising approximately 2500 described species (Lewis 1998). Southern Africa is home to 30 of these genera and approximately 106 described species. More than half (≈ 65 species) are endemic to South Africa (SA) (modified from Segerman 1995). The families Xiphiopsyllidae (endemic to East Africa) and Chimaeropsyllidae, including two subfamilies of Hystrichopsyllidae (Dinopsyllinae and Listropsyllinae), are unique to the Afrotropical region (Segerman 1995). What gave rise to this diversity is not certain, but historically flea taxonomy was solely based on morphology. This trend has seen a shift more recently particularly with the advances in molecular techniques (e.g. Luchetti et al. 2005, 2007; Lawrence et al. 2014). Combining data from both sources has led to greater insights. Through the use of molecules, it was possible to construct a molecular phylogeny (Whiting 2002; Whiting et al. 2008; Zhu et al. 2015) and to revise some of the taxonomy based both on morphology and genetics (see Luchetti et al. 2005, 2007; Lawrence et al. 2014).

Recent studies of fleas in other parts of the world have found high levels of intraspecific phylogenetic diversity (e.g. de la Cruz & Whiting 2003; Brinkerhoff et al. 2011; Lawrence et al. 2014; Lin et al. 2014), emphasizing the probability of undetected cryptic diversity. In SA, similar uncertainties exist and a case in point can be presented in the form of the endemic Afrotropical subgenus *Dinopsyllus* (see Chapter 3) (Hopkins & Rothschild 1966; de Meillon et al. 1961; Segerman 1995). The distribution of *Dinopsyllus ellobius* is confined to the southern African subregion and in SA it appears to be restricted to more mesic biomes such as those found along the coastal belt and in-land grassland habitats (Hopkins & Rothschild 1966). Following morphological descriptions, Jordan (1930) recognized two distinct sub-species (*D. e. ellobius* and *D. e. abaris*), based on the size and shape of sternite VIII in males. Subsequently, while using the same distinguishing character, the two lineages were regarded as full species by Hopkins & Rothschild (1966). However, the two taxa overlap partly in ranges and have reportedly similar host

Stellenbosch University https://scholar.sun.ac.za

associations, and this single subtle morphological difference is thus also regarded by some as merely geographic variation within the same species, *D. ellobius* (de Meillon et al. 1961; Segerman 1995). These uncertainties can be blamed on the subtle morphological differences coupled with the absence of genetic data for fleas in the region.

Despite the discovery of two new flea species (based on morphological differences) near Garies in the Northern Cape Province of SA (Hastriter 1998, 2000), the taxonomy of the other flea species in the region has not received much attention. Most research to date is in fact heavily biased towards species that are of veterinary and medical importance (Mitchell 1921; Davis 1953; de Meillon et al. 1961; Sheperd & Leman 1983; Sheperd et al. 1983; Scwan 1986; Njunwa et al. 1989; Gage & Kosoy 2005; Stenseth et al. 2008). Flea host lists and distribution maps are available for SA, but are limited and/or outdated (see Segerman 1995). Descriptive studies have been conducted on fleas and other parasites associated with small and medium sized mammals in SA (Braack et al. 1996; Anderson & Kok 2003; Horak & Fourie 1986; Horak et al. 1999, 2004; Matthee et al. 2010a, 2011) and particularly those found on scrub hare (*Lepus saxatilis*) (Louw et al. 1993, 1995), the four striped mouse (*Rhabdomys* spp.) (Tipton 1960; Zumpt 1961; de Meillon et al. 1961; Till 1963; Ledger 1980; de Graaf 1981; Horak et al. 1986, 2005; Howell et al. 1989; Segerman 1995; Horak & Boomker 1998; Petney et al. 2004; Matthee et al. 2007, 2010; van der Mescht 2011), elephant shrews (*Macroscelidopsylla albertyni*) (Beaucournu et al. 2003) and several small carnivores (Horak et al. 1987, 1999, 2000, 2004; Matthee et al. 2011). To gain a better understanding of flea taxonomy and distribution in the region, more data are clearly needed.

1.2 Flea biology

The flea life cycle consists of six developmental phases, initiated when the fertilized egg is produced. Eggs hatch into larvae, then go through three larval moults, and develop into an inactive pupal phase (Marshall 1981; Krasnov 2008). Immature flea development occurs mostly in the off-host environment where larvae feed on organic debris in the hosts nest/burrow (Marshall 1981; Krasnov 2008). Adult fleas emerge from the inactive pupae as soon as conditions are adequate and both sexes then require a blood meal before they can start mating (Marshall 1981; Krasnov 2008).

Stellenbosch University https://scholar.sun.ac.za

It has been suggested that fleas are permanent satellites of their hosts, referring to the association between fleas and their host species (Medvedev & Krasnov 2006; Krasnov 2008). Adult fleas are non-permanent parasites and can be classified according to their microhabitat preference; "fur" fleas spend most of their time in the fur of the host, whereas "nest" fleas spend more time in the nest of the host (Ioff 1941; Marshall 1981; Krasnov 2008). It should be noted that classification of flea microhabitat preference seems to be rather a continuum than strict and that some fleas are also classified as "fur/nest" fleas which spend equal amounts of time in the fur and nest of the host (see Krasnov 2008). Nest fleas are expected to be mostly affected by the off-host environment (e.g. relative humidity and temperature within the nest/burrow), whereas fur fleas will possibly be more affected by regional environmental variables (e.g. climate) and on-host mechanisms (e.g. interactions between parasites and host defence mechanism such as grooming or MHC complexes). In SA for example, *Listropsylla agrippinae* is considered a fur flea whereas *Chiastopsylla rossi* is considered a nest flea (see Chapter 2) (Segerman 1995). Some discrepancies exists surrounding the fur versus nest dichotomy, and although mainly attributed to ambient temperature, other factors like host behaviour, physiological status of fleas, and circadian and seasonal variation in flea behaviour have also been implicated as factors causing variation in this dichotomy (Krasnov 2008 and references therein).

Ambient temperature, rainfall and relative humidity are considered the most important climatic variables for free-living arthropod ectoparasite survival, especially for fleas that are less associated with the host (Marshall 1981; Stark 2002). Ambient temperature and rainfall will be more constant at smaller spatial scales, while at larger spatial scales factors such as soil composition, nest behaviour of the host community and host species composition and abundance are more important determinants of flea diversity and species composition (Krasnov et al. 2015). Soil composition influences both the host and the vegetation of an area. Furthermore, soil structure affects the type of nests that can be constructed (Shenbrot et al. 2002). More specifically, soil texture is considered the most important environmental parameter for fleas that infest burrowing hosts as it affects relative humidity and ambient temperature within the host's burrow (Osacar-Jimenez et al. 2001; Shenbrot et al. 2002). This in turn, may severely affect adult nest fleas (Krasnov et al. 2002a, b) and the development of immature stages within the burrow (Krasnov et al. 2001a, b). Changes in

Stellenbosch University https://scholar.sun.ac.za

soil composition can influence the distribution and nesting behaviour of potential host species (Shenbrot et al. 2002).

Although a certain level of host preference is exhibited, fleas are predominantly generalist in host preference. Host specificity in fleas range from highly host specific (≤ 2 hosts) to -opportunistic/generalist (> 2 hosts) (Marshall 1981; Combes 2001; Krasnov 2008). In SA for example, *L. agrippinae* is considered more host specific whereas *C. rossi* is considered less host specific (Segerman 1995). *Listropsylla agrippinae* has two principle host taxa (*Rhabdomys* spp. and *Myotomys unisulcatus*) whereas *C. rossi* has at least four main host taxa (*Rhabdomys* spp., *Otomys irroratus*, *M. unisulcatus* and *Gerbilliscus brantsii*) documented (Segerman 1995; Froeschke et al. 2013). The level of host specificity is determined by the adaptation of a parasite to various conditions related to ecological, behavioural, physiological and biochemical traits of a particular host or range of host species (Krasnov et al. 2003b). The abundance of highly host specific flea species seems to be controlled mostly by abiotic factors such as climatic conditions, whereas moderate and non-specific fleas are less influenced by abiotic factors, but rather more by host related factors such as host species composition and density (Krasnov et al. 2010a). Flea abundances generally differ between principle and auxiliary host species as well as within auxiliary host species, which may be attributed to the effect of feeding either on a principle or auxiliary host species (Khokhlova et al. 2012). These differences in flea abundance could be explained by the level of similarity between its principle host and the auxiliary host that a flea feeds on (Poulin 2005). It has been surmised that the longer an evolutionary relationship exist between the flea and its principle host, the more efficient feeding and reproductive success will be for the flea while on its principle host (Krasnov 2008). Furthermore, generalist flea species commonly tend to have larger range sizes than specialist flea species (Krasnov et al. 2005a). It is therefore possible that the geographic range of generalist flea species will be determined mainly by the geographic distribution of their principal and auxiliary host species (Shenbrot et al. 2007).

Stellenbosch University https://scholar.sun.ac.za

2. Phylogeography

The majority of parasite-host evolution studies focus at the macro-evolutionary scale by describing the spatial variation of parasite communities (e.g. Page 1994; Johnson et al. 2003; Clayton et al. 2004). Micro-evolutionary processes, which are equally important, have received less attention (e.g. Nadler 1995; Clayton & Johnson 2003; Criscione et al. 2005). Since processes that act at a micro-evolutionary scale will in turn influence patterns at the macro-evolutionary scale (Huyse et al. 2005; Hoberg & Brooks 2008), more integration between the data sets are required to fully understand the mechanisms driving parasite-host evolution. Explaining micro-evolutionary processes in parasite host systems are however complex since the patterns are influenced amongst others by an interplay between abiotic factors, complex parasite life cycles, host biogeography, and variation in the level of parasite-host associations between taxa (see Criscione et al. 2005; Barret et al. 2008; Hoberg & Brooks 2008; Nieberding et al. 2004, 2008; Galbreath & Hoberg 2012; Cangi et al. 2013; du Toit et al. 2013a; Fermino et al. 2013; Poretta et al. 2013b; Martinů et al. 2014).

Phylogeographic studies have the potential to elucidate part of the micro-evolutionary processes through deciphering the spatial genetic structure of parasite and host lineages (see Hafner & Nadler 1988; de la Cruz & Whiting 2003; Nieberding et al. 2004, 2005, 2006, 2008; Gómez-Díaz et al. 2007; Jones & Britten 2010; Štefka et al. 2011; du Toit et al. 2013a; Lin et al. 2014). To date, studies comparing the phylogeography of parasites and their hosts have resulted in variable outcomes (e.g. Hafner & Nadler 1988; de la Cruz & Whiting 2003; Nieberding et al. 2004, 2005, 2006, 2008; Gómez-Díaz et al. 2007; Jones & Britten 2010; Štefka et al. 2011; du Toit et al. 2013a; Lin et al. 2014). Some trends tend to suggest that variability is brought about by evolutionary processes that take place in a highly dynamic geographic context over a long period of time. Although paleoclimatic events are indicated as one of the main mechanisms driving speciation and extinction of host evolutionary lineages (Hewit 2000, 2001, 2011; Erwin 2009), the effect of this on parasitic lineages is not well investigated.

Stellenbosch University https://scholar.sun.ac.za

Globally, only a few studies have investigated the phylogeography of fleas, especially across a large geographic scale (de la Cruz & Whiting 2003; Gómez-Díaz et al. 2007; Luchetti et al. 2007; Jones & Britten 2010; Lin et al. 2014). To date, no study has compared the phylogeography of flea species that vary in host specificity and level of host association. In SA, phylogeographic investigations into parasites are equally rare. Only one recent study exists that indicates partial congruence in spatial genetic structure of a host specific louse (*Polyplax*) and its small mammal host (*Rhabdomys*) (du Toit et al. 2013a). The more host specific fur flea, *L. agrippinae*, and the less host specific nest flea, *C. rossi*, provide us with the perfect opportunity to compare phylogeographic patterns within a generalist parasite. Using the subtle differences in life history between the two species, it is proposed that better substantiated inferences can be made regarding the factors affecting dispersal in this taxonomic group.

3. Factors that shape flea phylogeographic patterns

3.1 Biological traits/Life history of the parasite

Flea life history (e.g. host specificity and microhabitat preference) can influence the geographic distribution of a species (Poulin & Krasnov 2010) and should ultimately be reflected in its spatial genetic structure (Krasnov & Poulin 2010). Variation in level of host specificity could affect the spatial genetic structure of a parasite by influencing its distribution within a host population, community and environment (Krasnov et al. 1997; Medvedev & Krasnov 2006; Froeschke et al. 2013). For example, more host specific avian body lice show higher congruence in spatial genetic structure with the host compared to less host specific wing lice which can be attributed mainly to differences in dispersal ability (Clayton & Johnson 2003). Although both lice depend on host contact for dispersal, the less host specific wing lice sometimes disperse via a phoretic association with parasitic flies which most likely explain the lower congruence in spatial genetic structure with the host (Clayton & Johnson 2003). Given the multi-host behaviour of fleas one would expect lower levels of congruence in spatial genetic structuring relative to hosts, compared to a highly host specific and permanent parasite, such as lice (Marshall 1981; Clayton & Johnson 2003; Huyse et al. 2005; Barrett et al. 2008). The influence of the level of host specificity on flea spatial genetic structure remains unknown at the moment. Although fleas are generalist parasites, it is anticipated

Stellenbosch University https://scholar.sun.ac.za

that higher levels of host specificity could lead to higher dispersal potential with the host and thus higher spatial genetic structure with the host.

Furthermore, parasites also display variations in their life cycles, ranging from complex or indirect life cycles (e.g. some nematodes) to direct life cycles (e.g. lice). Permanent parasites with direct life cycles are generally more dependent on host movement (vertical transmission: host-to-host or host-mediated dispersal) for dispersal and are thus more likely to show congruent spatial genetic structure with their hosts (Blouin et al. 1995, 1999; Johnson et al. 2003). In contrast, parasites that have multiple free-living stages (e.g. ticks) not only depend on host movement (i.e. vertical transmission), but are also more mobile themselves and can come into contact with other host individuals and species (horizontal transmission: high parasite mobility) (Marshall 1981; Huyse et al. 2005). Although adult fleas have a closer association with the body of a host, large scale movement may be restricted by the dependence of the immature stages on the host nest. Thus, fleas should show lower levels of congruence with host spatial genetic structure when compared to direct life cycle species, because all life stages of these parasites do not track the movements of one host species (Gómez-Díaz et al. 2007; Jones & Britten 2010). However, the question remains as to how differences in microhabitat preference will influence spatial genetic structure within fleas. For example, fur fleas spend more time on the host which could increase their dispersal potential with the host and may thus result in lower spatial genetic structure. In contrast, nest fleas spend more time in the nest of the host which could mean lower dispersal potential and they may thus show higher spatial genetic structure across the landscape. Since higher levels of host association may also lead to higher levels of congruence between phylogeographic patterns obtained for the parasite and host (Clayton & Johnson 2003; Huyse et al. 2005; Barrett et al. 2008), it is predicted that the fur flea will show more host congruence than the nest flea.

3.2 Population demography: Effective population size of the parasite

The effective population size of parasites can also affect the level of congruence in spatial genetic structure of parasites and its hosts. Parasite prevalence and abundance are generally related and both are also correlated to effective population size. Hence, low prevalence and low abundance of

Stellenbosch University https://scholar.sun.ac.za

parasites is generally interpreted to reflect lower effective population sizes (Anderson et al. 1998; Blouin 1998; Blouin et al. 1999). Lower effective population size in parasites usually results in higher congruence in spatial genetic structure with the host (and also lower genetic diversity locally) when compared to parasites with larger effective population sizes. Low effective population sizes will increase the chances for genetic drift (Blouin et al. 1995; Huyse et al. 2005). Higher effective population size in parasites usually result in lower congruence in spatial genetic structure with the host as well as higher genetic diversity locally (Blouin et al. 1995; Huyse et al. 2005). Host specific permanent parasites, like lice, generally have low prevalence and abundance on their host (Marshall 1981). These parasites are thus expected to have smaller effective population sizes and higher probability of congruent spatial genetic patterns with the host (de Meeús 2000). However, the effective population size of parasites with multiple free-living stages and those that are less host specific can be more complicated to interpret because these parasites vary in prevalence and abundance on the host (Marshall 1981). The estimation of effective population size for fleas are further perplexed by offspring that could be released into the off-host environment where they can come into contact with other populations and possibly be recruited again at a later stage (Criscione et al. 2005). For example, it might be more difficult to predict effective population size for nest fleas (that have more individuals present in the nest of the host) from adult flea individuals sampled on the host as this may not be a good representative sample of the actual nest flea population.

3.3 Host related factors

The dispersal ability of species will affect genetic variation among presumably isolated distant populations. The ability to disperse will have a direct effect on gene flow which in turn is affecting the rate of genetic drift between populations. In the case of a parasite, gene flow is usually closely linked to the dispersal ability of the host (see Blouin et al. 1995; Johnson et al. 2003; McCoy et al. 2003; Huyse et al. 2005). For example, gene flow in parasites with direct life cycles (normally also more host specific) is primarily determined by host contact and movement, which could result in parasite speciation in response to host speciation (Blouin et al. 1995, 1999; Johnson et al. 2003; McCoy et al. 2003). While in the case of temporary parasites, gene flow is determined by host contact and movement as well as movement of the parasite itself.

Stellenbosch University https://scholar.sun.ac.za

Besides parasite life cycle and level of host association, the host biogeography also contributes to parasite spatial genetic patterns (Nieberding et al. 2008; du Toit et al. 2013a). In SA, taxon pulses and vicariant barriers have had similar influences on multiple small mammal host species (Matthee & Robinson 1996, 1997; Jansen van Vuuren & Robinson 1997; Rambau et al. 2003; Russo et al. 2006, 2010; Smit et al. 2007; Willows-Munro & Matthee 2011; Edwards et al. 2011; Engelbrecht et al. 2011; Montgelard & Matthee 2012; du Toit et al. 2012), but the influence on parasite phylogeography has only been studied once (du Toit et al. 2013a). The significant geological and climatic events in the natural history of SA will be discussed in the following section.

3.4 Abiotic factors

Parasites with free-living stages and particularly those that are less host specific are also affected by the off-host environment (Gómez-Díaz et al. 2007; Jones & Britten 2010). Geological events during the post-Gondwana period had arguably the most significant influence on the present-day environment of southern Africa. Uplift and tilting coupled with weathering and erosion during the middle Miocene resulted in significant changes in the southern African landscape (Partridge & Maud 1987). For example, the uplift of the Great Escarpment in SA, at the edge of the plateau separating the coast from the interior Highveld, is predicted to have happened between 20-500 million years ago (Dollar & Goudie 2000). Major geomorphic post-rifting events contributed to changes in climate. The change in climate from glacial to drier conditions toward the end of the Miocene seems to be attributed to the growth of the Antarctic ice sheet (Partridge 1990) as well as the formation of the Agulhas and Benguela currents along the coastline (Dollar & Goudie 2000). These macro-scale geological processes related to plate tectonics and subsequent climatic changes over evolutionary time has resulted in the establishment of modern biomes in SA (Coetzee 1978; Scott et al. 1997). Biomes each have their own set of environmental elements (e.g. rainfall pattern, plant species composition and vegetation structure), which have acted and are still acting as determinants in historical and contemporary biodiversity patterns. It is thus not surprising that studies on various vertebrate taxa in SA points toward divergences among lineages that coincide with the paleoclimatic events during the Pliocene and Pleistocene (Matthee & Flemming 2002;

Stellenbosch University https://scholar.sun.ac.za

Tolley et al. 2006, 2008; Smit et al. 2007; Willows-Munro & Matthee 2009, 2011; Russo et al. 2010; Montgelard & Matthee 2012; du Toit et al. 2012).

Biome-related vegetation changes and topographical variation across landscapes influence gene flow and therefore play an important role in shaping the range size and position of various plant and animal taxa in SA (Montgelard & Matthee 2012). A topographic example is the Great Escarpment in SA that seems to have acted as a contemporary barrier in reducing contact and subsequent gene flow between small mammals based on altitudinal differences (e.g. Edwards et al. 2011; du Toit et al. 2012). The 'Bedford Gap' is a good example of a biome-related vegetation barrier in SA (Lawes 1990). Vegetation in this area is known as the Albany Thicket and it is the meeting place of at least five biomes in SA (Mucina & Rutherford 2006). The area is considered a contact zone where ecologically and genetically differentiated small mammal individuals could occur sympatrically. A specific example is a phylogeographic study that was conducted on *Rhabdomys* that identified three distinct genetic lineages displaying overlap in their distribution in this region (du Toit et al. 2012). More recently, it has been suggested that both taxon pulses (episodes of range contraction and expansion) and vicariance (barrier leading to isolation) have been acting across micro- and macro-evolutionary scales (Hoberg & Brooks 2008). In SA, the majority of studies suggested that the mechanism responsible for the geographic genetic differentiation among clades can be correlated with biome-related paleoclimatic changes and, to some extent, also changes in topography (Matthee & Robinson 1996, 1997; Jansen van Vuuren & Robinson 1997; Rambau et al. 2003; Russo et al. 2006, 2010; Smit et al. 2007; Willows-Munro & Matthee 2011; Edwards et al. 2011; Engelbrecht et al. 2011; Montgelard & Matthee 2012; du Toit et al. 2012).

These biogeographic changes can significantly affect parasite-host interactions through evolutionary time (Hoberg & Brooks 2008). Parasites with direct life cycles and those that are more host specific, that have a shared biogeographic history with their host, could result in congruent spatial genetic patterns through similar responses to evolutionary events (e.g. vicariance) (Hoberg & Klassen 2002; Nieberding et al. 2004; Hoberg & Brooks 2008). However,

Stellenbosch University https://scholar.sun.ac.za

biogeographic changes can also directly affect the spatial genetic structure of parasites with free-living stages and those that are less host specific (Hoberg & Brooks 2008; Nieberding et al. 2008). The effect on flea spatial genetic structure remains as yet unstudied for SA.

4. Aims and objectives

The overall aim of the current study is to explore the mechanisms involved in shaping distribution, diversity and diversification of selected flea species occurring in SA. The main objectives of the PhD study are to test if:

1. Fleas with different levels of host association (microhabitat preference and host specificity) differ in spatial genetic structure. If niche breath (number of hosts it parasitize) and time spent on the host affect 1) genetic diversity of the parasite and 2) congruence in phylogeographic structure between parasites and hosts (McCoy et al. 2001; Gómez-Díaz et al. 2007; Jones & Britten 2010), it is predicted that the more host specific fur flea (*L. agrippinae*) will show lower levels of intraspecific genetic diversity and also more similarities to the vicariant patterns of the hosts they infest when compared to the less host specific nest flea (*C. rossi*). In other words, genetic distances among individuals from different sampling sites are expected to be higher for the nest flea (*C. rossi*) when compared to the fur flea (*L. agrippinae*) since the latter can utilize the host more for dispersion.

2. *Dinopsyllus ellobius* contains cryptic diversity throughout SA. Given the wide host range, the previous description of two species (*D. ellobius* and *D. abaris* by Hopkins & Rothschild 1966) based on a morphological character, and the potential effect of host and habitat vicariance on the evolution of fleas in the region, we predict that *D. ellobius* (as described by de Meillon et al. 1961; Segerman 1995) may contain cryptic diversity in the region.

3. Species distribution model performance and the relative importance of predictor variables differ between flea species with different levels of host association (microhabitat preference and host specificity). It is predicted that fur fleas will be more accurately modelled due to being more strongly affected by variables associated with regional environmental conditions (e.g.

Stellenbosch University https://scholar.sun.ac.za

climate), while nest fleas will be less accurately modelled due to being affected by conditions within the host nest or burrow (e.g. soil conditions and microclimate; since all life stages spend the majority of their life cycle off the host and have limited dispersal capabilities). Furthermore, we predict that host specific fleas will be more strongly associated with the abiotic variables constraining their host's distribution (and therefore will be more accurately modelled by SD modelling) than host opportunistic fleas in the region. This is expected because host specific fleas may have been adapted to the immediate environment of their specific host and thus are expected to tolerate a narrower range of physical conditions compared to host opportunistic fleas (Krasnov et al. 2003b, 2015; Krasnov 2008; Shenbrot et al. 2007).

Stellenbosch University https://scholar.sun.ac.za

Chapter 2

Comparative phylogeography between two generalist flea species reveal a complex interaction between parasite life history and host vicariance: Parasite-host association matters

**BMC Evolutionary Biology 15 (2015): 105*

Stellenbosch University https://scholar.sun.ac.za

1. Introduction

Paleoclimatic events have been put forward as one of the main mechanisms causing speciations and extinctions of evolutionary lineages (Holt 1990; Hewitt 2000, 2011; Erwin 2009). The differences in ecology and life history of taxa, however, often result in differential species-specific responses to similar climatic events (Davis et al. 2005; Araújo et al. 2008; Montgelard & Matthee 2012) and a simple explanation describing speciation processes hardly exists. For parasitic taxa the situation is further perplexed by complex life cycles, host biogeography, and the level of parasite-host associations (Criscione et al. 2005; Barrett et al. 2008; Hopberg & Brooks 2008; Nieberding et al. 2008; Cangi et al. 2013; du Toit et al. 2013a; Fermino et al. 2013; Poretta et al. 2013b; Martinů et al. 2014). Although some progress has been made regarding this topic (Nieberding et al. 2004; Galbreath & Hoberg 2012; Morand 2012; Olival et al. 2013) more quantifiable data are needed to make more accurate predictions on for example the factors affecting range expansions and connectivity among populations. For many parasitic taxa, this is desperately needed from a disease ecology perspective (Bitam et al. 2010; Daszak et al. 2010).

In southern Africa, several phylogeographic studies have been performed on small vertebrate taxa (for example see Matthee & Flemming 2002; Bauer & Lamb 2005; Smit et al. 2007; Russo et al. 2010; Engelbrecht et al. 2011; Willows-Munro & Matthee 2011; du Toit et al. 2012). From these studies, paleoclimatic changes are once again suggested as one of the main drivers of evolution. Congruent phylogeographic patterns among lizards, rodents, shrews and elephant shrews also suggested the existence of vicariant biogeographic barriers in the region (for review see Linder et al. 2010; Montgelard & Matthee 2012). Of particular relevance to the present study is the support for regional vicariance found in *Rhabdomys* (du Toit et al. 2012), *Myosorex* (Willows-Munro & Matthee 2011), *Otomys* (Engelbrecht et al. 2011) and *Micaelamys* (Russo et al. 2010). These small mammals show distinct genetic clades that can be associated with the xeric western succulent biomes and the more mesic eastern predominantly grassland biome of southern Africa (Fig. 2.1a; Fig. A.1; also see Montgelard & Matthee 2012). In addition, *Rhabdomys*, *Myosorex* and *Micaelamys* show differentiation to a larger or lesser extent between the winter rainfall zone and the aseasonal rainfall zone in the Cape Floristic Region (Fig. 2.1a; Fig. A.1; also see Linder et al.

Stellenbosch University https://scholar.sun.ac.za

2010). From these studies it is evident that the diversification of small mammals in the region is driven by complex interactions between life history, climate and topographic barriers.

The effect of vicariant barriers on the evolution of parasites occurring in southern Africa is virtually unknown. An exception to this is provided by (du Toit et al. 2013a), who recently indicated partial congruence in spatial genetic structure between *Rhabdomys* and its species specific lice, *Polyplax*. Distinct genetic clades in the parasites and in the hosts supported the documented mesic eastern and xeric western divide in the subregion (see above). Co-evolution analyses, however, failed to support a strong signal of geographic co-differentiation between parasite and host and the authors suggested that the resultant pattern is due to the synergistic effects of parasite traits (host specificity), host-related factors (the vagility and social behavior of *Rhabdomys*) and the biogeography (vicariance) of the region (du Toit et al. 2013a). To expand on these findings, we here present phylogeographic data on two generalist flea species occurring on *Rhabdomys* and other small mammal species in southern Africa.

Individual flea species are normally adapted to utilize a variety of hosts in the environment and are all obligatory blood feeders of endothermic vertebrates (Marshall 1981; Krasnov 2008). The life cycle consists of four stages (egg, larva, pupa, and adult) and fleas are dependent on the nest environments of the hosts to variable extent (Marshall 1981). Immatures develop entirely in the off-host environment while adult stages have to spend some time on the host to obtain a blood meal. The length of time that is spent on the host, however, varies between flea taxa (Ioff 1941; Marshall 1981; Krasnov 2008). This observation has provided an opportunity to coarsely classify fleas according to their microhabitat preference as either a “fur“ (adult stage spend more time on the host) or a “nest“ flea (adults spend more time in the nest/burrow of the host) (Ioff 1941; Marshall 1981; Krasnov 2008). These differences can have profound effects on the ability of the flea species to disperse over the landscape and we therefore predict that fur fleas will show more phylogeographic congruence with the host/s (adult fleas spend more time on the host and can thus disperse over the landscape) and nest fleas will show more restricted gene flow across the landscape (since all juveniles and the adults spend the majority of their life cycle off the host and have limited dispersal capabilities).

Stellenbosch University https://scholar.sun.ac.za

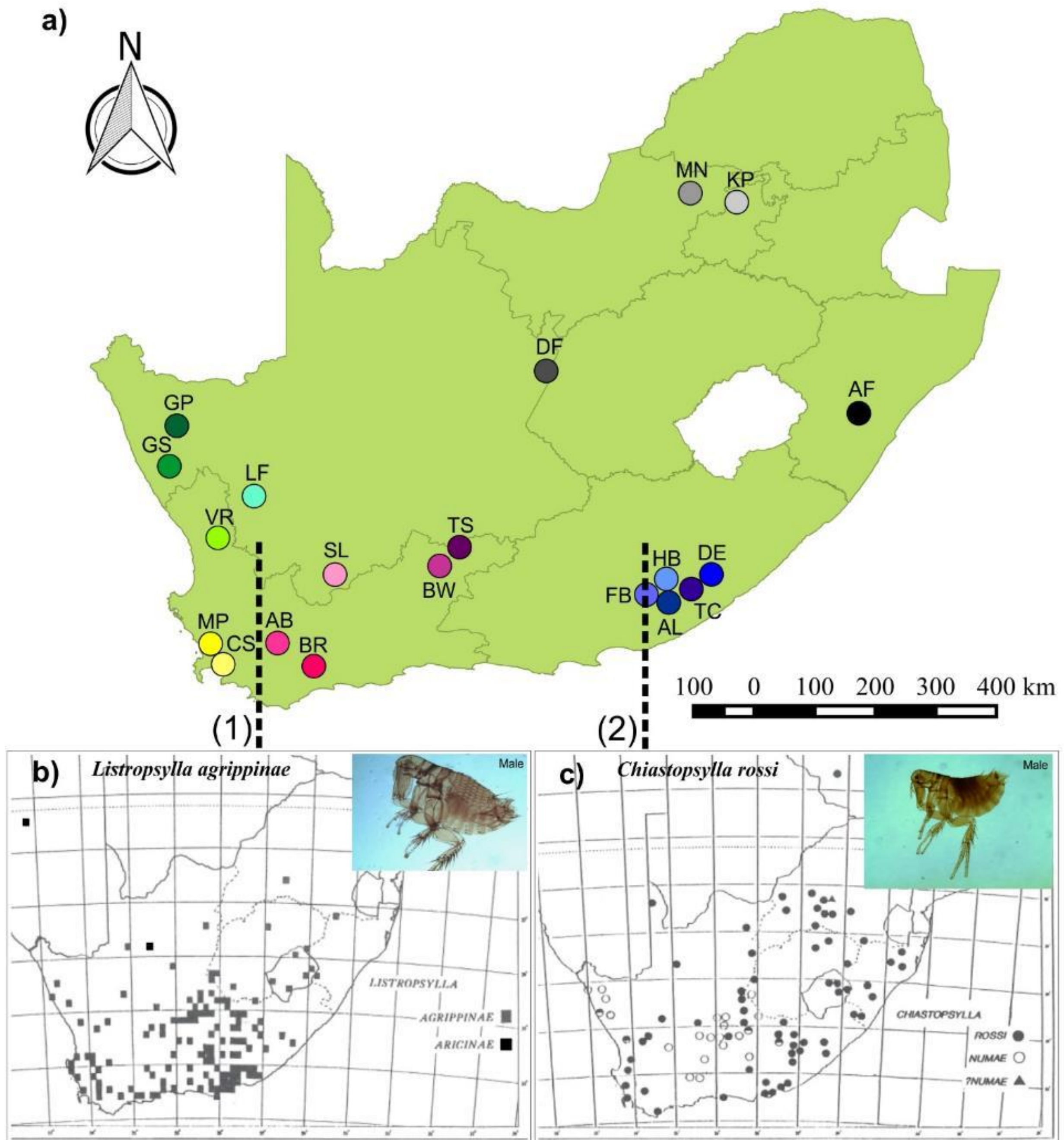


Figure 2.1 Map indicating **a)** localities sampled with their respective codes indicated as in Table 2.1. The vicariant breaks in South Africa (1) winter and aseasonal rainfall break and (2) xeric and mesic biomes are indicated by white dashed lines, **b)** the mapped distribution of *L. agrippinae* and **c)** *C. rossi* are taken directly from Segerman (1995).

Stellenbosch University https://scholar.sun.ac.za

To test the hypothesis that fleas with different levels of host association will differ in spatial genetic structure, the fur flea, *Listropsylla agrippinae*, and the nest flea, *Chiastopsylla rossi* (Segerman 1995; van der Mescht et al. 2013) were selected. Both *L. agrippinae* and *C. rossi* are widespread throughout southern Africa (Fig. 2.1) (de Meillon et al. 1961; van der Mescht et al. 2013) and span at least two well documented vicariant biogeographic barriers in the region (Fig. 2.1) (Linder et al. 2010; Montgelard & Matthee 2012). Apart from differences in the duration of time adults spend on the host, the two species also differ in the level of host specificity (niche breath). *Listropsylla agrippinae* has two principle host taxa (*Rhabdomys* spp. and *Myotomys unisulcatus*) whereas *C. rossi* has at least four main host taxa (*Rhabdomys* spp., *Otomys irroratus*, *M. unisulcatus* and *Gerbilliscus brantsii*) documented (Segerman 1995; Froeschke et al. 2013; van der Mescht et al. 2013). It has been suggested that host specificity may influence the level of intraspecific genetic divergences since more generalist parasite species will show a higher level of intraspecific genetic variation enabling them to infest a broader host range (McCoy et al. 2001; Gómez-Díaz et al. 2007; Jones & Britten 2010). Furthermore, microhabitat preference (specifically referring to fur vs nest) is not always related to host specificity, meaning that fleas may differ in host specificity irrespective of microhabitat preference (see de Meillon et al. 1961; Segerman 1995; Krasnov 2008). If niche breath and time spent on the host affect 1) genetic diversity of the parasite and 2) congruence in phylogeographic structure between parasites and hosts (McCoy et al. 2001; Gómez-Díaz et al. 2007; Jones & Britten 2010), we predict that *L. agrippinae*, when compared to *C. rossi*, will show lower levels of intraspecific genetic diversity and also more similarities to the vicariant patterns of the hosts they infest. Genetic distances among individuals from different sampling sites are expected to be higher for the nest flea, *C. rossi*, when compared to the fur flea, *L. agrippinae*, since the latter can utilize the host for dispersion. By considering the interplay between life history and geography, the present study should add valuable information needed to explain some of the evolutionary processes that shape ectoparasite distribution and diversity.

2. Materials and Methods

2.1 Sample collection

Small mammal trapping was performed during 2010 - 2013 at 20 localities (natural areas and low density grazing farms) in South Africa (SA) (Fig. 2.1a; Table 2.1). Baited Sherman-type live traps

Stellenbosch University https://scholar.sun.ac.za

were set in a line transect and sampling varied between 4 to 7 days per locality. All adult specimens of species that are listed as potential hosts for the two flea species were selected and juveniles were released at the trap site. Trapped animals were placed in a plastic bag before they were euthanized with sodium pentobarbital (200 mg/kg; ethical approval reference number SU-ACUM11-00004). The bodies were brushed over a white plastic tray and all fleas were collected. The brush was inspected and cleaned (using 96 % ethanol) after each animal was processed and new brushes were used for each host species at each locality. Individual fleas were placed in separate tubes filled with 96 % ethanol. Before DNA extraction, preliminary identification of *L. agrippinae* and *C. rossi* individuals was performed using a Leica stereoscopic microscope (Leica Microsystems, Wetzlar, Germany) and the taxonomic key of (Segerman 1995). After DNA extraction the exoskeletons of all extracted fleas were mounted (see van der Mescht et al. 2013) and a thorough morphological identification was done under a Leica DM 3000 light microscope (Leica Microsystems, Wetzlar, Germany) using the key of Segerman (1995). Most of the host species that were trapped during the study are quite common and widely distributed throughout SA. As a result voucher specimens of the species are readily available in several museums. In the case of the fleas, voucher specimens will be deposited in the Museum of the Department of Conservation Ecology and Entomology (Stellenbosch University) and the National Flea Collection in Johannesburg (both in SA).

2.2 DNA extraction, amplification and sequencing

Total genomic DNA was extracted with a Qaigen, DNeasy® Blood and Tissue kit (Qaigen, Valencia, CA, USA) following the protocol of the manufacturer. Whole flea specimens were placed in the extraction buffer containing Proteinase K (600 mAU/ml solution or 40 mAU/mg protein), and digested at 56°C overnight. After digestion, flea exoskeletons were removed for identification purposes (see above). Polymerase Chain Reactions (PCR) and sequencing were performed on the mitochondrial Cytochrome Oxidase II (COII) gene and the nuclear intron Elongation Factor 1 alpha (EF1-α) using published primers (Whiting 2002; Whiting et al. 2008; Table 2.1; Table B.2). Nuclear and mitochondrial regions were amplified using a GeneAmp® PCR 2700 thermal cycler (Applied Biosystems, Foster City, CA, USA). Mitochondrial COII and nuclear EF1-α regions were amplified following standard procedures (Table B.2). Sequencing was

Stellenbosch University https://scholar.sun.ac.za

performed on an ABI 3730 XL DNA analyzer (Applied Biosystems) using BigDye termination chemistry (version 3.1, Applied Biosystems).

Table 2.1 Locality information from where specimens were obtained for each of the two flea species, with codes corresponding to Fig. 2.1.

Province	Locality	Code	Geographic coordinates	*L. agrippinae* (n)		*C. rossi* (n)	
				COII	**EF1-α**	**COII**	**EF1-α**
Western Cape	Anysberg	AB	-33.46 S 20.59 E	5	1	14	13
	Beaufort West	BW	-32.22 S 22.80 E	13	7	14	6
	Buffeljagsrivier	BR	-34.05 S 20.53 E	6	0	10	9
	Kanu	CS	-33.95 S 18.83 E	10	5	7	6
	Mooiplaas	MP	-33.92 S 18.75 E	4	4	11	8
	Vanrhynsdorp	VR	-31.73 S 18.77 E	5	3	15	11
Northern Cape	Dronfield	DF	-28.74 S 24.77 E	*	*	4	3
	Garies	GS	-30.43 S 17.89 E	10	4	5	4
	Loeriesfontein	LF	-30.95 S 19.44 E	15	6	16	14
	Springbok	GP	-29.70 S 18.03 E	14	6	3	2
	Sutherland	SL	-32.40 S 20.90 E	14	4	8	8
	Three Sisters	TS	-31.89 S 23.15 E	10	2	*	*
Eastern Cape	Alice	AL	-32.79 S 26.85 E	*	*	5	2
	Dohne	DE	-32.53 S 27.46 E	11	2	10	5
	Fort Beaufort	FB	-32.78 S 26.63 E	9	3	17	8
	Hogsback	HB	-32.59 S 26.92 E	*	*	17	10
	The Croft	TC	-32.55 S 27.37 E	*	*	11	6
Gauteng	Kaalplaas	KP	-25.63 S 28.17 E	*	*	5	5
North West	Mooinooi	MN	-25.47 S 27.33 E	*	*	8	2
KwaZulu-Natal	Albert Falls	AF	-29.47 S 30.40 E	*	*	9	5
6 provinces	**20 localities**			**126**	**47**	**189**	**127**

Stellenbosch University https://scholar.sun.ac.za

2.3 Alignment and phylogenetic analyses

Sequences for each gene fragment were edited and aligned using BioEdit Sequence Alignment editor 7.2.5 (Whiting et al. 2008). Mitochondrial sequences were translated into amino acids using EMBOSStranseq (www.ebi.ac.uk/Tools/st/emboss_transeq) to confirm functionality. Heterozygous positions in the nuclear fragments were resolved in DNASP 5 (Librado & Rozas 2009) using PHASE 2.1.1 (Stephens et al. 2001, 2005). The algorithm was run for 1000 generations with a thinning interval of 1 and burn-in of 100 generations. Phases with a 0.9 probability or higher were considered resolved and the analysis was performed three times to see if there was any significant difference in results between runs (Stephens et al. 2001).

2.4 Diversity indices and population level analyses

Nucleotide diversity (π) and haplotype diversity (h) values were obtained using DNASP 5 (Librado & Rozas 2009). The evolutionary relationships between haplotypes were investigated by constructing statistical parsimony networks with 95 % confidence intervals in TCS 1.21 (Clement et al. 2000). The best-fit models of sequence evolution were determined for each fragment under the AICc (Akaike 1973; Posada & Buckley 2004) in jModelTest 2.1.4 (Posada 2008; Darriba et al. 2012). To further gain an evolutionary perspective on the association between clusters, individual HKY-corrected networks of each flea species were drawn for mitochondrial COII and nuclear EF1-α using the Neighbor-Net method (Bryant & Moulton 2004) implemented in SplitsTree 4.5 (Huson & Bryant 2006). HKY-corrected sequence distances among mitochondrial COII and nuclear EF1-α were calculated using PAUP* v4.0b10 (Swofford 2000). To investigate population structure without a-priori assumptions, a Bayesian Analyses of Population Structure (BAPS) was performed in BAPS 6.0 (Corander et al. 2008) on the mtDNA data. Spatial genetic mixture analyses of individuals and of groups were performed independently using a vector of maximum K values (each replicated 5 times; Excoffier et al. 1992, 2010). Analysis of molecular variance (AMOVA; Excoffier et al. 1992) and pairwise Φst statistics between sampled populations were performed in ARLEQUIN 3.5.1.2 (Excoffier et al. 2010). Only sampling localities with more than 5 individuals were included. The AMOVA higher level group differentiations were defined based on the subclusters obtained in the BAPS analysis (only sampling localities with more than 5 individuals were used). Mantel tests (Mantel 1967) were performed to test for isolation by

Stellenbosch University https://scholar.sun.ac.za

distance in ALLELES IN SPACE (Miller 2005). Spatial genetic structure was further explored by making use of genetic landscape shape interpolation surface plots constructed in ALLELES IN SPACE (Miller 2005).

2.5 Dating main phylogenetic events

We used a relaxed exponential Bayesian molecular clock as implemented in BEAST 2.1.3 (Bouckaert et al. 2014) to estimate divergence time between clusters. Siphonaptera lacks a useable fossil record and as a calibration point we employed the 2.3 % per million years estimated for various arthropod taxa (Brower 1994; Kandul et al. 2004). The HKY + G model was used and the birth-death process of speciation with exponential priors specified. MCMC simulation ran for 20 million generations, sampling every 10 000 generations for each of the two runs performed. Convergence and mixing were assessed in Tracer 1.6 (Bouckaert et al. 2014) and the first 25 % of the trees were discarded as burn-in. A maximum clade credibility tree was produced in TreeAnnotator 2.1.3 (Bouckaert et al. 2014).

3. Results

3.1 Parasite prevalence and distribution

Both *R. pumilio* and *R. dilectus* were trapped at a contact zone, Fort Beaufort (also see du Toit et al. 2012), and represent the only locality where more than one *Rhabdomys* species was trapped (see Table B.1). *Listropsylla agrippinae* was recorded at lower prevalence at the majority of localities when compared to *C. rossi* (Fig. 2.2; Table B.1). *Listropsylla agrippinae* was found on six host species and was most prevalent on *R. intermedius* (23.02 %) (Fig. 2.2a). *Chiastopsylla rossi* was found on nine host species, but was most prevalent on *R. intermedius* (38.89 %) and *O. irroratus* (37.31 %) compared to the other seven host taxa (Fig. 2.2b). Individuals selected for sequencing were representative of the various host species trapped at each locality (Fig. 2.2; Table B.1).

Stellenbosch University https://scholar.sun.ac.za

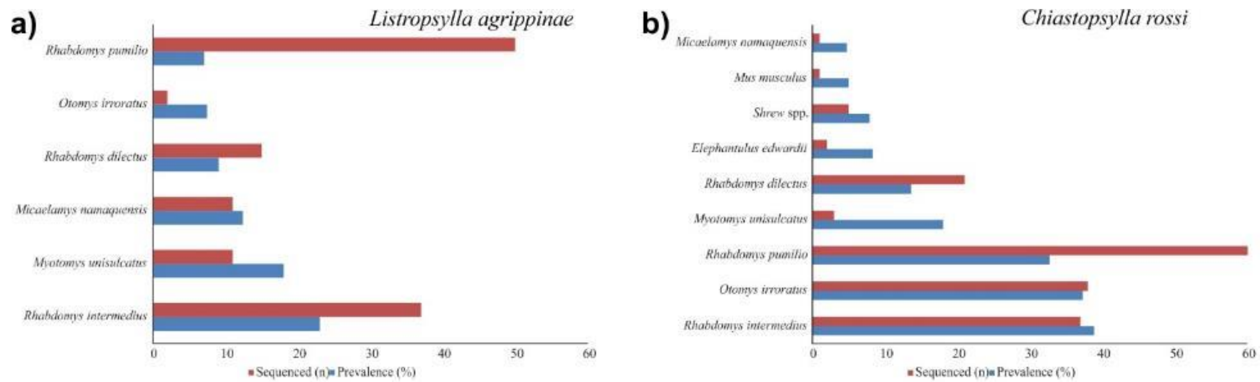


Figure 2.2 The number of individuals sequenced and prevalence of **a)** *L. agrippinae* and **b)** *C. rossi* sampled from different small mammal host species at 20 localities throughout South Africa.

3.2 Characteristics of molecular markers

Mitochondrial COII data for 126 *L. agrippinae* and 189 *C. rossi* specimens were generated and attempts were made to generate a similar nuclear EF1-α data set. Nuclear amplification was less successful but we nonetheless identified 94 *L. agrippinae* and 254 *C. rossi* nuclear alleles (Table 2.2) [GenBank: KR 263182-263844]. Haplotype and nucleotide diversity values were lower for *L. agrippinae* when compared to *C. rossi* for mitochondrial COII and nuclear EF1-α data (Table 2.2). Unexpectedly, the nuclear EF1-α data for both species resulted in higher diversity estimates in terms of haplotypic and nucleotide diversity when compared to the mitochondrial COII data (Table 2.2).

Stellenbosch University https://scholar.sun.ac.za

Table 2.2 Summary statistics for the mitochondrial DNA (COII) and nuclear intron (EF1-α) markers sequenced from *L. agrippinae* and *C. rossi*. The number of sequences/alleles, total number of haplotypes, total number of singleton haplotypes, fragment length (bp), number of polymorphic sites (*P*), nucleotide diversity (π) and haplotype diversity (*h*) is indicated.

	Sequences/ Alleles	Total haplotypes	Singleton haplotypes	bp	*P*	π ± SD	*h* ± SD
COII							
L. agrippinae	126	35	13	576	62	0.013 ± 0.003	0.925 ± 0.015
C. rossi	189	51	26	513	69	0.021 ± 0.006	0.958 ± 0.005
EF1-α							
L. agrippinae	47/94	54	42	593	88	0.010 ± 0.003	0.946 ± 0.017
C. rossi	127/254	177	90	579	184	0.022 ± 0.004	0.994 ± 0.001

3.3 Phylogeographic analyses

3.3.1 *Listropsylla agrippinae*

TCS statistical parsimony analysis based on the mitochondrial COII data showed two distinct clusters (L1 and L2; Fig. 2.3a and b) that could not be connected within the 95 % confidence interval. These two clusters are separated by an average corrected mtDNA sequence divergence of 3.02 % (± 0.36) and this large differentiation is further supported by the Neighbor-Net tree (Fig. 2.3c). The clusters correspond to the xeric western (L1) and eastern mesic (L2) zones of the country. The BAPS analysis corroborated the two clusters detected by TCS and the Neighbor-Net tree, and also indicate some additional substructure within cluster L1 (P = 1.00; log likelihood of optimal partition = - 1179.80) (Fig. 2.3a and c). These two subclusters loosely correspond to the winter and aseasonal rainfall divide (Fig. 2.1a).

Stellenbosch University https://scholar.sun.ac.za

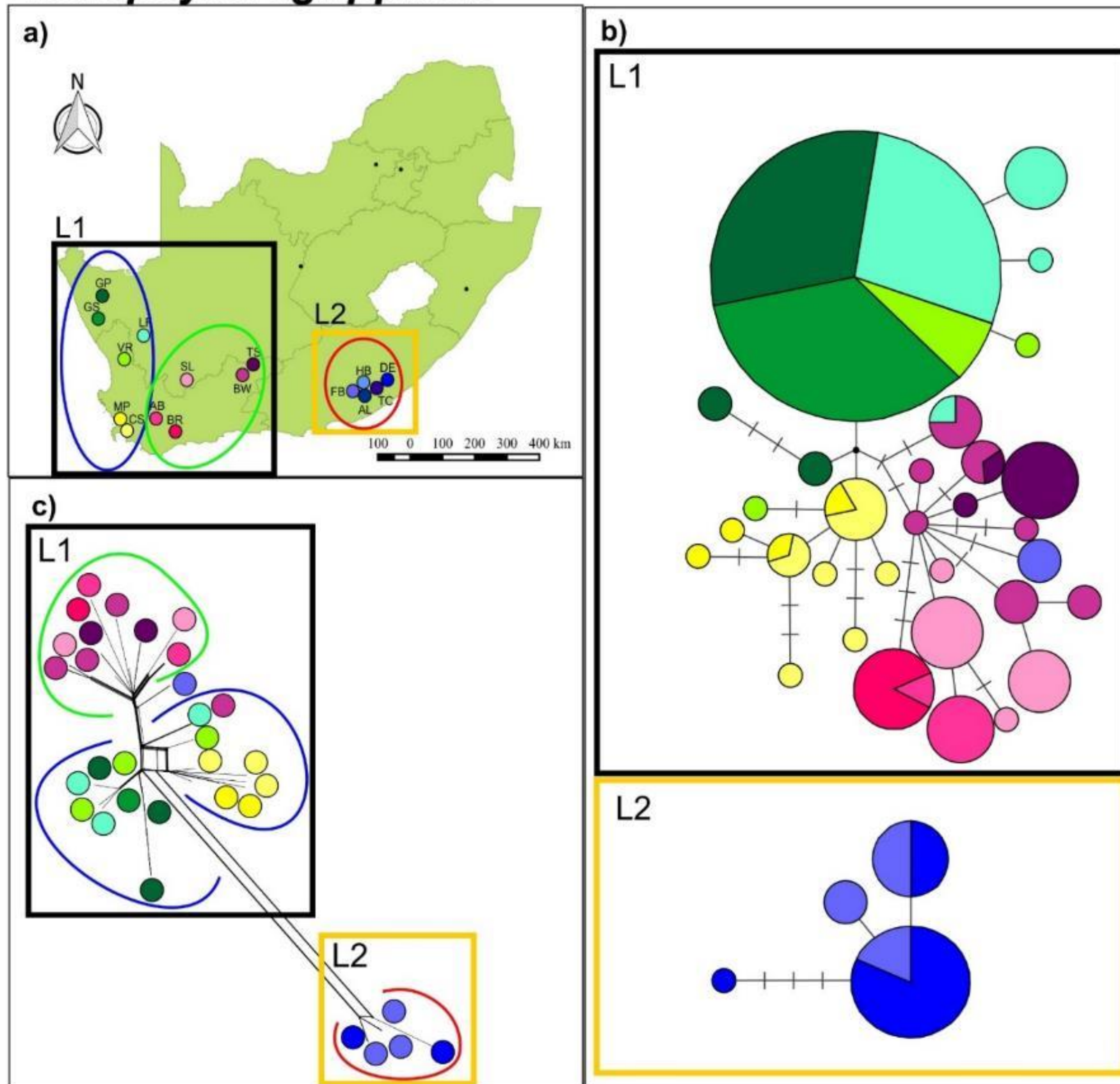


Figure 2.3 a) A map of geographical distribution of sampling localities in South Africa for *L. agrippinae*. Bayesian analysis of population structure (BAPS) subclusters are indicated by green, blue and red circles. **b)** Statistical parsimony mitochondrial COII haplotype network colour coded according to sampling locality. Circle size depicts frequency, branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes/mutational steps. Clades are bound by black (L1) and deep yellow (L2) line boxes. **c)** Neighbour-Net phylogenetic network for *L. agrippinae* labelled according to haplotype groupings from statistical parsimony results and are bound by black (L1) and deep yellow (L2) line boxes. Bayesian analysis of population structure (BAPS) subclusters are indicated by green, blue and red lines.

Stellenbosch University https://scholar.sun.ac.za

The spatial genetic patterns were less pronounced when the nuclear EF1-α data is compared to the mitochondrial DNA network (Figs. 2.3 and 2.4). There is some evidence for regional clustering as indicated by the grouping of similarly shaded colors in the parsimony network and Neighbor-Net tree (Fig. 2.4) but more importantly, however, the most common DNA allele is shared between the far eastern and far western side of the sampling distribution. Mantel tests for mitochondrial COII data indicated that there was weak isolation by distance when calculated for all localities ($r = 0.11$; $P = 0.00$) and when calculated for each cluster L1 ($r = 0.11$; $P = 0.01$) and L2 ($r = 0.16$; $P = 0.01$) separately. The time of divergence between the two main clusters within *L. agrippinae* date back to the late Miocene which was estimated to be 5.27 Ma (95 % HPD interval: 2.31, 9.58 Ma).

Stellenbosch University https://scholar.sun.ac.za

Listropsylla agrippinae

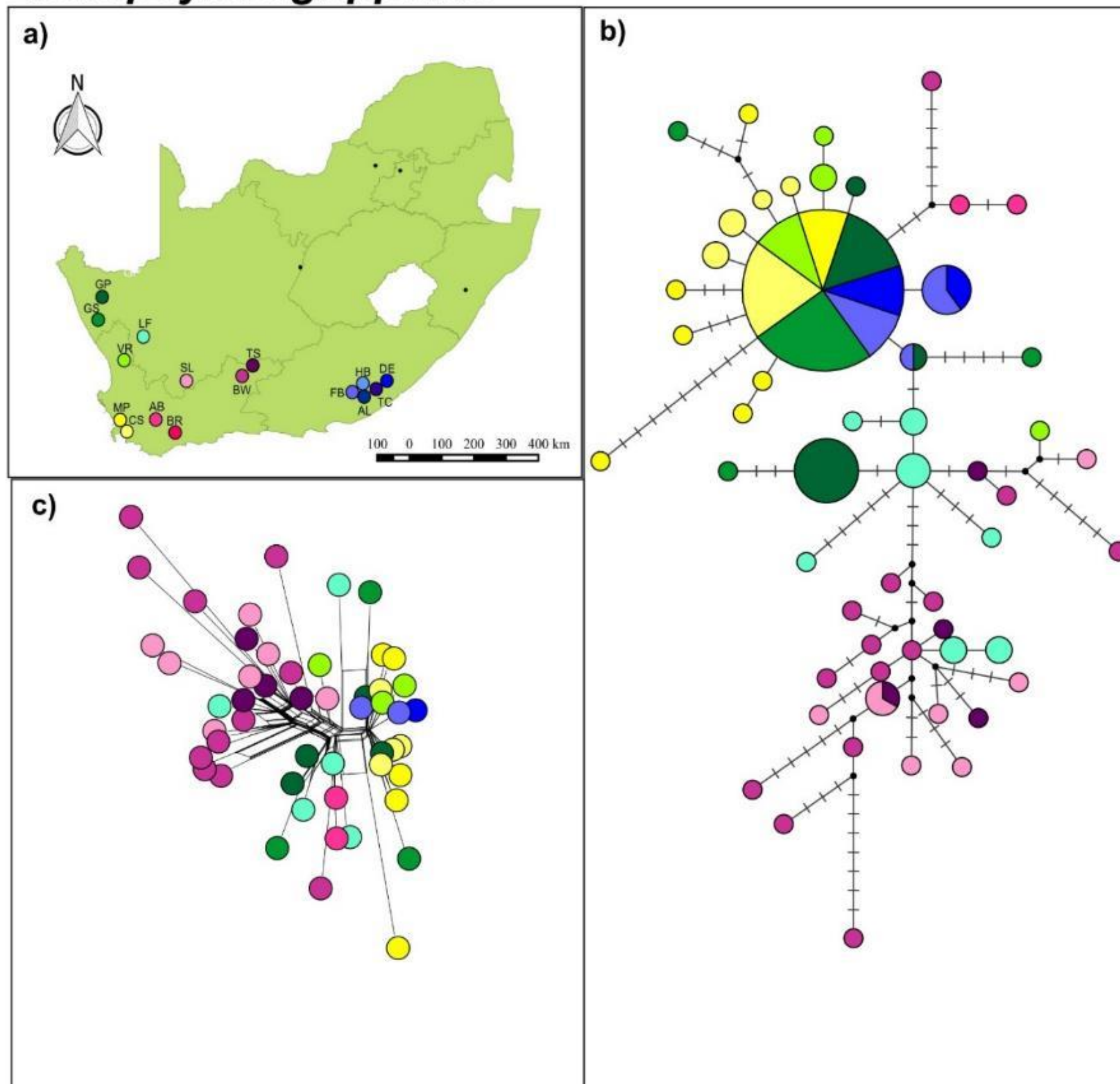


Figure 2.4 a) A map of geographical distribution of sampling localities in South Africa for *L. agrippinae*. **b)** Statistical parsimony nuclear EF1-α haplotype network colour coded according to sampling locality. Circle size depicts frequency, branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes/mutational steps. **c)** Neighbour-Net phylogenetic network for *L. agrippinae*.

Stellenbosch University https://scholar.sun.ac.za

Genetic landscape interpolation surface plots for the COII data indicated that there were marked differences between some sampling localities when the distribution of genetic diversity found at each is compared across the landscape (Fig. 2.5a). The graphical representation suggests that the area of greatest genetic diversity was found in the contact zone close to the 'Bedford gap' (Fig. 2.5a).

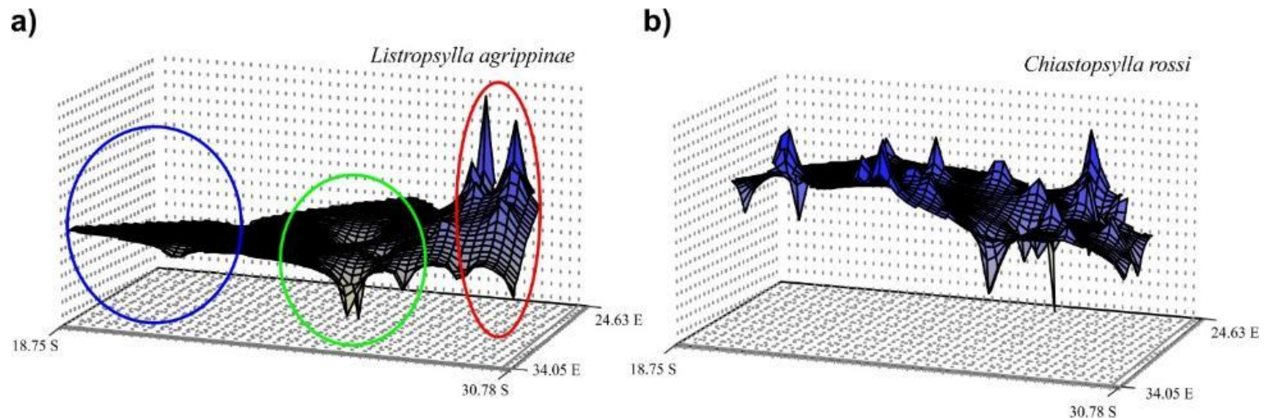


Figure 2.5 Genetic landscape interpolation plot for total distribution of **a**) *L. agrippinae*, and **b**) *C. rossi*.

Fixation index values of mitochondrial COII for *L. agrippinae* were significant at all levels and the highest level of differentiation was recovered among subclusters (60.00 % of variation) (Table 2.3). Fixation index values of nuclear EF1-α for *L. agrippinae* were significant at all levels but in this instance the highest level of variation was recovered within localities (63.26 % of variation) (Table 2.3). Pairwise Φst values among sampling localities showed that the majority of sampling sites are significantly differentiated from each other for both the nuclear and mtDNA data (Table 2.4).

Stellenbosch University https://scholar.sun.ac.za

Table 2.3 Hierarchical analysis of molecular variance (AMOVA) of mitochondrial DNA (COII) and nuclear intron (EF1-α) in two flea species for BAPS subclusters. Fixation index values are given for the hierarchical levels examined: F_{ST}, within localities; F_{CT}, among subclusters; and F_{SC}, among localities within subclusters. Significant ($p < 0.05$) tests are indicated in bold.

		Fixation indices			**Percentage variation**		
	Number of groups	F_{CT}	F_{SC}	F_{ST}	**Among subclusters**	**Among localities within subclusters**	**Within localities**
L. agrippinae							
COII	3	**0.600**	**0.439**	**0.775**	60.00	17.54	22.45
EF1-α	3	**0.239**	**0.169**	**0.367**	23.91	12.83	63.26
C. rossi							
COII	6	**0.513**	**0.382**	**0.699**	51.29	18.63	30.08
EF1-α	6	**0.145**	**0.429**	**0.512**	14.51	36.70	48.79

Table 2.4 Pairwise Φst values among *L. agrippinae* sampled localities for mitochondrial DNA (COII) below and nuclear intron (EF1-α) above the diagonal. Significant values ($p < 0.05$) are highlighted in bold. (Locality codes as in Table 2.1).

Localities	GP	GS	LF	VR	MP	CS	SL	AB	BR	TS	BW	FB	DE
GP		**0.18**	**0.26**	**0.35**	**0.26**	**0.41**	**0.46**	**0.62**	-	**0.55**	**0.31**	**0.43**	**0.38**
GS	**0.18**		**0.27**	0.04	0.00	**0.04**	**0.37**	**0.42**	-	**0.43**	**0.26**	0.05	-0.02
LF	**0.17**	0.09		**0.35**	**0.31**	**0.44**	**0.23**	**0.41**	-	0.21	**0.12**	**0.41**	**0.37**
VR	0.12	**0.23**	0.13		0.03	**0.16**	**0.40**	0.63	-	**0.54**	**0.26**	**0.24**	0.15
MP	**0.63**	**0.88**	**0.72**	**0.50**		**0.04**	**0.37**	**0.40**	-	**0.42**	**0.28**	0.04	-0.03
CS	**0.59**	**0.77**	**0.67**	**0.49**	0.07		**0.52**	**0.78**	-	**0.69**	**0.36**	**0.25**	0.15
SL	**0.65**	**0.75**	**0.70**	**0.60**	**0.71**	**0.71**		**0.30**	-	-0.06	0.03	**0.46**	**0.42**
AB	**0.71**	**0.88**	**0.77**	**0.64**	**0.75**	**0.76**	**0.27**		-	0.55	0.17	**0.84**	0.81
BR	**0.85**	**1.00**	**0.89**	**0.86**	**0.93**	**0.88**	**0.71**	**0.77**		-	-	-	-
TS	**0.77**	**0.92**	**0.82**	**0.77**	**0.85**	**0.82**	**0.51**	**0.66**	**0.90**		-0.06	**0.68**	**0.64**
BW	**0.52**	**0.60**	**0.56**	**0.42**	**0.57**	**0.59**	**0.13**	**0.30**	**0.59**	**0.39**		**0.30**	**0.26**
FB	**0.59**	**0.61**	**0.62**	**0.46**	**0.52**	**0.60**	**0.59**	**0.50**	**0.66**	**0.64**	**0.51**		-0.26
DE	**0.91**	**0.97**	**0.93**	**0.91**	**0.93**	**0.92**	**0.90**	**0.92**	**0.97**	**0.95**	**0.85**	**0.25**	

Stellenbosch University https://scholar.sun.ac.za

3.3.2 *Chiastopsylla rossi*

TCS analysis based on mitochondrial COII data showed a remarkably different genetic pattern for *C. rossi* when compared to *L. agrippinae*. The majority of the localities sampled are characterized by unique divergent haplotypes (Fig. 2.6a and b), and although two distinct genetic clusters separated by an average sequence divergence of 3.01 % (± 0.32) were also obtained for this species (C1 and C2; Fig. 2.6b), these were not congruent with vicariant host patterns (Fig. 2.6). Instead, limited haplotype sharing was evident among distant sampling sites and the majority of populations were characterized by unique/closely related locality specific haplotypes (Fig. 2.6a and b). The lack of clear geographic structure is further supported by the Neighbor-Net tree (Fig. 2.6c) and moreover by the BAPS analyses suggesting at least six distinct subclusters present within *C. rossi* (P = 0.99; log likelihood of optimal partition = -1862.85).

Stellenbosch University https://scholar.sun.ac.za

Chiastopsylla rossi

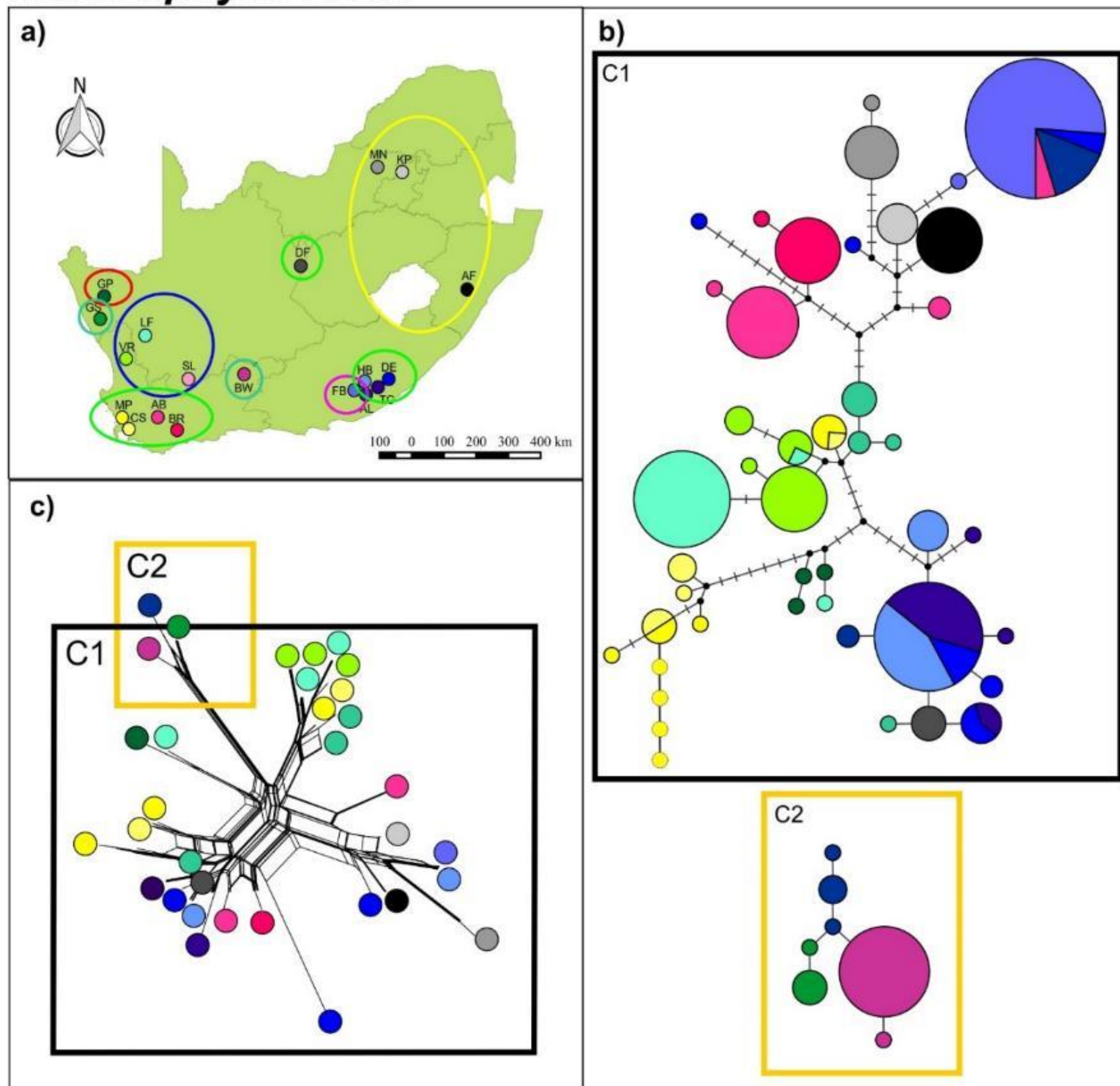


Figure 2.6 a) A map of geographical distribution of sampling localities in South Africa for *C. rossi*. Bayesian analysis of population structure (BAPS) revealed six subclusters indicated by different colours. **b)** Statistical parsimony mitochondrial COII haplotype network colour coded according to sampling locality. Circle size depicts frequency, branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes/mutational steps. Clusters are bound by black (C1) and deep yellow (C2) line boxes. **c)** Neighbour-Net phylogenetic network for *C. rossi* labelled according to haplotype groupings from statistical parsimony results and are bound by black (C1) and deep yellow (C2) line boxes.

Stellenbosch University https://scholar.sun.ac.za

The spatial genetic pattern of the nuclear EF1-α network and the Neighbor-Net tree suggest more lineage sharing across the landscape when compared to the mitochondrial DNA COII data (Figs. 2.6 and 2.7). Similar to *L. agrippinae*, the nuclear data indicated some level of population differentiation (closely related haplotypes confined to single localities). The nuclear TCS analyses suggest a distinct clade comprising individuals from KP, MN and AF and this assemblage is supported as one of the six mtDNA BAPS subclusters. This subcluster is mainly found in the north eastern part of SA, and although it may have some biological meaning, the Neighbor-Net analysis (Fig. 2.7c) show that comparatively the level of differentiation is low. The majority of the subclusters identified by the mtDNA and nDNA data do not corresponded to any of the known biogeographic breaks previously documented for the majority of the host species and in fact are scattered across the landscape (Fig. 2.6a). Mantel tests indicated that there is virtually no isolation by distance within *C. rossi* when calculated for all localities (r = 0.06; P = 0.00).

The scattered distribution coupled to high inter-populational genetic variation in *C. rossi* is supported when the genetic landscape surface plot is considered. In contrast to *L. agrippinae*, *C. rossi* show several genetic discontinuities (peaks) throughout the total sampling distribution (Fig. 2.5b). Fixation index values of the mitochondrial COII for *C. rossi* were significant at all levels and the highest level of differentiation was recovered among subclusters (51.29 % of variation; Table 2.3). The highest level of differentiation for the EF1-α was recovered within localities (48.79 % of variation; Table 2.3). Pairwise Φst values for the mitochondrial and nuclear data also showed significant differentiation among the majority of sampling localities and the highest level of differentiation was again generally detected between KP, MN and AF and the remainder of the populations (Table 2.5).

Stellenbosch University https://scholar.sun.ac.za

Chiastopsylla rossi

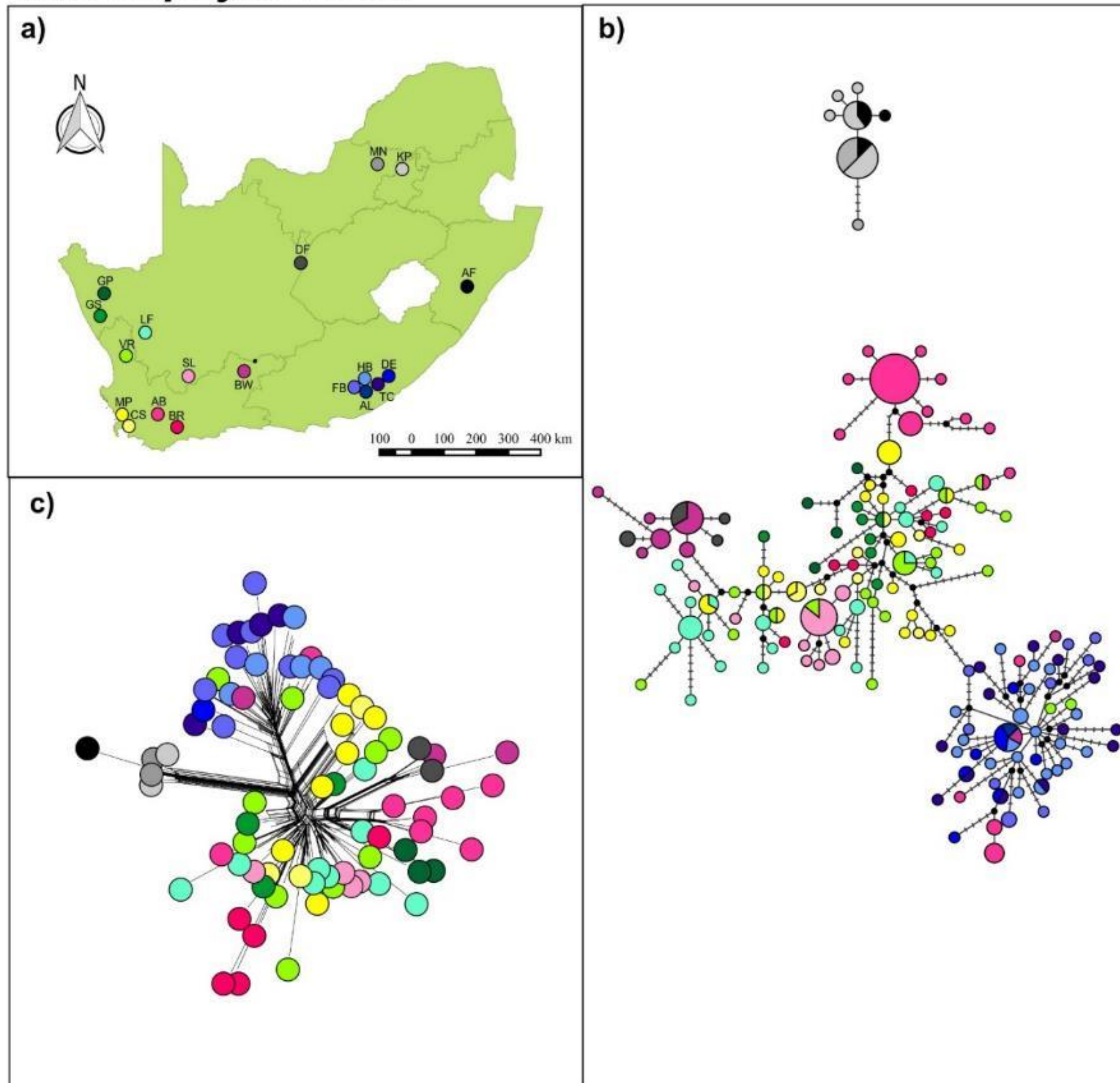


Figure 2.7 a) A map of geographical distribution of sampling localities in South Africa for *C. rossi*. **b)** Statistical parsimony nuclear EF1-α haplotype network colour coded according to sampling locality. Circle size depicts frequency, branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes/mutational steps. **c)** Neighbour-Net phylogenetic network for *C. rossi.*

Table 2.5 Pairwise Φst values among *C. rossi* sampled localities for mitochondrial DNA (COII) below and nuclear intron (EF1-α) above the diagonal. Significant values ($p < 0.05$) are highlighted in bold. (Locality codes as in Table 2.1).

Localities	GP	GS	LF	VR	MP	CS	SL	AB	BR	BW	HB	AL	FB	TC	DE	AF	DF	MN	KP
GP		**0.47**	**0.43**	**0.56**	**0.62**	**0.64**	**0.34**	**0.47**	**0.59**	**0.50**	**0.39**	**0.45**	**0.50**	**0.59**	**0.52**	**0.82**	**0.43**	**0.73**	**0.44**
GS	**0.31**		**0.08**	**0.42**	**0.60**	**0.65**	**0.04**	**0.43**	**0.65**	**0.60**	0.06	**0.22**	**0.49**	**0.73**	**0.52**	**0.84**	0.04	**0.81**	**0.43**
LF	**0.50**	**0.88**		**0.20**	**0.57**	**0.58**	**0.05**	**0.37**	**0.56**	**0.49**	**0.10**	**0.21**	**0.51**	**0.56**	**0.53**	**0.70**	0.03	**0.67**	**0.46**
VR	**0.50**	**0.89**	**0.46**		**0.64**	**0.71**	**0.15**	**0.40**	**0.71**	**0.65**	**0.27**	**0.44**	**0.56**	**0.74**	**0.60**	**0.82**	**0.26**	**0.80**	**0.53**
MP	**0.33**	**0.68**	**0.66**	**0.63**		**0.06**	**0.45**	**0.62**	-0.04	**0.64**	**0.49**	**0.61**	0.02	**0.73**	0.03	**0.76**	**0.55**	**0.72**	**0.23**
CS	**0.25**	**0.78**	**0.75**	**0.74**	0.06		**0.45**	**0.61**	0.00	**0.65**	**0.48**	**0.62**	0.03	**0.77**	-0.02	**0.84**	**0.58**	**0.80**	**0.19**
SL	**0.33**	**0.83**	**0.63**	**0.55**	**0.41**	**0.56**		**0.34**	**0.41**	**0.45**	**0.05**	**0.22**	**0.38**	**0.52**	**0.40**	**0.64**	0.00	**0.59**	**0.34**
AB	**0.24**	**0.52**	**0.55**	**0.51**	**0.17**	**0.18**	**0.32**		**0.60**	**0.51**	**0.36**	**0.41**	**0.57**	**0.57**	**0.58**	**0.72**	**0.34**	**0.69**	**0.51**
BR	**0.36**	**0.98**	**0.87**	**0.88**	**0.39**	**0.55**	**0.77**	0.12		**0.64**	**0.45**	**0.59**	-0.04	**0.79**	-0.05	**0.88**	**0.55**	**0.83**	0.15
BW	**0.54**	**0.93**	**0.92**	**0.92**	**0.78**	**0.87**	**0.89**	**0.64**	**0.99**		**0.46**	**0.52**	**0.55**	0.02	**0.56**	**0.83**	**0.51**	**0.79**	**0.51**
HB	**0.39**	**0.54**	**0.67**	**0.63**	**0.27**	**0.32**	**0.47**	**0.19**	**0.46**	**0.62**		**0.24**	**0.41**	**0.55**	**0.43**	**0.69**	-0.01	**0.63**	**0.35**
AL	**0.19**	**0.74**	**0.77**	**0.76**	**0.38**	**0.46**	**0.58**	**0.20**	**0.62**	**0.86**	**0.35**		**0.54**	**0.59**	**0.56**	**0.73**	**0.18**	**0.69**	**0.49**
FB	**0.57**	**0.99**	**0.94**	**0.94**	**0.80**	**0.89**	**0.92**	**0.61**	**0.99**	**0.99**	**0.76**	**0.49**		**0.63**	0.00	**0.68**	**0.46**	**0.62**	**0.17**
TC	**0.43**	**0.92**	**0.86**	**0.86**	**0.36**	**0.54**	**0.75**	**0.20**	**0.80**	**0.94**	**0.20**	**0.55**	**0.95**		**0.64**	**0.92**	**0.62**	**0.90**	**0.60**
DE	**0.25**	**0.66**	**0.70**	**0.67**	**0.18**	**0.26**	**0.46**	**0.04**	**0.35**	**0.78**	**0.15**	0.10	**0.71**	**0.11**		**0.71**	**0.49**	**0.64**	**0.16**
AF	**0.45**	**0.99**	**0.90**	**0.91**	**0.73**	**0.85**	**0.86**	**0.51**	**0.99**	**0.99**	**0.70**	**0.66**	**0.99**	**0.94**	**0.66**		**0.76**	0.02	**0.48**
DF	**0.23**	**0.98**	**0.88**	**0.88**	**0.24**	**0.44**	**0.77**	0.06	**0.95**	**0.99**	**0.22**	0.50	**0.99**	**0.31**	0.06	**1.00**		**0.71**	**0.40**
MN	**0.47**	**0.99**	**0.93**	**0.93**	**0.78**	**0.87**	**0.90**	**0.58**	**0.98**	**0.99**	**0.76**	**0.72**	**0.98**	**0.94**	**0.71**	**0.99**	**0.99**		**0.37**
KP	**0.29**	**0.99**	**0.89**	**0.89**	**0.64**	**0.77**	**0.81**	**0.43**	**0.98**	**0.99**	**0.61**	**0.42**	**0.97**	**0.91**	**0.51**	**1.00**	**1.00**	**0.98**	

Stellenbosch University https://scholar.sun.ac.za

4. Discussion

Parasites are generally considered to have a high mutation rate, small effective population size and a limited dispersal ability. From this, it is predicted that pronounced spatial genetic structuring will be evident among sampling sites due to reduced gene flow and increased genetic drift (Criscione et al. 2005; Huyse et al. 2005; Nieberding & Morand 2006; Nieberding et al. 2010). Overall, the results presented for the two generalist parasite species in this study conform to these suggestions. Most of the geographically distinct populations sampled show significant differentiation among sampling sites at both the mtDNA and nDNA level (Table 2.4 and 2.5) a scenario attributed to reduced gene flow across the landscape (also see Blouin et al. 1995; Gómez-Díaz et al. 2007).

Fleas are considered generalist parasites and are thus predicted to show very little phylogeographic congruence with host genetic structure (Blouin et al. 1995; Gómez-Díaz et al. 2007; Jones & Britten 2010). The present study provides the first evidence for the converse and, by making use of a comparative approach between two species, highlights the importance of the level of association between parasite and host in shaping genetic diversity across the landscape. The fur flea (*L. agrippinae*) which seems to have a narrower niche breath (i.e. narrower host range) (Segerman 1995; Froeschke et al. 2013; van der Mescht et al. 2013), and spends longer time periods on the host, comprised three distinct mtDNA phylogeographic clades which are markedly congruent to the previously published regional vicariant biogeographic regions based on the patterns obtained in *Rhabdomys* (du Toit et al. 2012), *Micaelamys* (Russo et al. 2010), *Otomys* (Engelbrecht et al. 2011) and *Myosorex* (Willows-Munro & Matthee 2011) (see Fig. A.1). Further support for the notion that the large scale dispersal of *L. agrippinae* is closely dependent on host movement can be obtained from the observation that the 95% confidence interval of the time of divergence between the two major clades found in *L. agrippinae* (5.27 Ma; 95 % HPD interval: 2.31, 9.58 Ma) corresponds reasonably well with the timing of the divergence of the majority of rodent host lineages in the region (Montgelard & Matthee 2012). In contrast, the more host opportunistic nest flea (*C. rossi*) shows, in agreement with other phylogeographic studies on fleas (Gómez-Díaz et al. 2007; Jones & Britten 2010), virtually no congruence in phylogeographic patterns between host and parasite. The distinct clustering of *C. rossi* fleas sampled in the north east of the country (KP, MN, AF: Fig. 2.6a and 2.7) is not clearly depicted by longer branches in

Stellenbosch University https://scholar.sun.ac.za

the Neighbor-Net analyses (Fig. 2.7c) and also not congruent with well documented biogeographic provinces. We argue that although this finding is interesting, it may simply be an artefact of the sampling distances between the various sampling sites.

The documented contact zone between *R. pumilio* and *R. dilectus* at Fort Beaufort (du Toit et al. 2012; Fig. 2.1; Fig. A.1) provides an interesting scenario to further explore host specificity among the two flea species (no contact zones have been described for the other hosts species sampled in this study). In the zone of contact, the two *Rhabdomys* species harbored distinct *L. agrippinae* mtDNA lineages represented by L1 (found exclusively on *R. pumilio*) and L2 (found exclusively on *R. dilectus*; Fig. 2.3b). In sharp contrast, the host opportunistic *C. rossi* show a high level of haplotype sharing among closely related localities in the same region (Figs. 2.6 and 2.7). *Rhabdomys* species is considered the principle hosts of *L. agrippinae* and *M. unisulcatus* is considered to be an auxiliary host, whereas *C. rossi* seems to have a larger host range (Segerman 1995; Froeschke et al. 2013; van der Mescht et al. 2013). Albeit based on a very small sample size, the tighter host association displayed by *L. agrippinae* might explain why we found distinct mtDNA lineages for *L. agrippinae* on *Rhabdomys* spp., but no differentiation in *C. rossi* in the contact zone.

When the intraspecific phylogeographic structures of *L. agrippinae* and *C. rossi* are compared to each other additional discrepancies are evident. First, our data provide supplementary support for the hypothesis that more generalist parasites will show higher levels of genetic diversity when compared to more specific parasites (Martinů et al. 2014; Table 2.2). The connectivity among sampling sites is also markedly different between the two flea species. The more host specific fur flea *L. agrippinae*, displays lower inter-populational divergence within clades (Fig. 2.3) when compared to the more host opportunistic nest flea, *C. rossi* (Fig. 2.6). The same trend is evident when the nDNA data are considered (Figs. 2.4 and 2.7). The higher level of inter-populational divergences of *C. rossi* are best reflected by the large number of site changes among locality specific haplotypes (Fig. 2.6) and the numerous peaks on the landscape interpolation plots (Fig. 2.5b). These findings could be interpreted to be the direct result of differences in the dispersal abilities of the two fleas and can best be ascribed to differences in host association. Both parasites

Stellenbosch University https://scholar.sun.ac.za

show isolation by distance but in both cases the correlation is extremely weak (as indicated by the r values). In the case of the host specific fur flea, *L. agrippinae*, host movement provides hitchhiking possibilities within clades resulting in a higher level of connectivity among sampling sites. In the host opportunistic nest flea, *C. rossi*, the parasite has less opportunity to spread via host movement, resulting in a pattern of distantly related haplotypes at most sites. At specific geographic sites, it is possible that populations experience high levels of genetic drift that is homogenizing the locality specific signals (Criscione et al. 2005; Huyse et al. 2005; Nieberding & Morand 2006; Nieberding et al. 2010). Long distance dispersals of the host opportunistic nest flea studied herein is a rare event and occurs in a random manner utilizing a wide niche breath (in the absence of strong isolation by distance), culminating in a pattern where some shared haplotypes are found on opposite ends of the geographic scale (for example see haplotype sharing between the group of localities with different shades of blue (FB, AL, TC, HB, DE) and geographically distant localities (BR and also DF; Fig. 2.6)).

From our study it was also evident that *L. agrippinae* occurred at a lower overall prevalence than *C. rossi* and that this pattern was apparent at most localities where the distribution of the two species overlaps. The same trend in prevalence and abundance for *L. agrippinae* and *C. rossi* was recorded in previous studies in SA (van der Mescht et al. 2013; Froeschke et al. 2013; Matthee et al. 2007). The two flea species differ in terms of body length with *L. agrippinae* being larger (on average 3650 µm) compared to *C. rossi* (on average 1750 µm) (Matthee et al. 2010b; also supported by body size index in van der Mescht et al. (2013)). Studies on free-living taxa and more recently on ectoparasitic mites of small mammals recorded a negative relationship between body size and abundance (Gaston & Blackburn 1995; Blackburn & Gaston 1999; Krasnov et al. 2013b). In the latter study it was suggested that the pattern may be due to higher host-induced mortality (grooming) associated with larger bodied ectoparasites (Krasnov et al. 2013b). *Listropsylla agrippinae* is almost twice the size of *C. rossi* and it is possible that host grooming (allo- and autogrooming) resulted in lower *L. agrippinae* prevalence. Interestingly, in the present study the difference in body size between the two flea species seems to support their level of host specificity. Studies on microbes and diatoms and more recently on ectoparasitic mites recorded a negative relationship between body size and niche breadth (Passy & Larson 2011; Passy 2012; Krasnov et

Stellenbosch University https://scholar.sun.ac.za

al. 2013b). It appears that smaller bodied species are more adaptable to environmental fluctuations which results in a larger niche breath or in the case of parasites a larger host range. This is in contrast to larger species that have a smaller niche breath/host range due to narrower tolerance levels (Passy & Larson 2011; Passy 2012; Krasnov et al. 2013b).

If prevalence is positively correlated to abundance in fleas (Krasnov et al. 2004, 2005c), then it could be argued that the higher abundance (and niche breath) of the more host opportunistic nest flea, *C. rossi*, on the hosts will facilitate dispersal among sampling sites. This pattern is however contrary to what we found in the present study where *C. rossi* is more structured between sampling sites when compared to *L. agrippinae*. One possible explanation for this may relate to differences in the effective population sizes between the two flea species. A higher genetic diversity, as found in *C. rossi*, is expected for populations with larger effective population sizes (Blouin et al. 1995; Huyse et al. 2005) and it is furthermore reasonable to predict that nest fleas will predominantly have higher abundances in the nests of their hosts (Krasnov 2008; Jones & Britten 2010). Given the nest bound nature of *C. rossi*, the number of dispersing individuals on the hosts may thus not be large enough to overcome the effects of drift when introduced into the nests of the hosts elsewhere (containing a large population of local genotypes in the new environment). In contrast, for *L. agrippinae*, the comparatively lower effective population sizes can facilitate the signature of more haplotype sharing among localities within clades.

In the light of re-emerging flea borne diseases worldwide, it is important to have a thorough understanding of the mechanisms that are involved in shaping flea distribution and movement (Bitam et al. 2010). From our study it is evident that host association (microhabitat preference and host specificity) plays an important role in flea dispersal and subsequent gene flow within and between geographic locations. This study also provides the first evidence of congruent phylogeographic vicariant patterns between a generalist parasitic flea and its hosts. Unfortunately our conclusions are based on a single study comprising two distinct life histories. More parasite taxa and gene fragments need to be evaluated in order to formulate stronger hypotheses in this regard.

Stellenbosch University https://scholar.sun.ac.za

Chapter 3

A genetic perspective on the taxonomy and evolution of the medically important flea, *Dinopsyllus ellobius* (Siphonaptera: Dinopsyllinae), and the resurrection of *Dinopsyllus abaris*

**Biological Journal of the Linnaean Society 116: 541-557*

Stellenbosch University https://scholar.sun.ac.za

1. Introduction

The fauna on the southern African terrestrial landscape is characterised by a rich and complex evolutionary history where speciation is often linked to habitat vicariance and climate driven diversification among free living lineages (Matthee & Flemming 2002; Smit et al. 2007; Tolley et al. 2009; Russo et al. 2010; du Toit et al. 2012). Speciation processes for parasitic taxa are characterized by an added complexity surrounding host association and host life history (Krasnov & Poulin 2010; du Toit et al. 2013b; Engelbrecht et al. 2014). Since parasites are also generally characterized by a low intrinsic dispersal potential, small effective population sizes, short generation times (Huyse et al. 2005; Whiteman & Parker 2005) and faster evolutionary rates (Paterson & Banks 2001; Nieberding et al. 2004), they often show higher levels of intraspecific divergence when compared to their hosts (Hafner et al. 1994; Moran et al. 1995; Page et al. 1998; Paterson et al. 2000; Whiteman et al. 2007; du Toit et al. 2013b). Understanding the evolutionary mechanisms responsible for the geographic diversification in parasites is thus not trivial.

Apart from deciphering the reasons for the divergence among parasitic evolutionary lineages, it is also predicted that parasites harbour a large proportion of cryptic diversity (Nadler & Pérez-Ponce de León 2011; Poulin 2011; Perkins et al. 2011; du Toit et al. 2013b; Engelbrecht et al. 2014). These previously hidden taxonomic entities highlight the importance of further investigations to decipher true biodiversity (Poulin 2011). Along these lines, a high intraspecific phylogenetic diversity has been recorded for fleas (De la Cruz & Whiting 2003; Brinkerhoff et al. 2011; Lawrence et al. 2014; Lin et al. 2014) and vacillations in taxonomic treatments of the exclusively Afrotropical subgenus *Dinopsyllus* is no exception (e.g. *D. lypusus* and *D. longifrons*; *D. ellobius* and *D. abaris*) (Hopkins & Rothschild 1966; de Meillon et al. 1961; Segerman 1995). Much of the subgeneric controversies can be blamed on the absence of genetic data coupled to subtle differences in morphological characters (de Meillon et al. 1961). In southern Africa, the subgenus is represented by *Dinopsyllus ellobius* and is regarded as a host generalist ectoparasite of small mammals occurring primarily on gerbils (*Gerbilliscus* spp.), vlei rats (*Otomys* spp.), multimammate mice (*Mastomys* spp.) and striped mice (*Rhabdomys* spp.) (Segerman 1995), while secondary hosts include the karoo bush rat (*Myotomys unisulcatus*), red rock rat (*Aethomys chrysophilus*), namaqua rock rat (*Micealamys namaquensis*) and the white-tailed mouse

Stellenbosch University https://scholar.sun.ac.za

(*Mystromys albicaudatus*) (Segerman 1995). This life history pattern is in contrast to certain flea taxa that have narrower host ranges (Krasnov 2008) or other parasitic insects such as sucking lice (Anoplura) that are regarded as highly host specific (Marshall 1981). In general, the immature stages of fleas occur in the nest of the host while the adult stages occur along a gradient from being more associated with the host body (also referred to as "fur "fleas) to being more associated with the nest or burrow of the host ("nest" fleas) (Ioff 1941; Marshall 1981; Krasnov 2008). In this regard *D. ellobius* has been recorded on the host body and in the nest, though flea abundances seem to be higher in the nest when compared to the host body (de Meillon et al. 1961).

The distribution of *D. ellobius* is confined to the southern African subregion and in South Africa (SA) it appears to be restricted to more mesic biomes such as those found along the coastal belt and in-land grassland habitats (Hopkins & Rothschild 1966). Following morphological descriptions, Jordan (1930) recognized two distinct sub-species (*D. e. ellobius* and *D. e. abaris*), based on the size and shape of sternite VIII in males. Subsequently, using the same distinguishing character, the two lineages were regarded as full species by Hopkins & Rothschild (1966). However, the two taxa overlap in geographic ranges and have reportedly similar host associations, and this single subtle morphological difference is thus also regarded by some as merely geographic variation within the same species, *D. ellobius* (de Meillon et al. 1961; Segerman 1995). Since *D. ellobius* has been implicated as potentially one of the most significant plague (*Yersinia pestis*) vectors in the region (Ingram 1927; de Meillon et al. 1961), the correct identification of this taxon is not only important in terms of biodiversity conservation (Bickford et al. 2007; Cook et al. 2008), but also could have bearing on the epidemiology of zoonotic diseases (Bidochka et al. 2001; Koufopanou et al. 2001; Venturi et al. 2004; Poulin 2011).

The importance of adding genetic data in deciphering the taxonomy of fleas cannot be disputed and this approach has the added advantage to also provide novel insights into the evolutionary history of the species concerned (see for example Whiting 2002; Luchetti et al. 2005, 2007; Whiting et al. 2008; Lawrence et al. 2014). Although cryptic divergence as a result of host vicariance has recently been implicated as a contributing factor towards diversification among generalist flea lineages (Chapter 2 – van der Mescht et al. 2015a), the majority of evolutionary

Stellenbosch University https://scholar.sun.ac.za

studies to date show very limited or complex relationships between parasite and host genetic structures (Criscione et al. 2005; Gómez-Díaz et al. 2007; Barret et al. 2008; Hoberg & Brooks 2008; Nieberding et al. 2008; Cangi et al. 2013; du Toit et al. 2013a; Fermino et al. 2013; Poretta et al. 2013b; Martinů et al. 2014). Given the wide host range, the previous description of two species (*D. ellobius* and *D. abaris* by Hopkins & Rothschild 1966) based on a single distinguishing morphological character, and the potential effect of host and habitat vicariance on the evolution of fleas in the region, we predict that *D. ellobius* (as described by de Meillon et al. 1961; Segerman 1995) may contain cryptic diversity in the region. We sampled a large number of *D. ellobius* specimens from a wide range of hosts throughout much of the distribution of the species and utilized molecular and morphological data to provide a better understanding of the taxonomy of *D. ellobius* and also the mechanisms that potentially played a role in diversification of this lineage.

2. Materials and Methods

2.1 Taxon and gene sampling

Sherman-type live traps baited with a peanut butter and oats mixture were used to sample small mammal hosts at 31 localities during austral spring and summer of 2011-2013 and attention was given to sample throughout most of the documented range of the flea species in SA (Fig. 3.1; Table B.3). All hosts captured were placed in a plastic bag and euthanized with intraperitoneal injection of 0.2-0.4 ml sodium pentobarbitone (200 mg/kg; ethical approval was granted by the Animal Ethics Committee of Stellenbosch University: SU-ACUM11-00004). Fleas were sampled using a standardized method that included brushing of the host pelage several times over a white plastic tray using separate toothbrushes for each host species at each locality. Individual fleas were placed in separate tubes filled with 96 % ethanol. Before DNA extraction, morphologically defined *D. ellobius* individuals were identified using a Leica stereoscopic microscope (Leica Microsystems, Wetzlar, Germany) and the taxonomic key of Segerman (1995). Following DNA extraction, the exoskeletons of all extracted fleas were mounted (for detail on method see van der Mescht et al. 2013) and identified using a Leica DM 3000 light microscope (Leica Microsystems, Wetzlar, Germany) and the taxonomic keys by Segerman (1995) with additional reference to Hopkins & Rothschild (1966). The reference collection together with voucher specimens of the flea species will be housed in the Museum of the department of Conservation Ecology and Entomology at

Stellenbosch University and duplicate sets will also be available at the National Flea Collection in Johannesburg (both in SA).

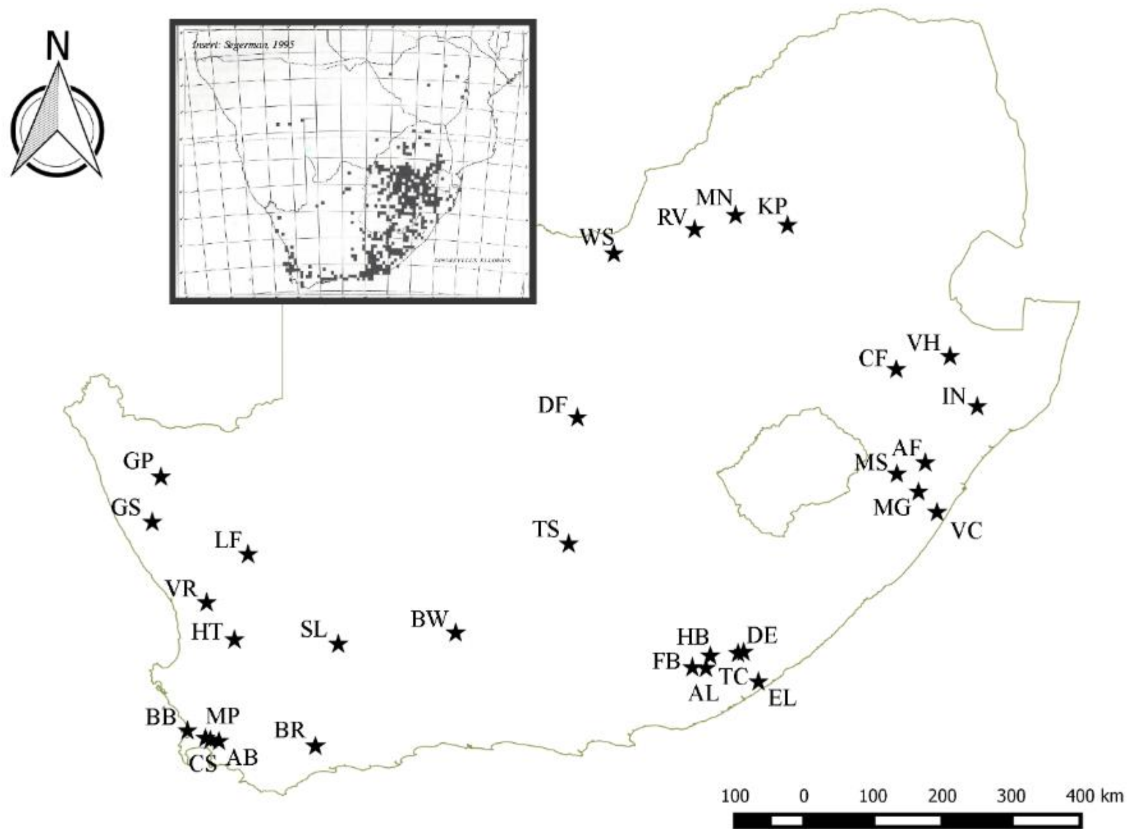


Figure 3.1 Localities from which small mammal hosts were sampled with locality codes as in Table B.3. (Insert: *D. ellobius* distribution from Segerman 1995)

Stellenbosch University https://scholar.sun.ac.za

2.2 DNA extraction, amplification and sequencing

Total genomic DNA was extracted with a Qaigen, DNeasy® Blood and Tissue kit (Qaigen, Valencia, CA, USA) following the protocol supplied. Whole flea specimens were placed in the extraction buffer containing Proteinase K (600 mAU/ml solution), and digested at 56°C overnight. After digestion, each flea exoskeleton was removed and stored in 70 % ethanol for identification purposes (see above). Polymerase Chain Reactions (PCR) were performed using KAPA Taq DNA polymerase (Kapa Biosystems, Wilmington, Massachusetts, USA) and published primers (Whiting 2002; Table B.4). Mitochondrial Cytochrome c oxidase subunit II (COII) and one intron of the nuclear Elongation factor 1 alpha (EF1-α) gene regions were amplified (GenBank accession numbers KR707331-707617) using a GeneAmp® PCR 2700 thermal cycler (Applied Biosystems, Foster City, CA, USA). Mitochondrial COII markers are frequently used to address taxonomic and phylogenetic questions in insects (Hillis & Dixon 1991; Caterino et al. 2000) and seem to be the preferred choice for fleas (Luchetti et al. 2005, 2007; Whiting et al. 2008; Lawrence et al. 2014; Lin et al. 2014). Sequences were analysed on an ABI 3730 XL DNA analyser (Applied Biosystems) after using BigDye termination chemistry (version 3.1, Applied Biosystems).

2.3 Sequence processing and alignment

Sequences for each gene fragment were edited and aligned using BioEdit Sequence Alignment editor 7.2.5 (Hall 2005). Mitochondrial and nuclear sequences were translated into amino acids using EMBOSStranseq (www.ebi.ac.uk/Tools/st/emboss_transeq) to confirm functionality. Heterozygous positions in the nuclear fragments were resolved in DNASP 5 (Librado & Rozas 2009) using PHASE 2.1.1 (Stephens et al. 2001; Stephens & Scheet 2005). The algorithm was run for 1000 generations with a thinning interval of 1 and burn-in of 100 generations. Phases with a 0.9 probability or higher were considered resolved and the analysis was performed three times to test for convergence (Stephens et al. 2001).

Stellenbosch University https://scholar.sun.ac.za

2.4 Phylogenetic analyses

Evolutionary relationships among individuals were investigated by constructing statistical parsimony haplotype networks with 95 % confidence intervals in TCS v1.21 (Clement et al. 2000). To obtain a better evolutionary perspective on the associations between geographic clusters the Neighbour-Net method (Bryant & Moulton 2004) as implemented in SplitsTree v4.5 (Huson & Bryant 2006) was used. For the latter, the best-fit models for un-partitioned mitochondrial COII and nuclear EF1-α alignments were calculated in jModeltest v2.1.4 (Posada 2008; Darriba et al. 2012) under the AICc (Akaike 1973; Posada & Buckley 2004). The resultant optimal models were used to obtain corrected sequence distances among gene fragments using PAUP* v4.0b10 (Swofford 2000). Analysis of molecular variance (AMOVA; Excoffier et al. 1992) and pairwise Φst statistics between sampled populations were performed in ARLEQUIN v3.5.1.2 (Excoffier & Lischer 2010). Only sampling localities with more than 5 individuals were included. AMOVA was also used to test for differentiations among groups as defined by the mtDNA haplotype networks.

2.5 Divergence time estimation

Divergence dates between mtDNA evolutionary lineages were obtained by using a relaxed exponential Bayesian molecular clock, the mtDNA data only, HKY + I + G model as prior and the birth-death process (Kendall 1948) as the tree prior in BEAST v1.8.0 (Drummond & Rambaut 2007; Drummond et al. 2012). Fossil calibrations were used employing exponential priors with hard minimum and soft maximum bounds. Fossil calibration points were obtained from Zhu et al. (2015) (1) Age of *Pulex* (*Pulex larimerius* that was found in Dominican amber; Lewis & Grimaldi, 1997) set as a hard minimum of 15 Mya and soft maximum of 20 Mya (Iturralde-Vinent & MacPhee, 1996). (2) Age of *Rhopalopsyllus* as implemented by Zhu et al. (2015) and also set at a hard minimum of 15 Mya and soft maximum of 20 Mya and (3) Age of *Plaeopsylla* found in Baltic amber 35-40 Mya (Dunlop & Giribet 2003; Ritzkowski 1997; Weitschat & Wichard 2010). Two species per outgroup were used for each of the three calibration points in our analysis (Table B.5). Test runs were performed to ensure that priors had sufficient ESS values (>200) and they were continuously assessed in Tracer v1.5 (Rambaut & Drummond 2007). Final MCMC simulations ran for 50 million generations, sampling every 5 000 generations. Convergence and mixing were assessed in Tracer v1.5 (Rambaut & Drummond 2007) and burn-in discarded. A maximum lineage

Stellenbosch University https://scholar.sun.ac.za

credibility tree was produced in TreeAnnotator v1.8.0 (Drummond & Rambaut 2007; Drummond et al. 2012) following the 50 % majority rule.

2.6 Morphological analysis

2.6.1 *Size*

Male and female fleas are known to differ in size (Krasnov et al. 2003a) and were thus analysed separately. To avoid missing data, only male (*N*=52) and female (*N*=78) individuals that had no damage to any part of the body were measured. The head, coxa, femur, and tibia length of the left hind leg (Fig. A.2) were measured using a camera mounted on a Leica DM 3000 light microscope (Leica Microsystems, Wetzlar, Germany) and Leica software v3.1.0 (Leica Microsystems, Wetzlar, Germany) (see Tripet et al. 2002; Krasnov et al. 2003a; van der Mescht et al. 2013). Raw size measurements were log transformed and an analysis of variance (ANOVA) performed with the *aov* function of the *stats* package in R v3.1.1 (R Foundation for Statistical Computing, Vienna, AT, AU) to investigate differences between characters measured. Co-variation between the characters was investigated by an analysis of co-variance (ANCOVA). A principal component analysis (PCA) was performed by the *prcomp* function of the *stats* package. Results were visualized using the *ggbiplot* package in R v3.1.1 (R Foundation for Statistical Computing, Vienna, AT, AU). The continuous variables were centred and scaled to standardize the variables prior to the application of PCA. Boxplots were constructed for each sex by using the *plot* function in R v3.1.1 (R Foundation for Statistical Computing, Vienna, AT, AU).

2.6.2 *Characters*

The key to Siphonaptera of southern Africa was first used to assess morphological characters of the subfamily Dinopsyllinae and subsequently to classify specimens up to species level (Segerman, 1995) (Table B.6). Then, we assessed the consistency of the morphological character (variation in male sternite VIII) used to distinguish between *D. ellobius* and *D. abaris* as in Hopkins & Rothschild (1966) (Table B.6). A total of 96 female (28 D1 and 68 D2) and 55 male (14 D1 and 41 D2) individuals were assessed for the latter.

Stellenbosch University https://scholar.sun.ac.za

3. Results

3.1 Characteristics of taxon and gene sampling

A total of 830 small mammal individuals comprising 23 different host species were trapped (Table B.7). *Dinopsyllus ellobius* was found on 11 of the 23 host species trapped and these originated from 25 of the 31 localities. The species were mainly absent from the dryer western interior region of SA (specifically localities DF, TS, BW, SL, LF, GS, GP, codes as listed in Table 3.1), confirming earlier suggestions on the extent of the distribution (Fig. 3.1; Hopkins & Rothschild 1966). Mitochondrial COII (567 bp) sequence data were generated for 151 *D. ellobius* specimens from 19 of the localities (Table 3.1). To test for congruence between gene trees the mtDNA data set was augmented with data generated for the nuclear EF1-α gene. Due to difficulties in amplification success we only managed to obtain a geographic representative subsample of 68 sequences derived from 13 of the 19 sampled localities (Table 3.1). The analyses of the nuclear EF1-α resulted in higher nucleotide and haplotype diversity compared to the mitochondrial COII data (Table 3.2).

Stellenbosch University https://scholar.sun.ac.za

Table 3.1 Geo-referenced localities with codes and number of fleas sequenced for each gene from which *D. ellobius* were collected.

Province	Locality	Code	Geographic coordinates	*D. ellobius* (n)	
				COII	EF1-α
Western Cape	Anysberg	AB	-33.97 S 18.97 E	6	3
	Blaauwberg	BB	-33.80 S 18.46 E	8	5
	Hottentotsholland	HT	-32.33 S 19.22 E	8	6
	Vanrhynsdorp	VR	-31.73 S 18.77 E	7	3
Eastern Cape	Alice	AL	-32.79 S 26.85 E	12	5
	Dohne	DE	-32.53 S 27.46 E	14	5
	East London	EL	-33.01 S 27.70 E	6	0
	Fort Beaufort	FB	-32.78 S 26.63 E	8	3
	The Croft	TC	-32.55 S 27.37 E	13	0
Gauteng	Kaalplaas	KP	-25.63 S 28.17 E	3	3
	Rietvlei	RV	-25.70 S 26.27 E	2	0
North West	Mooinooi	MN	-25.47 S 27.33 E	5	5
	Woodside	WS	-26.09 S 25.36 E	8	5
KwaZulu-Natal	Chelmsford	CF	-27.96 S 29.93 E	5	0
	Inkunzi Lodge	IN	-28.56 S 31.24 E	16	10
	Mt Gilboa	MG	-29.94 S 30.29 E	15	11
	Mt Shannon	MS	-29.65 S 29.94 E	5	4
	Vernon Crookes	VC	-30.27 S 30.59 E	6	0
	Vryheid	VH	-27.75 S 30.80 E	4	0
5 provinces	**19 localities**			**151**	**68**

Stellenbosch University https://scholar.sun.ac.za

Table 3.2 Summary statistics for the mitochondrial COII and nuclear EF1-α markers sequenced from two lineages in South Africa, including the number of sequences/alleles, total number of singleton haplotypes, fragment length (bp), number of polymorphic sites (*P*), nucleotide diversity (π) and haplotype diversity (*h*).

	Sequences/ Alleles	Total haplotypes	Singleton haplotypes	bp	*P*	*π* ± SD	*h* ± SD
COII							
D1 (*D. abaris*)	109	37	23	567	41	0.011 ± 0.002	0.913 ± 0.016
D2 (*D. ellobius*)	42	20	12	567	35	0.018 ± 0.002	0.941 ± 0.017
EF-1α							
D1 (*D. abaris*)	46/92	78	69	522	104	0.018 ± 0.004	0.995 ± 0.003
D2 (*D. ellobius*)	22/44	40	37	522	73	0.020 ± 0.004	0.995 ± 0.007

3.2 Phylogenetic reconstructions

The parsimony network analysis based on mitochondrial data resulted in the detection of two distinct genetic clades (hereafter referred to as D1 (*D. abaris*) & D2 (*D. ellobius*); Fig. 3.2) that could not be connected with 95 % confidence. The two clades are similarly supported by distant clustering on the Neighbour-Net analyses that incorporated HKY+I+G (COII) and GTR+I+G (EF1-α) corrections (Fig. 3.2). An average mtDNA sequence divergence of 5.53 % (± 1.78; Table 3.3) was detected between the assemblages. Nucleotide and haplotype diversity values were lower for D1 (*D. abaris*) compared to D2 (*D. ellobius*) (Table 3.2). Fixation index values of mitochondrial data were significant at all levels and the highest level of differentiation was recovered among D1 (*D. abaris*) and D2 (*D. ellobius*) lineages (71.95 %; Table 3.4). These two main mitochondrial lineages diverged during the late Miocene 6.42 Mya (95 % HPD interval: 1.88, 12.82 Mya) (Fig. A.3).

Stellenbosch University https://scholar.sun.ac.za

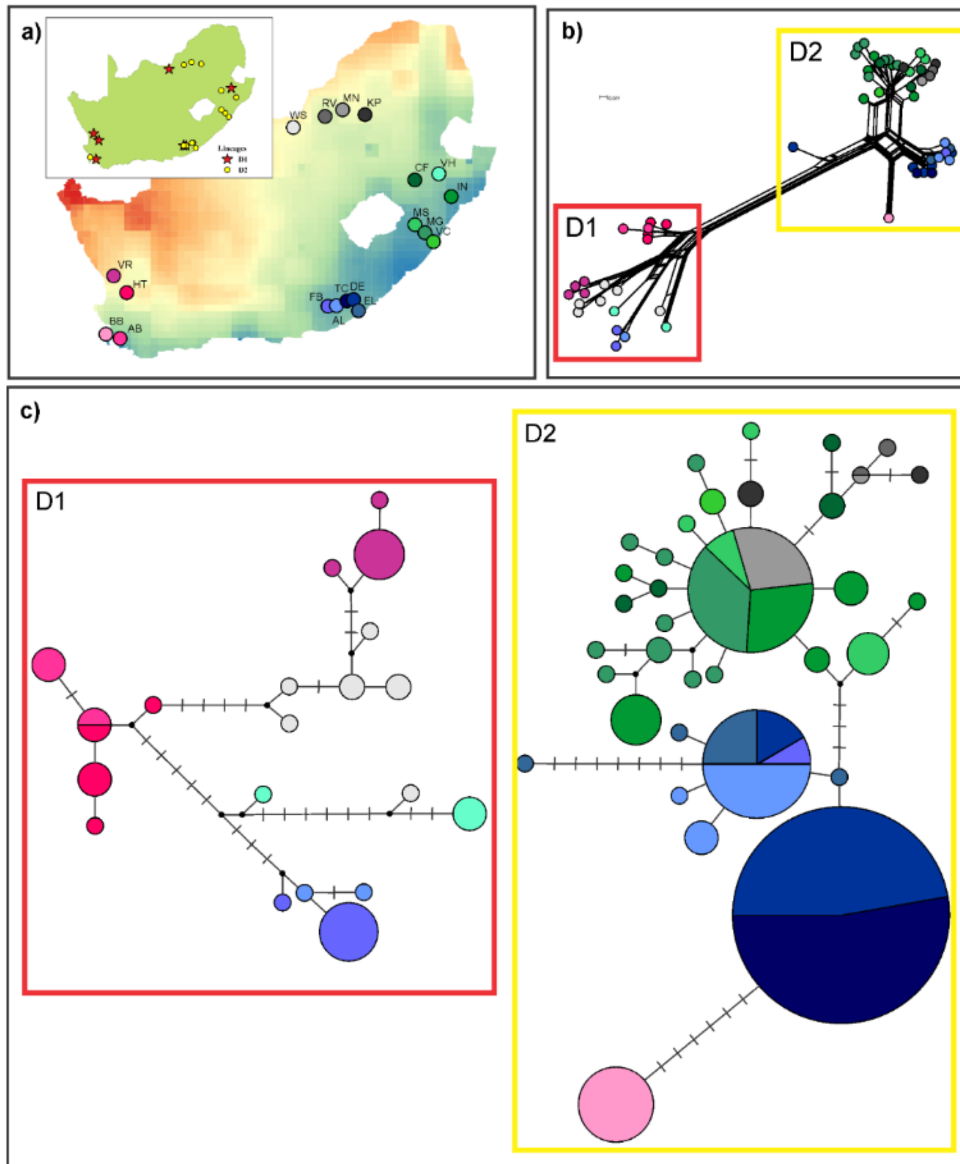


Figure 3.2 a) Geographical distribution of sampling localities in South Africa for *D. ellobius* plotted on a map of mean rainfall. Colour spectrum of the map vary from low values in red to high values in blue. (Insert: Map indicating where two lineages were found) **b)** Neighbour-Net phylogenetic network of mitochondrial COII for *D. ellobius* bound by red (D1 (*D. abaris*)) and deep yellow (D2 (*D. ellobius*)) line boxes. **c)** Statistical parsimony mitochondrial COII haplotype network colour coded according to sampling locality to represent relative frequencies. Circle size depicts frequency; branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes. The two lineages are bound by red (D1 (*D. abaris*)) and deep yellow (D2 (*D. ellobius*)) line boxes.

Stellenbosch University https://scholar.sun.ac.za

Table 3.3 Pairwise genetic divergence values for mitochondrial COII and nuclear EF1-α within and between D1 (*D. abaris*) and D2 (*D. ellobius*) mitochondrial lineages**.**

Sequence divergence	**COII (% ± SD)**	**EF1-α (% ± SD)**
Between D1 (*D. abaris*) and D2 (*D. ellobius*)	5.53 ± 1.78	1.98 ± 0.82
Within D1 (*D. abaris*)	1.82 ± 1.69	1.80 ± 0.78
Within D2 (*D. ellobius*)	1.12 ± 1.74	2.04 ± 0.85

Table 3.4 Hierarchical analysis of molecular variance (AMOVA) of mitochondrial COII and nuclear EF1-α of two mitochondrial lineages. Fixation index values are given for the hierarchical levels examined: F_{ST}, within populations; F_{CT}, among lineages; and $F_{SC,}$ among localities within lineages. Significant ($p < 0.05$) tests are indicated in bold.

		Fixation indices			**Percentage variation**		
	Number of groups	F_{CT}	F_{SC}	F_{ST}	**Among lineages**	**Among localities within lineages**	**Within localities**
Dinopsyllus							
COII	2	**0.720**	**0.807**	**0.946**	71.95	22.63	5.42
EF1-α	2	0.027	**0.337**	**0.319**	2.70	30.99	66.31

Apart from a single haplotype that is shared between two geographically close localities of lineage D1 (*D. abaris*) (AB & HT; Fig. 3.2c), all remaining sampled localities in this clade are characterized by unique distantly related haplotypes. Most localities are separated by a large number of nucleotide changes among haplotypes contributing to an average mitochondrial sequence diversity of 1.82 % (±1.69; Fig. 3.2c; Table 3.3) for this clade. In contrast, clade D2 (*D. ellobius*) is characterized by a fairly large amount of haplotype sharing between localities and an average mitochondrial sequence diversity of 1.12 % (±1.74; Fig. 3.2c; Table 3.3). Only one geographically distant locality where D2 (*D. ellobius*) was sampled (BB, Fig. 3.2c) consisted of a unique haplotype. Pairwise Φst values among sampling localities showed that the majority of sampling localities are significantly differentiated for both lineages (Table 3.5).

Table 3.5 Pairwise Φst values among sampled localities for mitochondrial COII below and nuclear EF1-α above the diagonal. Significant values ($p < 0.05$) are highlighted in bold. (Locality codes as in Table 3.1).

		D2 (*D. ellobius*)											D1 (*D. abaris*)				
		DE	BB	CF	EL	MG	MS	IN	MN	VC	AL	TC	VR	AB	FB	WS	HT
D2 (*D. ellobius*)	DE		**0.10**			0.01	0.02	0.00	0.01		0.00		0.00	0.00	0.00	0.05	0.00
	BB	**0.83**				**0.10**	**0.12**	**0.10**	**0.11**		**0.12**		0.11	**0.11**	**0.11**	**0.16**	**0.10**
	CF	**0.53**	**0.64**														
	EL	**0.51**	**0.65**	0.15													
	MG	**0.41**	**0.47**	**0.10**	**0.14**			0.01	**0.02**		0.03		**0.01**	0.01	0.01	**0.06**	0.01
	MS	**0.49**	**0.60**	0.05	0.10	-0.01		0.02	0.03		0.04		0.02	0.02	0.02	0.07	0.02
	IN	**0.44**	**0.49**	**0.14**	**0.18**	0.06	0.05		0.01		**0.02**		0.00	0.00	0.00	**0.05**	0.00
	MN	**0.70**	**0.85**	**0.35**	**0.39**	0.06	0.17	**0.17**			0.03		0.01	0.01	0.01	**0.06**	0.01
	VC	**0.64**	**0.77**	**0.29**	**0.33**	**0.25**	**0.25**	**0.28**	**0.53**								
	AL	**0.55**	**0.67**	**0.28**	0.01	**0.24**	**0.23**	**0.27**	**0.48**	**0.43**			0.02	0.02	0.01	**0.07**	**0.02**
	TC	0.07	**1.00**	**0.73**	**0.73**	**0.53**	**0.69**	**0.55**	**0.89**	**0.83**	**0.73**						
D1 (*D. abaris*)	VR	**0.64**	**0.75**	**0.31**	**0.34**	**0.26**	**0.26**	**0.29**	**0.53**	**0.47**	**0.43**	**0.81**		0.00	0.00	0.06	0.00
	AB	**0.62**	**0.74**	**0.26**	**0.30**	**0.22**	0.21	**0.26**	**0.49**	**0.43**	**0.40**	**0.81**	**0.44**		0.00	0.06	0.00
	FB	**0.73**	**0.87**	**0.45**	**0.47**	**0.35**	**0.40**	**0.38**	**0.67**	**0.60**	**0.54**	**0.90**	**0.60**	**0.57**		0.06	0.00
	WS	**0.46**	**0.54**	0.08	**0.13**	**0.09**	0.04	**0.12**	**0.30**	**0.25**	**0.24**	**0.62**	**0.27**	**0.22**	**0.38**		0.05
	HT	**0.52**	**0.61**	0.16	**0.21**	**0.15**	**0.12**	**0.19**	**0.38**	**0.33**	**0.31**	**0.68**	**0.34**	0.14	**0.45**	**0.14**	

The spatial genetic structure of the nuclear data was less pronounced when compared to the results from the mitochondrial data (Fig. 3.3). When the mtDNA clades are used to assign clade structure to the nuclear data, the sequence divergence among clades and the sequence diversity within clades are overlapping in value, supporting the lack of congruence between the two data sets (Table 3.3). Fixation index values of nuclear data for *D. ellobius* were significant at two levels, but as expected given the lack of geographic resolution, the highest level of variation was recovered within localities (66.31 % of variation) (Table 3.4).

Stellenbosch University https://scholar.sun.ac.za

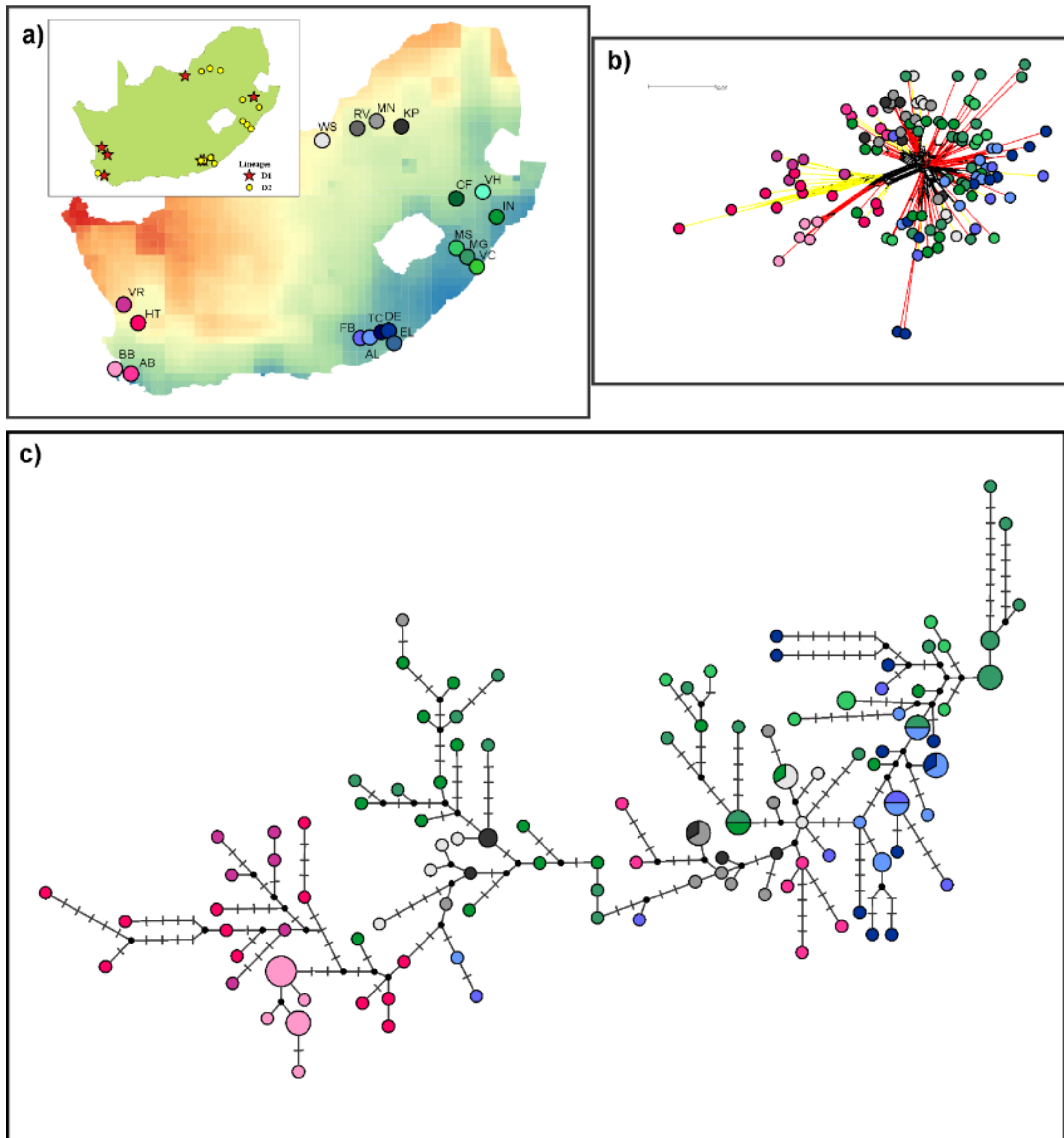


Figure 3.3 a) Geographical distribution of sampling localities in South Africa for *D. ellobius* plotted on a map of mean rainfall. Colour spectrum of the map vary from low values in red to high values in blue. (Insert: Map indicating where two lineages were found) **b)** Neighbour-Net phylogenetic network of subset nuclear EF1-α for *D. ellobius* with nodes coloured according to lineage. **c)** Statistical parsimony subset nuclear EF1-α haplotype network colour coded according to sampling locality to represent relative frequencies. Circle size depicts frequency; branches depict single mutational steps, small black circles display intersections and cross hatching indicate missing haplotypes.

Stellenbosch University https://scholar.sun.ac.za

3.3 Parasite and host demographics of lineages

Forty three individuals sampled belong to D1 (*D. abaris*) and 108 individuals to D2 (*D. ellobius*) (Table B.7). Most of the localities sampled showed segregation between the two lineages since D1 (*D. abaris*) was exclusively found at five localities and D2 (*D. ellobius*) at 12 localities. Individuals from both lineages were only found together on the same hosts at FB and AL (Table B.7) which represent a vicariant parapatric contact zone between at least three documented host species studied to date (Russo et al. 2010; Engelbrecht et al. 2011; du Toit et al. 2012; see Fig. A.1). Based on our sampling it seems that representatives from the two genetic lineages also differ in host occurrence. D1 (*D. abaris*) was isolated from seven host species with dominant representation on *M. namaquensis* (N = 14) and *G. brantsii* (N = 8) while D2 (*D. ellobius*) individuals were sampled from nine host species and was mainly collected from *R. dilectus* (N = 39) and *M. natalensis* (N = 28) (Fig. 3.4; Table B.7). Interestingly, in the zone of contact between some of the host species, D1 (*D. abaris*) was collected from *R. pumilio* and *R. dilectus*, whereas D2 (*D. ellobius*) was exclusively collected from the more mesic adapted *R. dilectus* (du Toit et al. 2012) (Fig. 3.4; Table B.7).

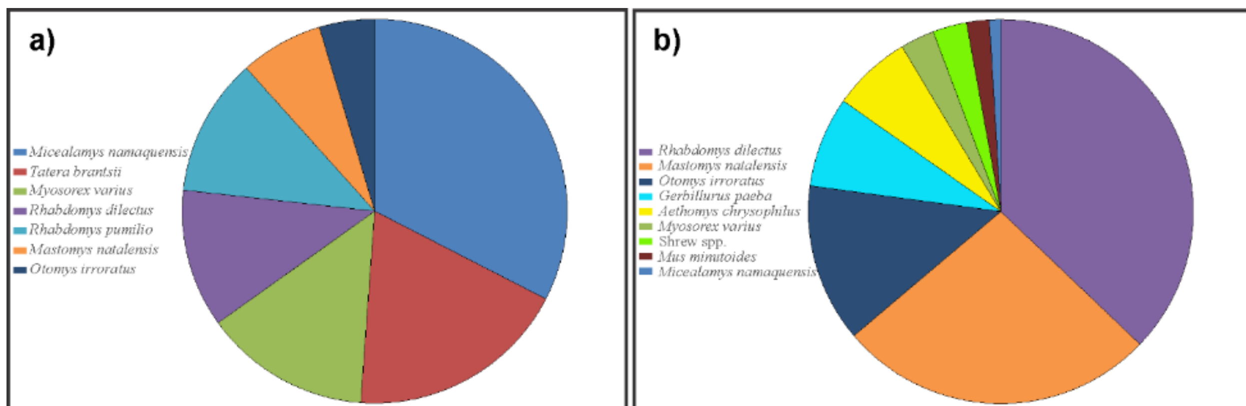


Figure 3.4 Small mammal host species and proportion of host species from which *D. ellobius* lineages, **a)** D1 (*D. abaris*) and **b)** D2 (*D. ellobius*) were sequenced.

Stellenbosch University https://scholar.sun.ac.za

3.4 Morphological analysis

3.4.1 *Size*

The ANOVA results indicated that there were significant differences between log transformed mean head (Fig. 3.5a), coxa (Fig. 3.5a), femur (Fig. 3.5c), and tibia (Fig. 3.5d) length of male individuals when grouped according to mtDNA clades. Analysis of co-variance indicated that male coxa, femur, and tibia length were the only measurements that significantly co-varied ($P < 0.01$) when again grouped into clades. No significant size difference was found for any of the measurements in female individuals between lineages (results not shown).

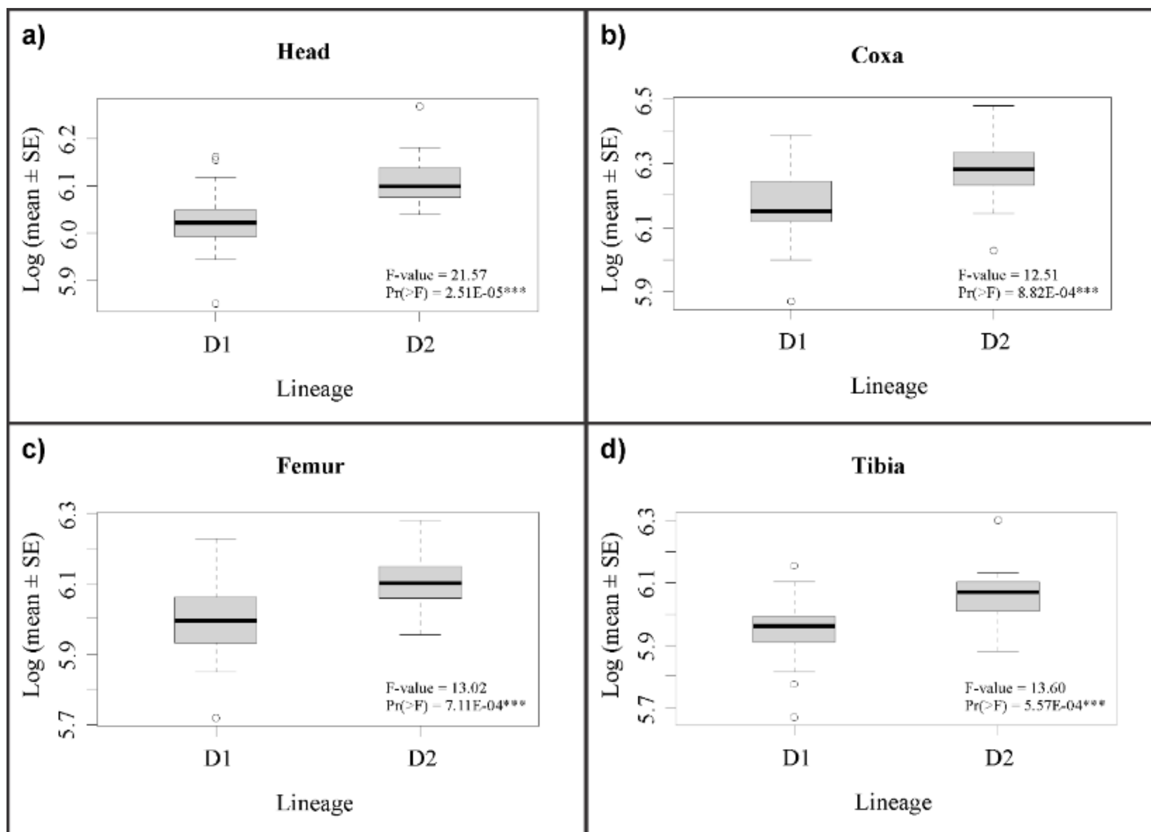


Figure 3.5 Boxplots for male individuals of mean log transformed length (µm) of **a)** head, **b)** coxa, **c)** femur, and **d)** tibia. ANOVA results are indicated in the bottom right corner of each boxplot. ***significant values

Stellenbosch University https://scholar.sun.ac.za

Principal component analysis on log transformed raw data resulted in two principal components that account for, respectively, 97.2 % and 97.5 % of the variance of male (*N* = 52) (Fig. 3.6a) and female (*N* = 78) (Fig. 3.6b) fleas. The first principal component (PC1) for males and females showed four variables (100 %) with moderate positive loadings suggesting that PC1 was mainly associated with overall size. Partial separation was evident when PC1 and PC2 was visualized for male individuals of D1 (*D. abaris*) (*N* = 14) and D2 (*D. ellobius*) (*N* = 38) (Fig. 3.6a). In contrast, no separation was evident when PC1 and PC2 was visualized for female individuals of D1 (*D. abaris*) (*N* = 24) and D2 (*D. ellobius*) (*N* = 54) (Fig. 3.6b). The average male femur size of the two lineages did not show a clear geographical pattern in variation (Fig. 3.7).

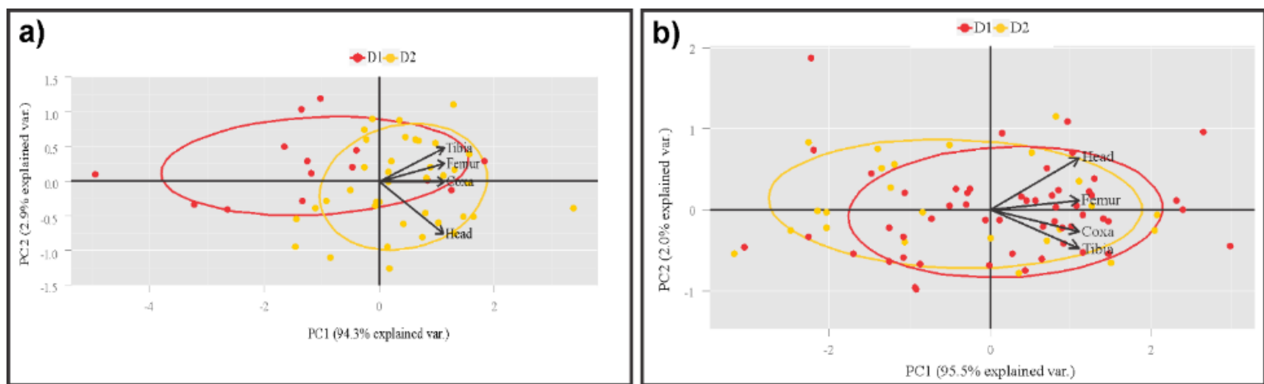


Figure 3.6 Principal component analysis of morphological measurements recorded for **a)** male *D. ellobius* lineages. D1 (*D. abaris*, red) (*N* = 14) and D2 (*D. ellobius*, deep yellow) (*N* = 38) and **b)** female *D. ellobius* lineages. D1, *D. abaris*, red) (*N* = 24) and D2, *D. ellobius*, deep yellow) (*N* = 54).

Stellenbosch University https://scholar.sun.ac.za

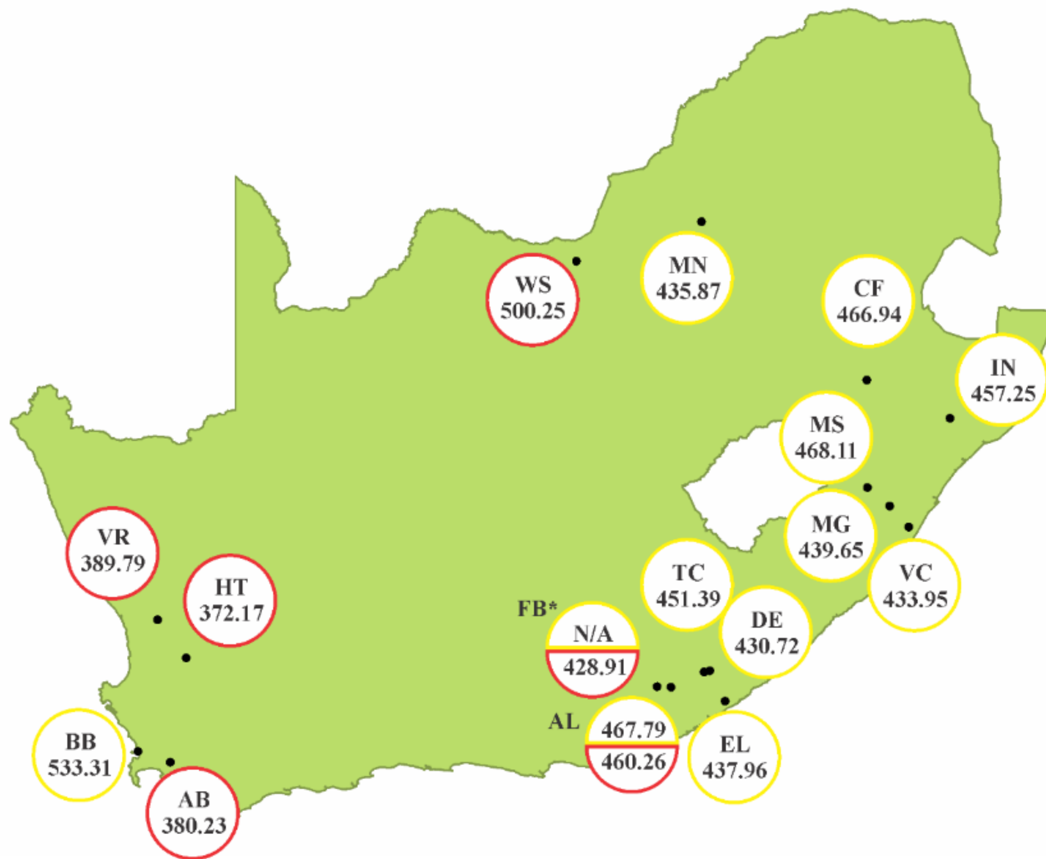


Figure 3.7 Mean femur length (µm) of male individuals at each locality. The two lineages are indicated by different colours (D1, *D. abaris*, red; D2, *D. ellobius*, deep yellow). *No males individuals were found at FB for D2 (*D. ellobius*, deep yellow).

3.4.2 *Characters*

The apical margin of male sternite VIII was very strongly rounded with a distinct small ventral sinus present in D2 (*D. ellobius*) and this corresponded with the description of *D. ellobius* (Hopkins & Rothschild 1966). Male sternite VIII of D1 (*D. abaris*) had a concave apical margin above a round-off ventral angle with a small ventral sinus present as described for *D. abaris* (Hopkins & Rothschild 1966). It was evident based on character scores (Table 3.6) that D1 and D2 reciprocally belong to *D. abaris* and *D. ellobius*.

Stellenbosch University https://scholar.sun.ac.za

Table 3.6 The morphological characters assessed for individuals of the two lineages with resulting character scores (CS). (also see Table B.6).

Character	D1	D2	*D. abaris*	*D. ellobius*
Genal comb	1	1	1	1
5th tarsal segment in males	1	1	1	1
Frons	0	0	0	0
Vestigial tubercle	0	0	0	0
Abdominal bristles	0	0	0	0
Size	1	0	1	0
Male sternite VIII	1	0	1	0
Male sternite IX	0	0	0	0
Female spermetheca	0	0	0	0
Male crochet phallosome	1	1	1	1
Pronotum	1	1	1	1
Total	6	4	6	4

4. Discussion

The study provides: (1) multiple lines of evidence (mtDNA, morphological, and ecological) showing support for the initial classification of *D. abaris* and *D. ellobius* as two valid species (or at least sub-species), and (2) new insights into the distribution and speciation process that may contribute to the genetic differentiation between *D. abaris* and *D. ellobius*.

Several lines of evidence can be put forward to reject the claim that the subtle morphological differences between *D. ellobius* and *D. abaris* are merely geographic variation within the same species (de Meillon et al. 1961; Segerman 1995). The study provides molecular (mtDNA) and morphological data to support the existence of two distinct taxa within the region and thus rather supports the suggestion by Hopkins & Rothschild (1966) that the taxa be recognized as distinct species or at least subspecies as described by Jordan (1930). The two distinct monophyletic mtDNA clades detected herein suggest the absence of female gene flow between the two assemblages and the sequence divergence between these lineages is almost equal to or exceeds the distance between other recognized flea species belonging to *Pulex* (5.8 %; De la Cruz & Whiting 2003) and *Oropsylla* (2.1 %; Brinkerhoff et al. 2011). The morphological differences recorded for male individuals of the two lineages further corresponded to the classification of *D. abaris* and *D.*

Stellenbosch University https://scholar.sun.ac.za

ellobius (Hopkins & Rothschild 1966) and given that the hind femur is considered as the preferred proxy of flea body size (Krasnov et al. 2003a; Khokhlova et al. 2010) all three measurements of the left hind leg (coxa, femur and tibia) co-varied and were larger in lineage D2 (*D. ellobius*) than in D1 (*D. abaris*). Female individuals did not show any morphological differences and, although the reason for this is unclear, it has been documented in many cases that variation in size is greater in males than in females of the same lineages (Fairburn 1997, 2005; Colwell 2000). It is unfortunate that the pattern recorded using mtDNA and morphological characteristics is not supported by nuclear DNA data in this study, but the absence of differentiation at the nuclear level is more than likely due to the close evolutionary relationship between the two clades and consequently incomplete lineage sorting at the nuclear DNA level (Funk & Omland 2003; Felsenstein 2004; Hernández-Vera et al. 2010). Incomplete lineage sorting is more prominent in nuclear DNA due to the higher effective population size and lower mutation rate for this marker system when compared to mtDNA. Furthermore, gene flow could be higher between nuclear sequences compared to mitochondrial sequences (Monteiro & Pierce 2001; Hernández-Vera et al. 2010).

Differences in the intraspecific genetic signatures of parasite species may be useful to predict different life history strategies (Gómez-Díaz et al. 2007; Whiteman et al. 2007**;** Poulin & Krasnov 2010). Previous phylogeographic investigations focussing on ectoparasites indicated that in general, the dispersion and gene flow of fleas are not correlated with their hosts (Brinkerhoff et al. 2011; Lin et al. 2014). After studying two flea species with different life histories, van der Mescht et al. 2015a (Chapter 2) indicated that a generalist fur flea, *Listropsylla agrippinae,* showed lower genetic diversity while a more host opportunistic nest flea, *Chiastopsylla rossi,* showed a higher level of mtDNA spatial genetic structure at the inter-populational scale. This pattern fitted the predictions that fleas that are more closely associated with the nest of the host are more restricted in dispersal (due to limited nests visits) and may thus show a higher level of genetic differentiation across the landscape when compared to a flea that spends more time on the host body ("fur" flea). If this tendency holds, it seems reasonable to speculate that the differences in the genetic diversity values between *D. ellobius* and *D. abaris* may also reflect similar differences in life history characteristics between these two species. The spatial genetic structure of D1 (*D. abaris*) is

Stellenbosch University https://scholar.sun.ac.za

comparable to a more generalist nest flea (higher inter-populational divergences due to more restrictions to dispersal) while D2 (*D. ellobius*) is comparatively reflective of a more specialist fur flea (showing a lower overall genetic diversity and inter-populational divergences).

Based on general ecological predictions for ectoparasites, it is also possible to speculate further on some other life history and ecological differences between D1 (*D. abaris*) and D2 (*D. ellobius*). It has furthermore been recorded that the level of host association (host specificity and time spent on the host) exhibited by a parasite can influence the distribution of parasites within a host population, community and environment (Krasnov et al. 1997; Medvedev & Krasnov 2006; Krasnov 2008; Froeschke et al. 2013). It is therefore possible that the specific life history strategy that is used by the parasite can influence the species' geographic distribution (Poulin & Krasnov 2010) and this will be reflected in its spatial genetic structure across the landscape (Krasnov & Poulin 2010). In support of this prediction, it has also been suggested that the overall genetic diversity may be elevated in parasite taxa that have more habitat-specific climatic requirements (Morand et al. 2010). Unfortunately, we could not calculate accurate indices of host specificity for each lineage (we only sequenced a subsample of all *D. ellobius* individuals), and counting the number of host species utilized by a parasite as a measure of host specificity can lead to biased results (Poulin et al. 2011). Nevertheless, evenness of proportional abundances on host species, seems to provide supporting evidence that D1 (*D. abaris*) is less host specific than D2 (*D. ellobius*). Interestingly, in the zone of contact D1 (*D. abaris*) was also collected from *R. pumilio* and *R. dilectus*, whereas D2 (*D. ellobius*) was exclusively collected from the more mesic adapted *R. dilectus* (du Toit et al. 2012) (Fig. 3.4; Table B.7).

The environmental factors that could potentially be implicated as causing the speciation between D1 (*D. abaris*) and D2 (*D. ellobius*) is not clear since the phylogeographic structure of the fleas did not correspond perfectly to the documented known vicariant biogeographic patterns for the host species occurring in the region (Russo et al. 2010; Engelbrecht et al. 2011; du Toit et al. 2012). Although D1 (*D. abaris*) was recorded at multiple localities in the western xeric region of the country it appears to display a broader geographic range compared to D2 (*D. ellobius*), which is more often associated with the mesic eastern region of the country (Fig. 3.2). It has been

Stellenbosch University https://scholar.sun.ac.za

hypothesized previously that the sporadic occurrence of *D. ellobius* (senso stricto) in semi-arid and arid areas may be explained by patchy survival in refugia of favourable habitats from a humid epoch when grassland extended further west in SA (Segerman 1995). Since precipitation gradients have also been shown to be important in diversification of other flea species (Krasnov et al. 1998, 2005b and references therein) it seems reasonable to suggest that the two species have different ecological requirements in terms of humidity at least. It is also interesting to note that the estimated 6.42 Mya time of divergence between the two mtDNA lineages corresponds broadly to the rainfall and vegetation changes characterizing the continent around the Miocene-Pliocene boundary (Partridge 1997; Tyson & Partridge 2000; McCarthy & Rubidge 2005). McCarthy & Rubidge (2005) suggested that the western side of southern Africa experienced a decrease in rainfall which resulted in a change in vegetation from sub-tropical to the modern dryer biome (Coetzee 1978; Scott et al. 1997; Mucina & Rutherford 2006). This is in contrast to the eastern side of the subcontinent that recorded higher rainfall during the same period (McCarthy & Rubidge 2005). Since the same vegetational/climatic changes most likely also affected several of the rodent hosts in the region in a similar fashion (Mongelard & Matthee 2012), it is possible that the evolution of *Dinopsyllus* was equally affected.

Apart from being geographically isolated, the two flea lineages seldom occurred sympatrically on the same host species. The exception being at two localities (AL and FB) that are regarded as the contact zone (Fig. 3.2 and 3.3; Table B.7) where the two lineages were found to occur on the same host individual. Based on host prevalence, it seems that the two lineages prefer (considering highest proportional prevalence) different host species and each lineage prefer host species that occur throughout their particular distribution range. This may be indicative of interspecific competition for rodent host species, ultimately leading to the occupation of different niches caused by ecological fitting (Hoberg & Brooks 2008).

The study underlines the value of molecular markers, specifically in our instance the mtDNA COII marker, to address morphologically challenging questions in the field of flea taxonomy (Luchetti et al. 2005, 2007; Whiting et al. 2008; Lawrence et al. 2014; Lin et al. 2014). The use of an extensive (multiple localities spanning the parasite's geographic range) and intensive (all possible

Stellenbosch University https://scholar.sun.ac.za

host species per locality) sampling approach in addition to a combination of markers (molecular and morphological) provides us with an opportunity to resolve two distinct lineages representing two putative species (*D. ellobius* & *D. abaris*; Hopkins & Rothschild 1966) currently classified as a single species (*D. ellobius*; Segerman 1995). These two species can only be morphologically identified based on males which seems to be a common phenomenon for many flea species in SA (Segerman 1995). The only way single sex morphological classification can be overcome is by classifying females with DNA barcoding. The sampling approach and the use of a mitochondrial marker provide novel ecological predictions open for future investigations. Additional sampling, and a further in depth analyses of habitat suitability, would provide more evidence to explain the evolutionary mechanisms involved.

Stellenbosch University https://scholar.sun.ac.za

Chapter 4

The influence of life history characteristics on flea (Siphonaptera) species distribution models

Stellenbosch University https://scholar.sun.ac.za

1. Introduction

Ectoparasites exhibit pronounced variation in life history strategies, with parasite-host associations ranging from one-to-one symbiosis (host specific) to multi-partner symbiosis (host generalist). Furthermore, life history also differs between taxa with some parasite species being only temporarily associated with the body of a host (e.g. ticks and fleas), while others are more permanently linked with a host (e.g. lice) (Marshall 1981; Morand et al. 2006; Alvarez et al. 2010). Consequently, it can be argued that the distribution of parasite species with multiple free-living stages (i.e. temporary parasites) are likely more strongly affected by the off-host environment (e.g. climatic and landscape features), whereas the distribution of permanent parasites may be indirectly driven by factors affecting host assemblages (e.g. shelter and food) (see Marshall 1981; Morand et al. 2006). Temporary parasite taxa, however, are also characterized by varying levels of host association (Marshall 1981; Morand et al. 2006), with these subtle differences in life history characteristics also potentially adding further complexity to patterns of parasite species distributions (Morand et al. 2006; Froeschke et al. 2013). These complex relationships between parasite, host and environment may make it difficult to achieve accurate range predictions for ectoparasites.

The increasing threat posed by emerging infectious diseases (Daszak & Cunningham 2000; Bitam et al. 2010), coupled with an increase in the availability of species occurrence records, has stimulated renewed interest in predicting the current and future distributions of arthropod vectors. Species distribution (SD) modelling has proved useful for this purpose (e.g. Pérez-Rodríguez et al. 2013; Poretta et al. 2013a; Estrada-Peña et al. 2013), with particular success for several medically and veterinary important arthropod vectors (e.g. Adjemian et al. 2006; Rose & Wall 2011; Poretta et al. 2013a; Wardhana et al. 2014), and also a single study on fleas (Adjemian et al. 2006). In the latter study the authors used standard climatic variables (temperature, relative humidity and precipitation) in a GARP modelling approach to estimate the regional distribution of 18 flea species that act as vectors of *Yersinia pestis*, the bacterial agent of plague. Given the recent re-emergence of certain flea-borne diseases (Bitman et al. 2010), any improvements in our understanding of changes in flea vector distributions is valuable for the field of epidemiology.

Stellenbosch University https://scholar.sun.ac.za

Fleas are obligate ectoparasites of terrestrial vertebrates, and are regarded as "permanent satellites" of their hosts, due to the intimate association between fleas and hosts (Medvedev & Krasnov 2006; Krasnov 2008). In general, fleas spend part of their life cycle in the host's nest (egg, larvae and pupae) while adults occur on the body of the host. The length of time that adults spend on the host varies between flea taxa (Ioff 1941; Marshall 1981; Krasnov 2008), and this difference in microhabitat preference allows fleas to be categorized either as “fur” (adults spend more time on the host), “nest” (adults spend more time in the nest of the host) or “fur/nest” species (adults spend more or less equal amounts on the host and in the nest of the host) (Ioff 1941; Marshall 1981; Krasnov 2008). Flea species also differ in terms of host specificity, which ranges from host specific (recorded from ≤ 2 host species) to host opportunistic (recorded from > 2 host species) (Marshall 1981; Combes 2001; Krasnov 2008). There appears to be no relationship between microhabitat preference and level of host specificity exhibited by flea species (i.e. nest fleas are not generally regarded as having a higher level of host specificity). However, differences in level of host association between a flea and a host may have profound implications for the level of exposure to environmental features (e.g. climate and landscape) (see Krasnov et al. 2010b, 2013a, 2014, 2015), highlighting the need to assess the importance of these life history characteristics in SD modelling studies.

The aim of this study was therefore to compare SD model performance and the relative importance of predictor variables between flea species with different microhabitat preferences (fur vs. nest) and levels of host specificity (opportunistic vs. specific). We predicted that fur fleas will be more accurately modelled due to being more strongly affected by variables associated with regional environmental conditions (e.g. climate), while nest fleas will be less accurately modelled due to being affected by conditions within the host nest (e.g. soil conditions and microclimate; since all life stages spend the majority of their life cycle off the host and have limited dispersal capabilities). Furthermore, we predict that host specific fleas will be more strongly associated with the abiotic variables constraining their host's distribution and therefore will be more accurately modelled than host opportunistic fleas. This is expected because host specific fleas are assumed to be adapted to the immediate environment of their specific host and thus are expected to tolerate a narrower range

Stellenbosch University https://scholar.sun.ac.za

of physical conditions compared to host opportunistic fleas (see Krasnov et al. 2003b; Krasnov 2008; Shenbrot et al. 2007).

2. Materials and Methods

A dataset comprising occurrence records for flea species that parasitize 83 small mammal species (rodents, elephant shrews and shrews) from 1109 localities across South Africa (SA) were compiled from published literature (Segerman 1995) and newly sampled sites (Chapters 2 and 3); van der Mescht et al. 2015a, b). Flea species were selected for modelling if occurrence records from 10 or more localities were available (following Stockwell & Peterson 2002; Wisz et al. 2008; Robe et al. 2014) and when their microhabitat preference and host specificity were also known (Table 4.1). Flea species were subsequently categorized by microhabitat preference (fur: adult stage spend more time on the host vs. nest: adults spend more time in the nest of the host vs. fur/nest: adults stage spend roughly equal time in the nest and on the host) (Ioff 1941; Marshall 1981; Krasnov 2008) and host specificity (opportunistic: recorded from > 2 host species vs. specific: recorded from ≤ 2 host species: Marshall 1981; based on de Meillon et al. 1961; Segerman 1995) (Table 4.1). Fleas that spend almost equal amounts of time in the nest and fur (indicated as fur/nest in Table 4.1) were only included when analysing species by host specificity. Furthermore, instances where a flea species was only recorded once from a host species were seen as accidental infestations and not considered as valid when classifying host specificity.

Stellenbosch University https://scholar.sun.ac.za

Table 4.1 Flea species analysed in the study. Microhabitat preference, host specificity, number of unique occurrence records and two measures of model performance (AUC and TSS) are indicated for each species.

Species	Microhabitat preference	Host specificity[a]	Occurrence records	AUC[b]	TSS[b]
Chiastopsylla coraxis	nest	opportunistic	51	0.907	0.694
Chiastopsylla mulleri simplex	fur/nest	specific	26	0.948	0.752
Chiastopsylla pitchfordi	nest	opportunistic	58	0.907	0.706
Chiastopsylla quadrisetis	fur/nest	specific	15	0.940	0.804
Chiastopsylla rossi	nest	opportunistic	88	0.545	0.067
Ctenopthalmus calceatus	fur/nest	opportunistic	85	0.848	0.518
Demeillonia granti	fur	specific	15	0.737	0.386
Dinopsyllus ellobius	fur/nest	opportunistic	439	0.646	0.231
Dinopsyllus lypusus	fur	opportunistic	34	0.894	0.545
Epirimia aganippes	fur	opportunistic	52	0.738	0.280
Listropsylla agrippinae	fur	opportunistic	150	0.760	0.360
Listropsylla chelura	nest	specific	50	0.821	0.454
Listropsylla dorripae	nest	specific	70	0.695	0.277
Listropsylla fouriei	nest	specific	13	0.665	0.272
Listropsylla prominens	fur	opportunistic	28	0.882	0.519
Praopsylla powelli	fur/nest	specific	15	0.885	0.560
Xenopsylla eridos	nest	opportunistic	134	0.810	0.499
Xenopsylla mulleri	nest	specific	10	0.978	0.763
Xenopsylla pirei	nest	opportunistic	284	0.737	0.423
Xenopsylla trifaria	nest	specific	24	0.733	0.363
Xenopsylla versuta	fur/nest	specific	32	0.890	0.636

[a] Opportunistic = recorded from > 2 host species and specific = recorded from ≤ 2 host species.

[b] AUC = area under the curve of the receiver operating characteristic (ROC), TSS = True Skill Statistic

Preliminary climate and landscape feature variables were selected based on our knowledge of flea ecology, limiting candidate variables to only predictors that are considered ecologically relevant to flea species (following Austin 2007; Elith et al. 2011; Merow et al. 2013). All predictor variables

Stellenbosch University https://scholar.sun.ac.za

and flea occurrence data were converted to Quarter Degree Grid Cell (QDGC) scale and cropped to the borders of SA. Six remotely-sensed climate-based variables (daytime land surface temperature (hereafter referred to as day temperature), Leaf Area Index (LAI), Normalised Difference Vegetation Index (NDVI), rainfall, water vapour, and soil characteristics) and one landscape feature variable (Topography) were extracted from the NASA-NEO website (http://neo.sci.gsfc.nasa.gov/about/) as potential predictor variables (missing values were estimated as the average of contiguous cells). Climate is known to generally influence flea populations to a greater extent than host species, especially at regional and local scales (Stark 2002; Krasnov et al. 2015), with air temperature, rainfall and relative humidity being important for flea survival (see Marshall 1981; Stark 2002; Adjemian et al. 2006; Krasnov et al. 2001a, b, 2002a, b). NDVI is widely used in arthropod vector distribution modelling and is a measure of primary productivity (plant photosynthetic activity), and therefore can be considered as a proxy for general arthropod habitat conditions (Reisen 2010; Pettorelli et al. 2011). Furthermore, NDVI has also been successfully used in small mammal resource and population dynamics studies (Andreo et al. 2009a, b; Pettorelli et al. 2011), and therefore may also be a surrogate for host availability. LAI is a measure of plant canopy structure and can influence incident radiation and evapotranspiration at the soil surface (Myneni et al. 2002). Additionally soil data, including soil organic carbon content, pH, cation exchange capacity, percentage sand and bulk density, were extracted from the SoilGrids database (Hengl et al. 2014) (http://www.soilgrids.org/) at a depth of 60-100 cm. These soil characteristics may be expected to have direct (via microhabitat) and indirect (small mammal burrowing conditions) effects on flea species distributions (Cook 1990; du Plessis et al. 1992; Osacar-Jimenez et al. 2001; Krasnov et al. 2002a; Shenbrot et al. 2002).

To remove collinearity between climate-based predictors and to summarize seasonality we performed harmonic regressions for all of the climate variables using monthly data from January 2001 until June 2014 (following the methods of Estrada-Peña et al. 2014). These Fourier-transformed variables ("harmonic variables" hereafter) represent key temporal climate trends, reflecting different measures of seasonality (Estrada-Peña et al. 2014). Maximum, minimum, mean, range, and coefficient of variation (CV) values were also calculated for each climate variable. The correlation between all predictor variables was then calculated to identify collinear

Stellenbosch University https://scholar.sun.ac.za

predictors, with the most strongly correlated variable excluded and the process repeated until the strongest correlation was weaker than |0.7|. When choosing between two strongly correlated variables, the preference was to drop harmonic variables rather than the other more biologically-interpretable variables (i.e. maximum, minimum, mean, range, CV, topography, and soil variables). Through this process an initial set of 68 predictor variables was reduced to 19 predictors with minimal collinearity and with clear ecological relevance, avoiding the inclusion of variables that are irrelevant and/or will inflate models (Cumming & van Vuuren 2006; Pérez-Rodríguez et al. 2013; Estrada-Peña et al. 2013, 2014). All analyses were conducted in R v3.1.3 (R Foundation for Statistical Computing, Vienna, AT, AU) and ArcGIS 10.1 (ESRI 2011).

Species distribution modelling relates species presence (or presence and absence) data to environmental variables to predict the distribution of species over a specified geographic range (Franklin 2009). In this study MaxEnt models, based on the maximum entropy algorithm, were used with presence-only data (MaxEnt v3.3.3; Phillips et al. 2006, 2008). To account for potentially spatially-biased sampling of fleas across our study region (e.g. due to more studies being conducted in protected areas), MaxEnt models were adjusted for uneven sampling by incorporating background data reflecting patterns in sampling effort (Elith et al. 2011; Merow et al. 2013). All 1109 localities from which flea species occurrences have been published were therefore included as background points to distinguish false absences (lack of survey data) from true absences (species was not recorded).

For each flea species variable importance (i.e. relative contribution of each predictor variable) was calculated using the full dataset. The mean importance of each variable type (the average of all related individual predictor variables) was determined for each group of flea species (microhabitat preference and host specificity). Model performance was evaluated using 10-fold cross-validation to calculate the area under the curve (AUC) of the receiver operating characteristic (ROC; Peterson et al. 2007) and the true skill statistic (TSS; Allouche et al. 2006). Analysis of variation (ANOVA) was used to test if AUC and TSS values differed according to species' microhabitat preference and host specificity. Non-metric multi-dimensional scaling (NMDS) and analysis of similarity

Stellenbosch University https://scholar.sun.ac.za

(ANOSIM; implemented from the *vegan* package in R) were used to test if differences in variable importance were related to microhabitat preference or host specificity.

3. Results

A total of 21 flea species from small mammals were selected with known microhabitat preference, host specificity, and sufficient occurrence records (Table 4.1). Overall, model performance was highly variable between flea species (AUC from 0.545 to 0.978; TSS from 0.067 to 0.804; Table 4.1), but was good to excellent on average (mean ± SE AUC = 0.799 ± 0.026, TSS = 0.464 ± 0.045).

Table 4.2 Final list of variable type and individual predictor variables used for modelling.

Variable type	Predictor variable
Day temperature	6th harmonic component of daytime land surface temperature
	7th harmonic component of daytime land surface temperature
	Minimum daytime land surface temperature
LAI	3rd harmonic component of daily LAI
	4th harmonic component of daily LAI
NDVI	4th harmonic component of daily NDVI
	7th harmonic component of daily NDVI
Rainfall	4th harmonic component of daily rainfall
	6th harmonic component of daily rainfall
	7th harmonic component of daily rainfall
	Minimum rainfall
	Maximum rainfall
Water vapour	4th harmonic component of daily water vapour
	5th harmonic component of daily water vapour
	6th harmonic component of daily water vapour
	Minimum water vapour
Soil	Soil percentage sand
	Soil organic carbon
	Soil pH

Stellenbosch University https://scholar.sun.ac.za

For flea microhabitat preference, there was no significant difference in AUC ($F_{1, 13}$ = 0.120, p = 0.735) or TSS ($F_{1, 13}$ = 0.101, p = 0.756) values between fur (AUC = 0.802 ± 0.035; TSS = 0.418 ± 0.050) and nest (AUC = 0.780 ± 0.041; TSS = 0.452 ± 0.070) fleas. In addition, importance of the predictor variables did not differ significantly between fur and nest fleas (ANOSIM: R^2 = 8.7 %, p = 0.204; Fig 4.1). Variable importance averaged by variable type (i.e. day temperature, LAI, NDVI, rainfall, water vapour, and soil; Table 4.2), revealed that temperature-related variables were most important for predicting species distributions, and this was particularly so for fur fleas (Fig 4.2). Rainfall was the second most important variable type for predicting the occurrence of nest fleas (having a significantly higher relative contribution to nest fleas than fur fleas; Fig 4.2). Considering variables individually, minimum day temperature and minimum water vapour contributed significantly more strongly to fur than nest fleas while minimum rainfall contributed significantly more towards predicting the distribution of nest fleas than fur fleas (Fig. A.4a).

Stellenbosch University https://scholar.sun.ac.za

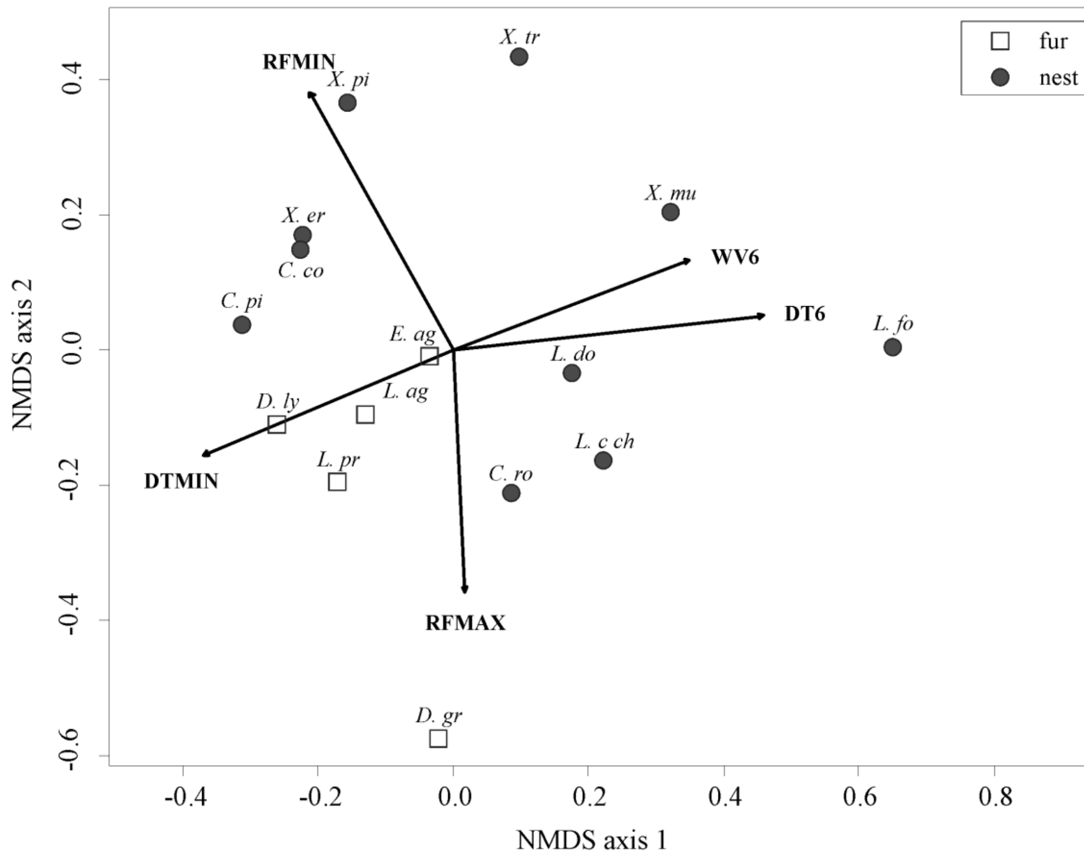


Figure 4.1 NMDS ordination plot showing the relationship between microhabitat preference of flea species and the variable importance of predictors included in each species distribution model. The best linear fit of all variables that had a significant ($p \leq 0.05$) influence are indicated. Variable codes: DT6, 6th harmonic component of daytime land surface temperature; DTMIN, minimum daytime land surface temperature; RFMIN, minimum rainfall; RFMAX, maximum rainfall; WV6, 6th harmonic regression component of water vapour.

Stellenbosch University https://scholar.sun.ac.za

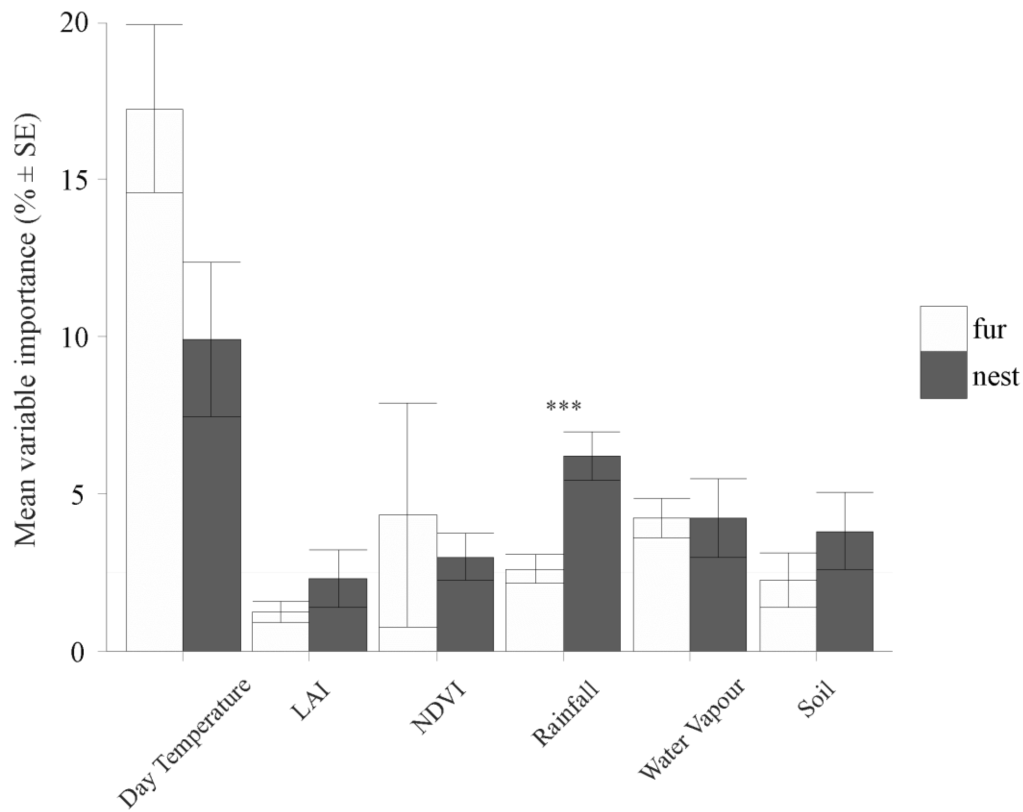


Figure 4.2 Variable importance (i.e. percent relative predictor variable type importance) in MaxEnt models, averaged across flea species for each variable type based on microhabitat preference (see Table 4.2 for variable category information). Significant differences in the contribution of predictor variable types between the two categories of species are indicated by asterisks: *** 0.001, ** 0.01, * 0.05.

Stellenbosch University https://scholar.sun.ac.za

In the case of host specificity, there was no significant difference in AUC ($F_{1, 24} = 0.427$, $p = 0.520$) or TSS ($F_{1, 24} = 0.699$, $p = 0.411$) values between host opportunistic (AUC = 0.789 ± 0.035; TSS = 0.440 ± 0.059) and -specific (AUC = 0.829 ± 0.036; TSS = 0.527 ± 0.065) species. However, there was a strong and significant difference in variable importance between host opportunistic and -specific fleas (ANOSIM: R^2 = 31.3 %, p = 0.003; Fig 4.3). When averaging variable importance by type, day temperature, followed by rainfall, was the most important predictor for both host opportunistic and -specific fleas. However, day temperature contributed more strongly towards host opportunistic than -specific fleas, whereas rainfall contributed equally towards both host opportunistic and –specific fleas (Fig 4.4). Further, NDVI contributed more strongly towards host specific than -opportunistic fleas (Fig 4.4). Considering variables individually, minimum day temperature contributed significantly more towards host opportunistic than –specific fleas (Fig. A.4b).

Stellenbosch University https://scholar.sun.ac.za

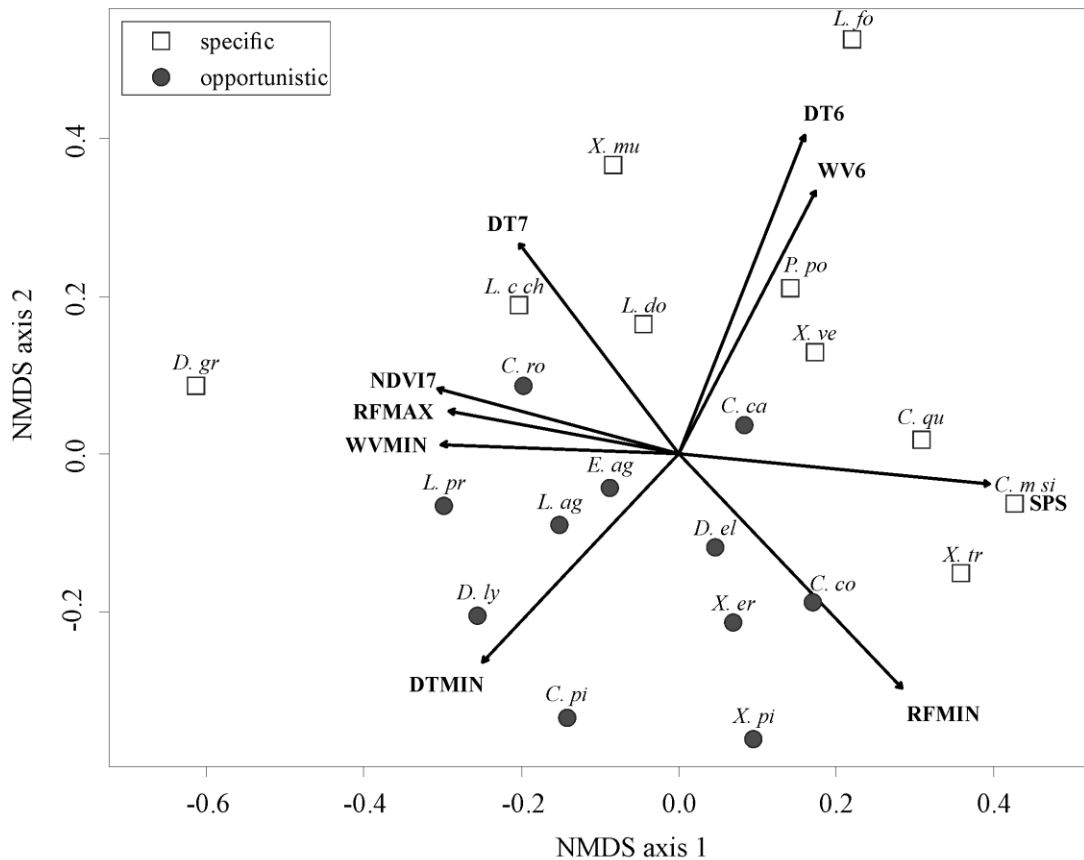


Figure 4.3 NMDS ordination plot showing the relationship between host specificity of flea species and the variable importance of predictors included in each species distribution model. The best linear fit of all variables that had a significant ($p \leq 0.05$) influence are indicated. Variable codes: DT6, 6th harmonic component of daytime land surface temperature; DT7, 7th harmonic component of daytime land surface temperature; DTMIN, minimum daytime land surface temperature; NDVI7, 7th harmonic component of NDVI; RFMIN, minimum rainfall; RFMAX, maximum rainfall; WV6, 6th harmonic regression component of water vapour; WVMIN, minimum water vapour; SPS, soil percentage sand.

Stellenbosch University https://scholar.sun.ac.za

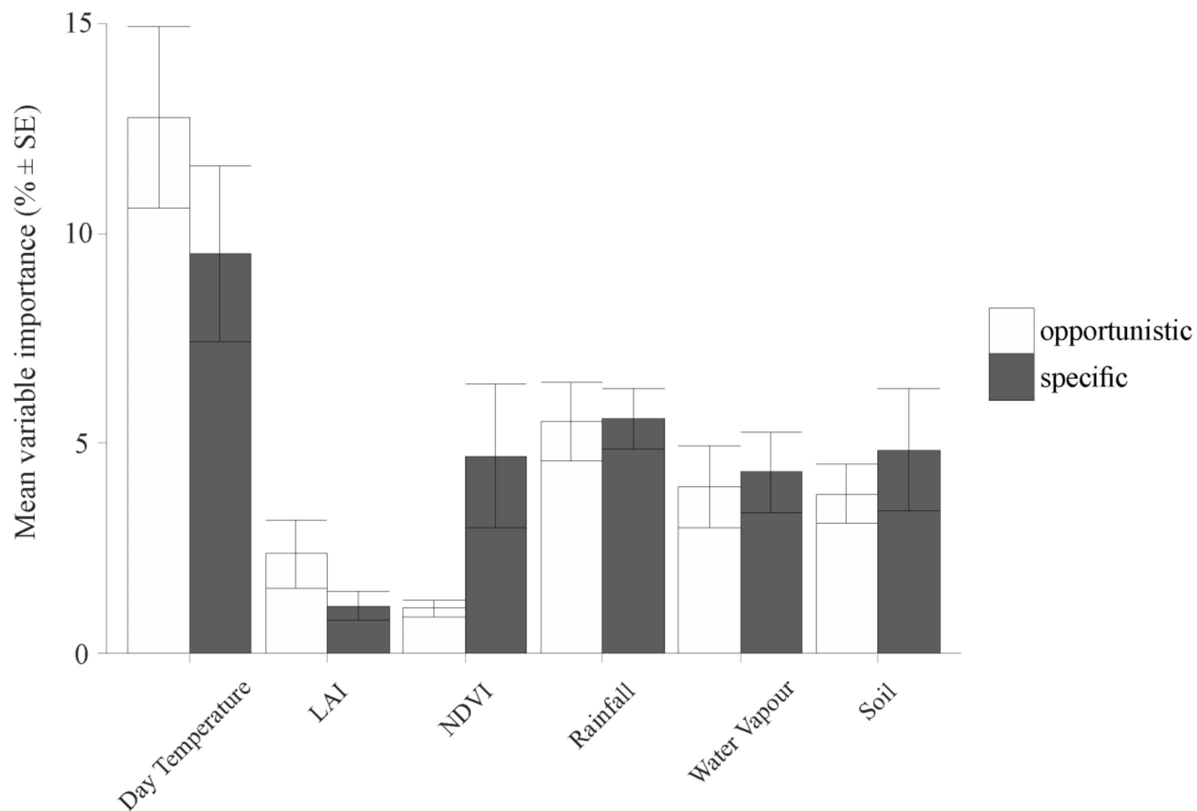


Figure 4.4 Variable importance (i.e. percent relative predictor variable type importance) in MaxEnt models, averaged across flea species for each variable type based on host specificity (see Table 4.2 for variable category information). Significant differences in the contribution of predictor variable types between the two categories of species are indicated by asterisks: *** 0.001, ** 0.01, * 0.05.

4. Discussion

The study confirms that contemporary SD modelling, that includes climatic and landscape variables, has potential for improving predictions of changes in the distribution of flea species. The model performance for fleas was good overall, with life history not having a significant effect on model performance. However, the importance of predictor variables differed considerably between species with different life history strategies, suggesting differential sensitivity to climate (temperature and rainfall) and landscape feature (NDVI) variables among groups of fleas.

Stellenbosch University https://scholar.sun.ac.za

Abiotic conditions such as air temperature, relative humidity and precipitation are important for fleas (Marshall 1981; Krasnov et al. 2001a, b, 2002a, b; Stark 2002; Adjemian et al. 2006), especially at the regional scale (Wisz et al. 2013; Krasnov et al. 2015). In particular, air temperature and relative humidity can have a direct (influence on egg production and rate of oviposition, development and survival of pre-imaginal fleas, and survival of imagoes) (Krasnov et al. 2001a, b, 2002a, b) and/or indirect affect (influence hosts abundance and distribution) on flea survival (Marshall 1981; Krasnov 2008; Krasnov et al. 2015). However, flea species also vary in terms of preferred temperature range (reviewed by Krasnov 2008) and although there is evidence that phylogenetic relatedness may explain this pattern for certain flea taxa it appears not to be consistent across all taxa (Vashchenok 1988; Krasnov 2008, Krasnov et al. 2015). In the present study we found that temperature is an important predictor of flea species ranges, irrespective of life history, but that the relative importance of temperature differed between flea life histories. In particular, the higher importance of temperature for host opportunistic compared to -specific fleas may be related to their physiological limits (Krasnov et al. 2001a, b, 2002a, b). Host opportunistic fleas generally have a wider tolerance range (climate and host composition), are geographically more widespread and are likely to experience more environmental heterogeneity throughout their geographical range (Krasnov et al. 2003b, 2015; Krasnov 2008; Shenbrot et al. 2007). In contrast, host specific fleas generally have narrower tolerance ranges (climate and host composition), are less widespread and are likely to experience more environmental homogeneity throughout their geographical range (Krasnov et al. 2003b, 2015; Krasnov 2008; Shenbrot et al. 2007). Thus, although temperature is important, especially for the development of immature nest stages in general, it is evidently less important for host specific fleas in this study.

Linking on to this it seems that host specific fleas are more influenced by variables related to host availability as evident from the higher NDVI importance for host specific compared to –opportunistic fleas in the current study. NDVI has been used as a measure of the amount of suitable habitat for arthropod vectors (Reisen 2010; Pettorelli et al. 2011) and may also be a good proxy for several aspects of habitat quality that are of relevance to fleas. Abiotic factors influence vegetation and thus food supply (Singleton et al. 2001; Brown & Ernest 2002; Jaksic & Lima 2003;

Stellenbosch University https://scholar.sun.ac.za

Meserve et al. 2003; Zhang et al. 2003) and actual or perceived predation risk (Rosenzweig 1973; Norrdahl & Korpimäki 1995; Spencer et al. 2005) which will affect small mammal host abundance and distribution (Andreo et al. 2009a, b; Pettorelli et al. 2011), this in turn may directly influence flea abundance and distribution (Krasnov et al. 2015). In the current study, environmental stability could facilitate small mammal host population stability and subsequent flea population stability and specialization (MacArthur 1955, 1972), because host specific fleas tolerate restricted abiotic and biotic conditions (Brown 1995). In contrast, host opportunistic fleas can tolerate variable environments and conditions and are thus less affected by NDVI (Pettorelli et al. 2011).

Predictor importance for individual variables also differed between species with different microhabitat preferences. Temperature was again overall the most important predictor for both fur and nest fleas. However, the difference in the importance of temperature was even more pronounced when compared with level of host specificity (as discussed above), with temperature being more important for fur fleas as opposed to nest fleas. Fur fleas spend more time on the body of the host and as a result are potentially exposed to higher fluctuations in ambient temperature and relative humidity compared to nest fleas (see Shenbrot et al. 2002). It has been suggested that nest fleas may have evolved to spend more time in nests due to more constant and buffered microclimatic conditions brought about by the physical properties of nests (Rothschild & Clay 1952; Hopkins 1957; Marshall 1981). Although discrepancies in the fur versus nest dichotomy are mainly attributed to ambient temperature in the literature (see Krasnov 2008 and references within), it has also been suggested that within-host among-flea difference can be explained by this dichotomy, whereas between-host within-flea differences are better explained by between-host difference in nest construction (Krasnov et al. 1997, 1998).

Studies on different nest types (burrows, above-ground nests and nest within rock crevices) all demonstrate that microclimatic conditions are more stable in nests compared to the external environment (Cook 1990; du Plessis et al. 1992; Osacar-Jimenez et al. 2001; Shenbrot et al. 2002). Other factors that can facilitate higher and more stable humidity levels in the nest include, the presence of nest material (Shenbrot et al. 2002; Burda et al. 2006) and higher soil water capacity of the mineral and organic enriched nest soil (Hudson 1994), due to the activities of hosts in nests

Stellenbosch University https://scholar.sun.ac.za

(Shenbrot et al. 2002; Burda et al. 2006). In support of this, rainfall contributed significantly more towards explaining the distribution of nest compared to fur fleas in the current study. Rainwater naturally filters down into soil layers which can contribute to maintaining higher and more stable humidity levels in the nest of hosts (Cook 1990; Osacar-Jimenez et al. 2001; Shenbrot et al. 2002). In addition, the type and complexity of host nests are influenced by the soil texture (Laundre & Reynolds 1993; Shenbrot et al. 2002). It is therefore not surprising that soil percentage sand was more important for nest compared to fur fleas. It is evident from this study that the interplay between temperature, rainfall and potentially relative humidity with nest construction can facilitate the separation of fleas into different microhabitat types.

Our study suggests that SD modelling can be a useful tool for studying the drivers of flea species distributions and also the underlying ecology of these species, but caution needs to be taken when deciding which predictor variables to include. While our results highlight how contemporary models can perform well, it is unclear to what extent the inclusion of biotic interactions (e.g. host availability and competition) could further improve model accuracy and transferability (Pérez-Rodríguez et al. 2013; Wisz et al. 2013). Specifically, modelling and comparing flea species with different levels of host specificity could possibly benefit from including accurate host species occurrence data. Furthermore, it is important to remember that ecological patterns are affected by processes that act at different scales (Coreau & Martin 2007). For example, the assembly of flea compound (all species infesting a host community) communities is strongly affected by host filters (e.g. evolutionary, biogeographic and historical forces) at the continental scale, while at the regional and local scale it is more strongly affected by the abiotic filters (e.g. topography, NDVI and climate) (see Krasnov et al. 2010b, 2013a, 2014, 2015). Therefore, the value of the inclusion of host species data (and the predominant importance of temperature in our models) may be contingent on the spatial scale of analysis.

Five of the flea species in our study (*Chiastopsylla rossi*, *Dinopsyllus lypusus*, *Listropsylla dorripae*, *Xenopsylla pirei*, and *Xenopsylla versuta*) have been implicated as possible vectors of diseases in SA (e.g. plague, de Meillon et al. 1961). As a consequence, accurate forecasts of the future distributions of these species are valuable for the field of epidemiology. Our results suggest

Stellenbosch University https://scholar.sun.ac.za

that despite differences in their degree of host specificity, SD models should perform well for all of these species. However, due to differential sensitivity to different groups of climatic and landscape variables, host specific and -generalist flea species are likely to respond very differently to changes in abiotic conditions. As a result, our results suggest the importance of explicitly considering species life history as a potential mediating variable when predicting flea species distributions.

Stellenbosch University https://scholar.sun.ac.za

Conclusions

Stellenbosch University https://scholar.sun.ac.za

The study explored various mechanisms that are involved in shaping the diversification and distribution of fleas on small mammals across South Africa.

In Chapter 2, we investigate the influence of vicariance and host association on the genetic structure of two generalist flea species, *Listropsylla agrippinae*, and *Chiastopsylla rossi*. In parasitic taxa, life history traits such as microhabitat preference and host specificity can result in differential evolutionary responses to similar abiotic events. The taxa differ in the time spent on the host (predominantly fur vs. nest) and level of host specificity. The evolutionary history of the flea species could best be explained by the association between parasite and host (time spent on the host). The phylogeographic pattern of the fur flea with a narrower host range correspond to host spatial genetic structures, while the pattern in the host opportunistic nest flea correspond to higher genetic divergences between sampling localities that may also be associated with higher effective population sizes. These findings suggest that genetic exchange among localities are most likely explained by differences in the dispersal abilities and life histories of the flea species.

In chapter 3, we provide a genetic perspective on the taxonomy and evolution of the medically important flea, *Dinopsyllus ellobius*, and show strong evidence for the resurrection of *Dinopsyllus abaris*. *Dinopsyllus ellobius* is considered a common and widespread flea in southern Africa and can act as a vector for plague. Due to differences in the interpretation of geographic variation in male sternite VIII, the taxonomy of the species is characterised by uncertainty. The exact mechanisms that could have caused the diversification among lineages are not clear but the two lineages seem to be geographically separated and may have different ecological requirements. The present study strongly supports the notion that the two lineages are representative of *D. ellobius* (probably more associated with the host body and better adapted to mesic conditions) and *D. abaris* (most likely more associated with the host nest and diverse climatic conditions) as originally proposed based on the single morphological character confined to male sternite VIII.

In chapter 4, we investigate the influence of variation in life history strategies on species distribution (SD) modelling. Since SD modelling does not account for differences in life history,

Stellenbosch University https://scholar.sun.ac.za

the accuracy of predictions of current and future species' ranges could differ strongly between species groups. The current study provide supporting evidence that contemporary SD modelling, that includes climate and landscape feature variables, is a valuable tool to study the biogeography and future distributions of fleas. However, we provide evidence that flea species with different life histories may be differentially sensitive to environmental conditions.

From the study it is evident that parasite life history is an important aspect to consider in any evolutionary and ecological studies. In particular the level of host association (i.e. time spent on the host) has direct bearing on transmission of parasite species which can significantly influence the distribution, diversity and diversification of taxa.

Stellenbosch University https://scholar.sun.ac.za

References

Stellenbosch University https://scholar.sun.ac.za

Adjemian JCZ, Girvetz EH, Beckett L, Foley JE. 2006. Analysis of Genetic Algorithm for Rule-Set Production (GARP) Modeling Approach for Predicting Distributions of Fleas Implicated as Vectors of Plague, *Yersinia pestis*, in California. Journal of Medical Entomology 43:93-103.

Akaike H. 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov PN, Csaki F, (Eds). 1973. *Second International Symposium on Information Theory*. Adad. Kiado, Budapest, 267-281.

Allouche O, Tsoar A, Kadmon R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa, and the true skill statistic (TSS). Journal of Applied Ecology 43:1223-1232.

Alvarez N, McKey D, Kjelberg F, Hossaert-McKey M. 2010. Phylogeography and historical biogeography of obligate specific mutualism. In: Morand S, Krasnov BR, (Eds.). 2010. *The Biogeography of Host-Parasite Interactions*. New York: Oxford University Press, 31-39.

Anderson PC, Kok OB. 2003. Ectoparasites of springhares in the Northern Cape Province, South Africa. South African Journal of Wildlife Research 33:23-32.

Anderson TJ, Blouin MS, Beech RN. 1998. Population biology of parasitic nematodes: Applications of genetic markers. Advances in Parasitology 41:219–283.

Andreo V, Lima M, Provensal C, Priotto J, Polop J. 2009a. Population dynamics of two rodent species in agro-ecosystems of central Argentina: intra-specific competition, land-use, and climate effects. Population Ecology 51:297–306.

Andreo V, Provensal C, Scavuzzo M, Lamfri M, Polop J. 2009b. Environmental factors and population fluctuations of *Akodon azarae* (Muridae: Sigmontinae) in central Argentina. Austral Ecology 34:132–142.

Araújo MB, Nogués-Bravo D, Diniz-Filho JAF, Haywood AM, Valdes PJ, Rahbek C. 2008. Quaternary climate changes explain diversity among reptiles and amphibians. Ecography 31:8-15.

Austin M. 2007. Species distribution models and ecological theory: A critical assessment and some possible new approaches. Ecological Modelling 200:1-19.

Stellenbosch University https://scholar.sun.ac.za

Barrett LG, Thrall PH, Burdon JJ, Linde CC. 2008. Life history determines genetic structure and evolutionary potential of host-parasite interactions. Trends in Ecology and Evolution 23:678-685.

Bauer AM, Lamb T. 2005. Phylogenetic relationships of southern African geckos in the *Pachydactylus* group (Squamata: Gekkionidae). African Journal of Herpetology 48:53-62.

Beaucourno J-C, Horak IG, Fourie LJ. 2003. Fleas of elephant shrews (Mammalia, Macroscelididae), and a new host and locality record for *Macroscelidopsylla albertyni* De Meillon & Marcus, 1958 (Siphonaptera, Chimaeropsyllidae). Onderstepoort Journal of Veterinary Research 70:251-253.

Bidochka MJ, Kamp AM, Lavender TM, Dekoning J, De Croos JNA. 2001. Habitat association in two genetic groups of the insect- pathogenic fungus *Metarhizium anisopliae*: uncovering cryptic species? Applied and Environmental Microbiology 67:1335–1342.

Bitam I, de la Cruz KD, Parola P, Whiting MF, Raoult D. 2010. Flea and flea-borne diseases. International Journal of Infectious Disease 14:667-676.

Blackburn TM, Gaston KJ. 1999. The relationship between animal abundance and body size: a review of the mechanisms. Advances in Ecological Research 28:181–210.

Blouin MS, Liu J, Berry RE .1999. Life cycle variation and the genetic structure of nematode populations. Heredity 83:253–259.

Blouin MS, Yowell CA, Courtney CH, Dame JB. 1995. Host movement and the genetic structure of populations of parasitic nematodes. Genetics 141:1007-1014.

Blouin MS. 1998. Mitochondrial DNA diversity in nematodes. Journal of Helminthology 72:285–289.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Computational Biology 2014, 10:e1003537.

Stellenbosch University https://scholar.sun.ac.za

Braack LEO, Horak IG, Jordaan LC, Segerman J, Louw JP. 1996. The comparative host status of red veld rats (*Aethomys chrysophilus*) and bushveld gerbils (*Tatera leucogaster*) for epifaunal arthropods in the southern Kruger National Park. South Africa. Onderstepoort Journal of Veterinary Research 63:149-158.

Brinkerhoff RJ, Martin AP, Jones RT, Collinge SK. 2011. Population genetic structure of the prairie dog flea and plague vector, *Oropsylla hirsuta*. Parasitology 138:71-79.

Brower AVZ. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proceedings of the National Academy of Sciences 91:6491-6495.

Brown JH, Ernest SKM. 2002. Rain and rodents: complex dynamics of desert consumers. BioScience 52:979–87.

Brown JH. 1995. *Macroecology*. Chicago: University of Chicago Press.

Bryant D, Moulton V. 2004. NeighborNet: an agglomerative algorithm for the construction of planar phylogenetic networks. Molecular Biology and Evolution 21:255-265.

Burda H, Šumbera R, Begall S. 2007. Microclimate in burrows of subterranean rodents – Revisited. In: Begall S, Burda H, Schleich CE, (Eds.). 2007. *Subterranean rodents: News from underground*. Berlin: Springer-Verlag.

Cangi N, Horak IG, Apanaskevich DA, Matthee S, Das Neves LCBG, Estrada-Peña A, Matthee CA. 2013. The influence of interspecific competition and host preference on the phylogeography of two African Ixodid tick species. PLoS ONE 2013, 8:e76930.

Carlson TN, Ripley DA. 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sensing of Environment 62:241-252.

Caterino MS, Cho S, Sperling FAH. 2000. The current state of insect molecular systematics: a thriving tower of Babel. Annual Review of Entomology 45:1-54.

Stellenbosch University https://scholar.sun.ac.za

Cheng L, Connor TR, Sirén J, Aanensen DM, Corander J. 2013. Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. Molecular Biology and Evolution 5:1224-1228.

Clayton DH, Al-Tamimi S, Johnston KP. 2003. The ecological basis of coevolutionary history. In: Page RDM, (Ed.). 2003. Tangled trees: *Phylogeny, cospeciation and coevolution*. Chicago: University of Chicago Press, 310-342.

Clayton DH, Bush S, Johnson KP. 2004. Ecology of congruence: past meet present. Systematic Biology 53:165-173.

Clayton DH, Johnston KP. 2003. Linking coevolutionary history to ecological process: doves and lice. Evolution 57:2335-2341.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657-1659.

Coetzee JA. 1978. Climate and biological changes in south-western Africa during the Late Cainozoic. Palaeoecology of Africa 10:13-29.

Colwell RK. 2000. Rensch's rule crosses the line: convergent allometry of sexual size dimorphism in hummingbirds and flower mites. American Naturalist 156:495-510.

Combes C. 2001. *Parasitism. The ecology and evolution of intimate interactions*. Chicago: University of Chicago Press.

Cook BD, Page TJ, Hughes JM. 2008. Importance of cryptic species for identifying 'representative' units of biodiversity for freshwater conservation. Biological Conservation 141:2821-2831.

Cook BD. 1990. Rabbit burrows as environments for European rabbit fleas, *Spilopsyllus cuniculi* (Dale), in arid South Australia. Australian Journal of Zoology 38:317-325.

Corander J, Marttinen P, Sirén J, Tang J. 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics 2008, 9:539.

Stellenbosch University https://scholar.sun.ac.za

Coreau A, Martin J-L. 2007. Multi-scale study of bird species distribution and of their response to vegetation change: a Mediterranean example. Landscape Ecology 22:747-764.

Criscione CD, Poulin R, Blouin MS. 2005. Molecular ecology of parasites: elucidating ecological and microevolutionary processes. Molecular Ecology 14:2247-2257.

Cumming GS, van Vuuren DP. 2006. Will climate change affect ectoparasite species ranges? Global Ecology and Biogeography 15:486-497.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9:772.

Daszak P, Cunningham AA, Hyatt AD. 2000. Emerging infectious diseases of wildlife – threats to biodiversity and human health. Science 287:443-449.

Davis DHS. 1953. Plague in Africa from 1935 to 1949. Bulletin of the World Health Organization 9:665-700.

Davis MB, Shaw RG, Etterson JR. 2005. Evolutionary response to changing climate. Ecology 86:1704–1714.

de Graaf G. 1981. *The rodents of southern Africa*. Durban: Butterworths.

de la Cruz KD, Whiting MF. 2003. Genetic and phylogeographic structure of populations of *Pulex simulans* (Siphonaptera) in Peru inferred from two genes (*CytB* and *CoII*). Parasitology Research 91:55–59.

de Meeús T. 2000. Adaptive diversity, specialization, habitat preference and parasites. In: Poulin R, Morand S, Skorping A, (Eds.). 2000. *Evolutionary Biology of Host-Parasite relationships: Theory Meets Reality*. 2000. Elsevier, 27-42.

de Meillon B, Davis D, Hardy F. 1961. *Plaque in southern Africa, vol.1. The Siphonaptera (excluding Ischnopsyllidae).* Government Printer, Pretoria, South Africa.

Dollar E, Goudie A. 2000. Environmental change. In: Fox R, Rowntree K (Eds.), *The geography of South Africa in a changing world*. Oxford: Oxford University Press, 31-59.

Stellenbosch University https://scholar.sun.ac.za

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

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29:1969- 1973.

du Plessis A, Kerley GIH, deo Winter PE. 1992. Refuge microclimates of rodents: a surface nesting *Otomys unisulacatus* and a burrowing *Parotomys brantsii*. Acta Theriologica 37:351-358.

du Toit N, Jansen van Vuuren B, Matthee S, Matthee CA. 2012. Biome specificity of distinct genetic lineages within the four-striped mouse *Rhabdomys pumilio* (Rodentia: Muridae) from southern Africa with implications for taxonomy. Molecular Phylogenetics and Evolution 65:75-86.

du Toit N, Jansen van Vuuren B, Matthee S, Matthee CA. 2013a. Biogeography and host related factors trump parasite life history: limited congruence among the genetic structures of specific ectoparasitic lice and their rodent hosts. Molecular Ecology 22:5185–5204.

du Toit N, Matthee S, Matthee CA. 2013b. The sympatric occurrence of two genetically divergent lineages of sucking louse, *Polyplax arvicanthis* (Phthiraptera: Anoplura), on the four-striped mouse genus, *Rhabdomys* (Rodentia: Muridae). Parasitology 140:604-616.

Dunlop JA, Giribet G. 2003. The First Fossil Cyphophthalmid (Arachnida, Opiliones) from Bitterfeld Amber, Germany. Journal of Arachnology 31:371-378.

Edwards S, Claude J, Van Vuuren BJ, Matthee CA. 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.

Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ. 2011. A statistical explanation MaxEnt for ecologists. Diversity and Distributions 17:43-57.

Engelbrecht A, Matthee CA, Ueckermann EA, Matthee S. 2014. Evidence of cryptic speciation in mesostigmatid mites from South Africa. Parasitology 141:1322-1332.

Stellenbosch University https://scholar.sun.ac.za

Engelbrecht A, Taylor PJ, Daniels SR, Rambau RV. 2011. Cryptic speciation in the southern African vlei rat *Otomys irroratus* complex: evidence derived from mitochondrial *cyt b* and niche modelling. Biological Journal of the Linnean Society 104:192-206.

Erwin DH. 2009. Climate as a driver of evolutionary change. Current Biology 19:R575-R583.

ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Estrada-Peña A, Estrada-Sánchez A, de la Fuente J. 2014. A global set of Fourier-transformed remotely sensed covariates for the description of abiotic niche in epidemiological studies of tick vector species. Parasites & Vectors 2014, 7:302.

Estrada-Peña A, Estrada-Sánchez A, Estrada-Sánchez D, de la Fuente J. 2013. Assessing the effects of variables and background selection on the capture of the tick climate niche. International Journal of Health Geographics 2013, 12:43.

Excoffier L, Lischer HEL. 2010. ARLEQUIN suite version 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10:564-567.

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

Fairburn DJ. 1997. Allometry for sexual size dimorphism: pattern and process in the coevolution of body size in males and females. Annual Review of Ecology and Systematics 28:659-687.

Fairburn DJ. 2005. Allometry for sexual size dimorphism: testing two hypotheses for Rensch's rule in the water strider *Aquarius remigis*. American Naturalist 166:S69-S84.

Felsenstein J. 2004. *Inferring Phylogenies*. Sunderland: Sinauer Associates.

Fermino BR, Viola LB, Paiva F, Garcia HA, de Paula CD, Botero-Arias R, Takata CSA, Campaner M, Hamilton PB, Camargo EP, Teixeira MMG. 2013. The phylogeography of trypanosomes

Stellenbosch University https://scholar.sun.ac.za

from South American alligatorids and African crocodilids is consistent with the geological history of South American river basins and the transoceanic dispersal of *Crocodylus* at the Miocene. Parasites & Vectors 2013, 6:313.

Franklin J. 2009. *Mapping species distributions: spatial inference and prediction*. New York: Cambridge University Press.

Froeschke G, van der Mescht L, McGeoch M, Matthee S. 2013. Life history strategy influences parasite responses to habitat fragmentation. International Journal of Parasitology 45:1109-1118.

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

Gage KL, Kasoy MY. 2005. Natural history of plague: Perspectives from more than a centuary of research. Annual Review of Entomology 50:505-528.

Galbreath KE, Hoberg EP: Return to Beringia. 2012. Parasites reveal cryptic biogeographic history of North American pikas. Proceedings of the Royal Society of London B 279:371-378.

Gaston KJ, Blackburn TM. 1995. Birds, body size and the threat of extinction. Philosophical Transactions of the Royal Society London B 347:205–212.

Gaston KJ. 2003. *The structure and dynamics of geographic ranges*. Oxford: Oxford University Press.

Gómez-Díaz E, González-Solís J, Peinado MA, Page RDM. 2007. Lack of host-dependent genetic structure in ectoparasites of *Calonectris* shearwaters. Molecular Ecology 16:5204-5215.

Hafner MS, Nadler SA. 1988. Phylogenetic trees support the coevolution of parasites and their hosts. Nature 332:258-259.

Hafner MS, Sudman PD, Villablanca FX, Spradling TA, Demastes JW, Nadler SA. 1994. Disparate rates of molecular evolution in cospeciating hosts and parasites. Science 265:1087-1090.

Stellenbosch University https://scholar.sun.ac.za

Hafner, M. S., Demastes, J. W., Spradling, T. A. & Reed, D. L. 2003. Cophylogeny between pocket gophers and chewing lice. In: Page RD, (Ed.). 2003. *Tangled trees: phylogeny, cospeciation, and coevolution*. Chicago: University of Chicago Press, 195–220.

Hall T. 2005. BioEdit, Biological sequence alignment editor for Win95/98/NT/2K/XP. Available at: http://www.mbio.ncsu.edu/BioEdit/bioedit.html.

Hastriter MW. 1998. *Chiastopsylla tetratricha* (Siphonaptera: Chimaeropsyllidae), a new species from the Cape Province, South Africa. Journal of African Zoology 112:97-102.

Hastriter MW. 2000. *Echidnophaga suricatta* (Siphonaptera: Pulicidae), a new species of flea from the Northern Cape Province, South Africa. African Zoology 35:77-83.

Hengl T, de Jesus JM, MacMillan RA, Batjes NH, Heuvelink GBM, Ribeiro E, Samuel-Rosa A, Kempen B, Leenaars JGB, Walsh MG, Gonzalez MR. 2014. SoilGrids1km – Global soil information based on automated mapping. PLoS ONE 2014, 9:e105992.

Hernandez-Vera G, Mitrovic M , Jovic J , Tosveski I, Caldara R, Gassmann A, Emerson BC. 2010. Host-associated genetic differentiation in a seed parasitic weevil *Rhinusa antirrhini* (Coleptera: Curculionidae) revealed by mitochondrial and nuclear sequence data. Molecular Ecology 24:2286–2300.

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

Hewitt GM. 2001. Speciation, hybrid zones and phylogeography or seeing genes in space and time. Molecular Ecology 10:537–549.

Hewitt GM. 2011. Quaternary phylogeography: the roots of hybrid zones. Genetica 139:617-638.

Hillis DM, Dixon MT. 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. Quarterly Review in Biology 4:411-53.

Hoberg EP, Brooks DR. 2008. A macroevolutionary mosaic: episodic host-switching, geographical colonization and diversification in complex host-parasite systems. Journal of Biogeography 35:1533-1550.

Stellenbosch University https://scholar.sun.ac.za

Hoberg EP, Klassen GJ. 2002. Revealing the faunal tapestry: coevolution and historical biogeography of hosts and parasites in marine systems. Parasitology 124:S3–S22.

Holt RD. 1990. The microevolutionary consequences of climate change. Trends in Ecology and Evolution 5:311-315.

Hopkins GHE, Rothschild M. 1966. *An Illustrated Catalogue of the Rothschild Collection of Fleas (Siphonaptera) in the British Museum (Natural History). Vol. IV. Hystrichopsyllidae Concluded*. British Museum (Natural History), London.

Hopkins GHE. 1957. Host associations of Siphonaptera. In. Krasnov BR, (Ed.). 2008. *Functional and Evolutionary Ecology of Fleas: A Model for Ecological Parasitology*. New York: Cambridge University Press.

Horak IG, Beaucournu JC, and Braack LEO. 2004. Parasites of domestic and wild animals in South Africa. XLIV. Fleas (Insecta: Siphonaptera: Pulicidae) collected from 15 carnivore species. Onderstepoort Journal of Veterinary Research 71:9-14.

Horak IG, Boomker J. 1998. Parasites of domestic and wild animals in South Africa. XXXV. Ixodid ticks and bot fly larvae in the Bontebok National Park. Onderstepoort Journal of Veterinary Research 65:205-211.

Horak IG, Braack LEO, Fourie LJ, Walker JB. 2000. Parasites of domestic and wild animals in South Africa. XXXVIII. Ixodid ticks collected from 23 wild carnivore species. Onderstepoort Journal of Veterinary Research 67:239–250.

Horak IG, Chaparro F, Beaucournu J-C, Louw JP. 1999. Parasites of domestic and wild animals in South Africa. XXXVI. Arthropod parasites of yellow mongooses, *Cynictis penicillata*. Onderstepoort Journal of Veterinary Research 66:33-38.

Horak IG, Fourie LJ, Braack LEO. 2005. Small mammals as hosts of immature ixodid ticks. Onderstepoort Journal of Veterinary Research 72:255-261.

Stellenbosch University https://scholar.sun.ac.za

Horak IG, Fourie LJ. 1986. Parasites of domestic and wild animals in South Africa. XIX. Ixodid ticks and fleas on rock dassies (Procavia capensis) in the Mountain Zebra National Park. Onderstepoort Journal of Veterinary Research 53:123-126.

Horak IG, Guillard AJ, Moolman LL, De Vos V. 1987. Parasites of domestic and wild animals in South Africa. XXII. Ixodid ticks on domestic dogs and wild carnivores. Onderstepoort Journal of Veterinary Research 54:573-580.

Horak IG, Sheppey K, Knight MM, Beuthin CL. 1986. Parasites of domestic and wild animals in South Africa. XXI. Arthropod parasites of vaal ribbok, bontebok and scrub hares in the Western Cape Province. Onderstepoort Journal of Veterinary Research 53:187-197.

Howell, DJ, Petney TN, and Horak IG. 1989. The host status of the striped mouse, *Rhabdomys pumilio*, in relation to the tick vectors of heartwater in South Africa. Onderstepoort Journal of Veterinary Research 56:289-291.

Hudson BD. 1994. Soil organic matter and available water capacity. Journal of Soil and Water Conservation 49:189.

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

Huyse T, Poulin R, Théron A. 2005. Speciation in parasites: a population genetics approach. Trends in Parasitology 21:469-475.

Ingram A. 1927. Plague investigations in South Africa from an etiological aspect. Publications of the South African Institute for Medical Research 3:222-256.

Ioff IG. 1941. Ecology of Fleas in Relevance to Their Medical Importance. Pyatygorsk Publishers, Pyatygorsk, USSR (in Russian). In: Krasnov BR, Burdelova N, Shenbrot GI, Khokhlova IS. 2002. Annual cycles of four flea species in the central Negev desert. Medical and Veterinary Entomology 16:266-276.

Iturralde-Vinent MA, MacPhee RDE. 1996. Age and Paleogeographical origin of Dominican amber. Science 273:1850-1852.

Stellenbosch University https://scholar.sun.ac.za

Jaksic FM, Lima M. 2003. Myths and facts on ratadas: bamboo blooms, rainfall peaks and rodent outbreaks in South America. Austral Ecology 28:237–51.

Jansen van Vuuren B, Robinson TJ. 1997. Genetic population structure in the yellow mongoose, *Cynictis penicillata*. Molecular Ecology 6:1147-1153.

Johnson KP, Adams RJ, Page RDM, Clayton DH. 2003. When do parasites fail to speciate in response to host speciation? Systematic Biology 52:37-47.

Jones PH, Britten HB. 2010. The absence of concordant population genetic structure in the black-tailed prairie dog and the flea, *Oropsylla hirsuta*, with implications for the spread of *Yersinia pestis*. Molecular Ecology 19:2038-2049.

Jordan K. 1930. Siphonaptera: on some South African fleas. In: Rothschild FRS, Hartert E, Jordan K, (Eds). 1930. *Novitates Zoologicae* A Journal of Zoology. Vol. XXXVI, 1930-31. Zoological Museum, Tring, 129-138.

Kandul NP, Lukhtanov VA, Dantchenko AV, Coleman JWS, Sekercioglu CH, Haig D, Pierce NE. 2004. Phylogeny of *Agrodiaetus* Hübner 1822 (Lepidoptera: Lycaenidae) inferred from mtDNA sequences of *COI* and *COII* and nuclear sequences of *EF1-α*: Karyotype diversification and species radiation. Systematic Biology 53:278-289.

Kendall DG. 1948. On the generalized “birth-and-death” process. *Annals of Mathematical* Statisitics 19:1-15.

Khokhlova IS, Fielden LJ, Degen AA, Krasnov BR. 2012. Digesting blood of an auxiliary host in fleas: effect of phylogenetic distance from a principal host. The Journal of Experimental Biology 215:1259-1265.

Khokhlova IS, Serobyan V, Degen AA, Krasnov BR. 2010. Host gender and offspring quality in a flea parasitic on a rodent. Journal of Experimental Biology 213:680-689.

Koufopanou V, Burt A, Szaro T, Taylor JW. 2001. Gene genealogies, cryptic species, and molecular evolution in the human pathogen *Coccidioides immitis* and relatives (Ascomycota, Onygenales). Molecular Biology and Evolution 18:1246–1258.

Stellenbosch University https://scholar.sun.ac.za

Krasnov BR, Burdelov SA, Khokhlova IS, Burdelova NV. 2003a. Sexual size dimorphism, morphological traits and jump performance in seven species of desert fleas (Siphonaptera). Journal of Zoology (London) 261:181-189.

Krasnov BR, Sarfati M, Arakelyan MS, Khokhlova IS, Burdelova NV, Degen AA. 2003b. Host specificity and foraging efficiency in blood-sucking parasite: feeding patterns of the flea *Parapulex chephrenis* on two species of rodents. Parasitological Research 90:393-399.

Krasnov BR, Khokhlova IS, Fielden LF, Burdelova N. 2002a. The effect of substrate on survival and development of two species of desert fleas (Siphonaptera: Pulicidae). Parasite 9:135-142.

Krasnov BR, Khokhlova IS, Fielden LJ, Burdelova NV. 2002b. Time of survival under starvation in two flea species (Siphonaptera: Pulicidae) at different air temperatures and relative humidities. Journal of Vector Ecology 27:70-81.

Krasnov BR, Khokhlova IS, Fielden LJ, Burdelova NV. 2001a. The effect of temperature and humidity on the survival of pre-imaginal stages of two flea species (Siphonaptera: Pulicicdae). Journal of Medical Entomology 38:629-637.

Krasnov BR, Khokhlova IS, Fielden LJ, Burdelova NV. 2001b. Development rates of two Xenopsylla flea species in relation to air temperature and humidity. Medical and Veterinary Entomology 15:249-258.

Krasnov BR, Mouillot D, Shenbrot GI, Khokhlova IS, Poulin R. 2010a. Deconstructing spatial patterns in species composition of ectoparasite communities: the relative contribution of host composition, environmental variables and geography. Global Ecology and Biogeography 19:515-526.

Krasnov BR, Mouillot D, Shenbrot GI, Khokhlova IS, Vinarski MV, Korallo-Vinarskaya NP, Poulin R. 2010b. Similarity in ectoparasite faunas of Palaearctic rodents as a function of host phylogenetic, geographic or environmental distances: Which matters the most? International Journal for Parasitology 40:807–817.

Stellenbosch University https://scholar.sun.ac.za

Krasnov BR, Pilosof S, Shenbrot GI, Khokhlova IS, Degen AA. 2014. Phylogenetic structure of host spectra in Palaearctic fleas: stability versus spatial variation in widespread, generalist species. Parasitology 141:181-191.

Krasnov BR, Pilosof S, Shenbrot GI, Khokhlova IS. 2013a. Spatial variation in the phylogenetic structure of flea assemblages across geographic ranges of small mammalian hosts in the Palearctic. International Journal for Parasitology 43:763-770.

Krasnov BR, Vinarski MV, Korallo-Vinarskaya NP, Khokhlova IS. 2013b. Ecological correlates of body size in gamasid mites parasitic on small mammals: abundance and niche breadth. Ecography 36:1042-1050.

Krasnov BR, Poulin R, Shenbrot GI, Mouillot D, Khkhlova IS. 2005a. Host specificity and geographic range in haematophagous ectoparasites. Oikos 108:449-456.

Krasnov BR, Shenbrot GI, Khokhlova IS, Poulin R. 2005b. Diversification of ectoparasite assemblages and climate: an example of fleas parasitic on small mammals. Global Ecology and Biogeography 14:167-175.

Krasnov BR, Stanko M, Miklisova D, Morand S. 2005c. Distribution of fleas (Siphonaptera) among small mammals: Mean abundance predicts prevalence via simple epidemiological model. International Journal of Parasitology 35:1097-1101.

Krasnov BR, Poulin R. 2010. Ecological properties of a parasite: species-specific stability and geographical variation. In: Morand S, Krasnov BR, (Eds.). 2010. *The Biogeography of Host-Parasite Interactions*. New York: Oxford University Press, 99-113.

Krasnov BR, Shenbrot GI, Khokhlova IS, Stanko M, Morand S, Mouillot D. 2015. Assembly rules of ectoparasite communities across scales: combining patterns of abiotic factors, host composition, geographic space, phylogeny and traits. Ecography 38:184-197.

Krasnov BR, Shenbrot GI, Medvedev SG, Khokhlova IS, Vatschenok VS. 1998. Habitat dependence of a parasite-host relationship: Flea (Siphonaptera) assemblages in two gerbil species of the Negev Desert. Journal of Medical Entomology 35:303-313.

Stellenbosch University https://scholar.sun.ac.za

Krasnov BR, Shenbrot GI, Medvedev SG, Vatschenok VS, Khokhlova IS. 1997. Host-habitat relations as an important determinant of spatial distribution of flea assemblages (Siphonaptera) on rodents in the Negev Desert. Parasitology 114:159-173.

Krasnov BR. 2008. *Functional and Evolutionary Ecology of Fleas: A Model for Ecological Parasitology*. New York: Cambridge University Press.

Krasnov BR. Poulin R. Shenbrot GI. Mouillot D. Khokhlova IS. 2004. Ectoparasitic “Jack-of-All-Trades”: Relationship between Abundance and Host Specificity in Fleas (Siphonaptera) Parasitic on Small Mammals. American Naturalist 164:506-516.

Laundre JW, Reynolds TD. 1993. Effects of soil structure on burrow characteristics of five small mammal species. Great Basin Naturalist 53:358-366.

Lawrence AL, Brown GK, Peters B, Spielman DS, Morin-Adeline V, Šlapeta J. 2014. High phylogenetic diversity of the cat flea (*Ctenocephalides felis*) at two mitochondrial DNA markers. Medical and Veterinary Entomology 28:330-336.

Ledger JA. 1980. The arthropod parasites of vertebrates in Africa south of the Sahara. Volume IV. Phthiraptera (Insecta). Publications of the South African Institute for Medical Research No.56. South African Institute for Medical Research, Johannesburg, SA.

Lewis RE, Grimaldi D. 1997. A pulicid flea in Miocene amber from the Dominican Republic (Insecta: Siphonaptera: Pulicidae). American Museum Novitates 3205:1-9.

Lewis RE. 1998. Résumé of the Siphonaptera (Insecta) of the World. Journal of Medical Entomology 35:377-389.

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

Lin G, Zhao F, Chen H, Deng X, Su J, Zhang T. 2014. Comparative phylogeography of the plateau zokor (*Eospalax baileyi*) and its host-associated flea (*Neopsylla paranoma*) in the Qinghai-Tibet Plateau. BMC Evolutionary Biology 2014, 14:180.

Stellenbosch University https://scholar.sun.ac.za

Linder HP, Johnson SD, Kuhlman M , Matthee CA, Nyffeler R, Swartz ER. 2010. Biotic diversity in the Southern African winter-rainfall region. Current Opinion in Environmental Sustainability 2:109-116.

Louw JP, Horak IG, Braack LEO. 1993. Fleas and lice on scrub hares (*Lepus saxatilis*) in South Africa. Onderstepoort Journal of Veterinary Research 60:95-101.

Louw JP, Horak IG, Horak ML, Braack LEO. 1995. Fleas, lice and mites on scrub hares (*Lepus saxatilis*) in Northern and Eastern Transvaal and in Kwazulu-Natal, South Africa. Onderstepoort Journal of Veterinary Research 62:133-137.

Luchetti A, Mantovani B, Pampiglione S, Trentini M. 2005. Molecular characterization of *Tunga trimamillata* and *T. penetrans* (Insecta, Siphonaptera, Tungidae): Taxonomy and genetic variability. Parasite 12:123-129.

Luchetti A, Trenti M, Pampiglione S, Fioravanti ML, Montovani B. 2007. Genetic variability of *Tunga penetrans* (Siphonaptera, Tungidae) sand fleas across South America and Africa. Parasitology Research 100:593-598.

MacArthur RH. 1955. Fluctuations of animal populations and a measure of community stability. Ecology 36:533-536.

MacArthur RH. 1972. *Geographical ecology*. Princeton: Princeton University Press.

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

Marshall AG. 1981. *The ecology of ectoparasite insects*. London: Academic Press.

Martinů J, Sychra O, Literák I, Capek M, Gustafsson DL, Stefka J. 2014. Host generalists and specialists emerging side by side: an analysis of evolutionary patterns in the cosmopolitan chewing louse genus *Menacanthus*. International Journal of Parasitology 45:63-73.

Matthee CA, Flemming AF. 2002. Population fragmentation in the southern rock agama, *Agama atra*: more evidence for vicariance in Southern Africa. Molecular Ecology 11:465-471.

Stellenbosch University https://scholar.sun.ac.za

Matthee CA, Robinson TJ. 1996. Mitochondrial DNA differentiation among geographic populations of Smith's red rock rabbit, *Pronolagus rupestris* (Mammalia: Lagomorpha). Heredity 76:514-523.

Matthee CA, Robinson TJ. 1997. Mitochondrial DNA phylogeography and comparative cytogenetics of the springhare, *Pedetes capensis* (Mammalia: Rodentia). Journal of Mammalian Evolution 4:53-73.

Matthee S, Horak IG, Beaucournu J-C, Durden LA, Ueckermann EA, McGeoch MA. 2007. Epifaunistic arthropod parasites of the four-striped mouse, *Rhabdomys pumilio*, in the Western Cape Province, South Africa. Journal of Parasitology 93:47-59.

Matthee S, Horak IG, van der Mescht L, Ueckermann EA, Radloff FGT. 2010a. Ectoparasite diversity on rodents at De Hoop Nature Reserve, Western Cape Province. African Zoology 45:213-224.

Matthee S, McGeogh MA, Krasnov BR. 2010b. Parasite-specific variation and the extent of male-biased parasitism; an example with a South African rodent and ectoparasitic arthropods. Parasitology 137:651-660.

Matthee S, van der Mescht L, Wilson B, Lamberski N. 2011. Flea diversity on small cranivores in the Northern Cape Province, South Africa. African Zoology 46:27-31.

McCarthy T, Rubidge B. 2005. *The Story of Earth and Life: A Southern African Perspective on a 4.6-Billion-Year Journey*. South Africa: Struik Publishers.

McCoy KD, Bouliner T, Tirard C, Michalakis Y. 2001. Host specificity of a generalist parasite: genetic evidence of sympatric host races in the seabird tick *Ixodes uriae*. Journal of Evolutionary Biology 14:395-405.

McCoy KD, Boulinier T, Tirard C, Michalakis Y. 2003. Host-dependent genetic structure of parasite populations: differential dispersal of seabird tick host races. Evolution 57:288-296.

Stellenbosch University https://scholar.sun.ac.za

Medvedev SG, Krasnov BR. 2006. Fleas: permanent satellites of small mammals. In: Morand S, Krasnov BR, Poulin R, (Eds.). 2006. *Micromammals and Macroparasites: From Evolutionary Ecology to Management*. New York: Springer-Verlag, 161-177.

Merow C, Smith MJ, Silander JA. 2013. A practical guide to MaxEnt for modelling species' distributions: what it does, and why inputs and settings matter. Ecography 36:1058-1069.

Meserve PL, Kelt DA, Milstead WB, Gutiérrez JR. 2003. Thirteen years of shifting top-down and bottom-up control. BioScience 53:633–46.

Miller MP. 2005. Alleles In Space: computer software for the joint-analysis of inter-individual spatial and genetic information. Journal of Heredity 96:722-724.

Mitchell JA. 1921. Plague in South Africa: Perpetuation and Spread of Infection by Wild Rodents. The Journal of Hygiene 20:377-382.

Monteiro, A, Pierce NE. 2001. Phylogeny of *Bicyclus* (Lepidoptera: Nymphalidae) inferred from COI, COII, and EF-1a gene sequences. Molecular Phylogenetics and Evolution 18:264–281.

Montgelard C, Matthee CA. 2012. Tempo of genetic diversification in southern African rodents: The role of Plio-Pleistocene climatic oscillations as drivers for speciation. Acta Oecologica 42:50-57.

Moran NA, van Dohlen CD, Baumann P. 1995. Faster evolutionary rates in endosymbiotic bacteria than in cospeciating insect hosts. Journal of Molecular Evolution 41:727–731.

Morand S, Bordes F, Pisanu B, Goüy de Bellocq J, Krasnov BR. 2010. The geography of defence. In: Morand S, Krasnov BR, (Eds.). 2010. *The Biogeography of Host-Parasite Interactions*. New York: Oxford University Press, 159-172.

Morand S, Krasnov BR, Poulin R, (Eds.). 2006. *Micromammals and Macroparasites*: *From Evolutionary Ecology to Management*. Berlin: Springer-Verlag.

Morand S. 2012. Phylogeography helps with investigating the building of human parasite communities. Parasitology 139:1966-1974.

Stellenbosch University https://scholar.sun.ac.za

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

Myneni RB, Hoffman S, Knyazikhin Y, Privette JL, Glassy J, Tian Y, Running SW. 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sensing of Environment 83:214-231.

Nadler SA, Pérez-Ponce de León G. 2011. Integrating molecular and morphological approaches for characterizing parasite cryptic species: implications for parasitology. Parasitology 138:1688-1709.

Nadler SA. 1995. Microevolution and the genetic structure of parasite populations. Journal of Parasitology 81:395-403.

Nieberding C, Jousselin E, Desdevises Y. 2010. The use of co-phylogeographic patterns to predict the nature of host–parasite interactions, and vice versa. In: Morand S, Krasnov BR, (Eds.). 2010. *The Biogeography of Host-Parasite Interactions*. Oxford: Oxford University Press; 631–641.

Nieberding C, Morand S, Douady CJ, Libois R, Michaux J. 2005. Phylogeography of a nematode (*Heligmosomoides polygyrus*) in the western Palearctic region: Persistence of northern cryptic populations during ice ages? Molecular Ecology 14:765-779.

Nieberding C, Morand S, Douady CJ, Libois R, Michaux J. 2006. Parasites and the island syndrome: the colonization of the western Mediterranean islands by Heligmosomoides polygyrus (Dujardin, 1845). Journal of Biogeography 33:1212-1222.

Nieberding C, Morand S, Libois R, Michaux J. 2004. A parasite reveals cryptic phylogeographical history of its host. Proceedings of the Royal Society of London B 271:2559-2568.

Nieberding CM, Durette-Desset M-C, Vanderpoorten A, Casanova JC, Ribas A, Deffontaine V, Feliu C, Morand S, Libois R, Michaux JR. 2008. Geography and host biogeography matter for understanding the phylogeography of a parasite. Molecular Phylogenetics and Evolution 47:538-554.

Stellenbosch University https://scholar.sun.ac.za

Nieberding CM, Morand S. 2006. Comparative phylogeography: The use of parasites for insights into host history. In: Morand S, Krasnov BR, Poulin R, (Eds.). 2006. *Micromammals and Macroparasites*. Japan: Springer; 277-293.

Njunwa KJ, Mwaiko GL, Kilonzo BS, Mhina JIK. 1989. Seasonal patterns of rodents, fleas and plague status in the Western Usambara Mountains, Tanzania. Medical and Veterinary Entomology 3:17-22.

Norrdahl K, Korpimäki E. 1995. Does predation risk constrain maturation in cyclic vole populations? Oikos 72:263–72.

Olival KJ, Dick CW, Simmons NB, Morales JC, Melnick DJ, Dittmar K, Perkins SL, Daszak P, Desalle R. 2013. Lack of population genetic structure and host specificity in the bat fly, *Cyclopodia horsfieldi*, across species of *Pteropus* bats in Southeast Asia. Parasites & Vectors 2013, 6:231.

Osacar-Jimenez JJ, Lucientes-Curdi J, Calvete-Margolles C. 2001. Abiotic factors influencing the ecology of wild rabbit fleas in north-eastern Spain. Medical and Veterinary Entomology 15:157-166.

Page RDM, Lee PL, Becher SA, Griffiths R, Clayton DH. 1998. A different tempo of mitochondrial DNA evolution in birds and their parasitic lice. Molecular Phylogenetics and Evolution 9:276–293.

Page RDM. 1994. Parallel phylogenies: reconstructing the history of host-parasite assemblages. Cladisitics 10:155-173.

Partridge TC, Maud RR. 1987. Geomorphic evolution of southern Africa since the Mesozoic. South African Journal of Geology 90:179-208.

Partridge TC. 1990. Cainozoic environmental changes in Southern Africa. South African Journal of Science 86:315-317.

Partridge TC. 1997. Evolution of landscapes. In: Cowling, RM, Richardson DM, Pierce SM, (Eds.). 1997. *Vegetation of Southern Africa*. Cambridge: Cambridge University Press.

Stellenbosch University https://scholar.sun.ac.za

Passy SI, Larson CA. 2011. Succession in stream biofilms is an environmentally-driven gradient of stress tolerance. Microbial Ecology 62:414–424.

Passy SI. 2012. A hierarchical theory of macroecology. Ecological Letters 15:923–934.

Paterson AM, Banks J. 2001. Analytical approaches to measuring cospeciation of hosts and parasites: through a glass, darkly. International Journal for Parasitology 31:1012-1022

Paterson AM, Wallis GP, Wallis LJ, Gray RD. 2000. Seabird and louse coevolution: complex histories revealed by 12S rRNA sequences and reconciliation analysis. Systematic Biology 49:383–399.

Pérez-Rodríguez A, Fernández-González S, De le Hera I. 2013. Finding the appropriate variables to model the distribution of vector-borne parasites with different environmental preferences: climate is not enough. Global Change Biology 19:3245-3253.

Perkins SL, Martinsen ES, Falk BG. 2011. Do molecules matter more than morphology? Promises and pitfalls in parasites. Parasitology 138:1664-1674.

Peterson AT, Papeş M, Eaton M. 2007. Transferability and model evaluation in ecological niche modelling: a comparison of GARP and MaxEnt. Ecography 30:550-560.

Petney TN, Horak IG, Howell DJ, Meyer S. 2004. Striped mice, *Rhabdomys pumilio*, and other murid rodents as hosts for immature ixodid ticks as in the Eastern Cape Province. Onderstepoort Journal of Veterinary Research 71:313-318.

Pettorelli N, Ryan S, Mueller T, Bunnefeld N, Jedrzejewska B, Lima M, Kausrud K. 2011. The Normalized Difference Vegetation Index (NDVI): unforeseen successes in animal ecology. Climate Research 46:15-27.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modelling of species geographic distributions. Ecological Modelling 190:231-259.

Phillips SJ, Dudik M. 2008. Modeling of species distributions with MaxEnt: new extensions and comprehensive evaluation. Ecography 31:161-175.

Stellenbosch University https://scholar.sun.ac.za

Poretta D, Mastrantonio V, Amendolia S, Gaiarsa S, Epis S, Genchi C, Bandi C, Otranto C, Urbanelli S. 2013a. Effects of global changes on the climatic niche of the tick *Ixodes ricinus* inferred by species distribution modelling. Parasites & Vectors 2013, 6:271.

Porretta D, Mastrantonio V, Mona S, Epis S, Montagna M, Sassera D, Bandi C, Urbanelli S. 2013b. The integration of multiple independent data reveals an unusual response to Pleistocene climatic changes in the hard tick *Ixodes ricinus*. Molecular Ecology 22:1666-1682.

Posada D, Buckley TR. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Systematic Biology 53:793-808.

Posada D. 2008. jModelTest: phylogenetic model averaging. Molecular Biology and Evolution 25:1253-1256.

Poulin R, Krasnov BR, Mouillot D. 2011. Host specificity in phylogenetic and geographic space. Trends in Parasitology 27:355-361.

Poulin R, Krasnov BR. 2010. Similarity and variability of parasite assemblages across geographical space. In: Morand S, Krasnov BR, (Eds.). 2010. *The Biogeography of Host-Parasite Interactions*. New York: Oxford University Press, 115-127.

Poulin R. 2005. Relative infection levels and taxonomic distances among the host species used by a parasite: insights into parasite specialization. Parasitology 130:109-115.

Poulin R. 2011. Uneven distribution of cryptic diversity among higher taxa of parasitic worms. Biology Letters 7:241–244.

R Development Core Team. 2014. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rambau RV, Robinson TJ, Stanyon R. 2003. Molecular genetics of *Rhabdomys pumilio* subspecies boundaries: mtDNA phylogeography and karyotypic analysis by fluorescence in situ hybridization. Molecular Phylogenetics and Evolution 28:564-575.

Stellenbosch University https://scholar.sun.ac.za

Rambaut A, Drummond AJ. 2007. Tracer v1.5, Available from http://beast.bio.ed.ac.uk/Tracer .

Reisen WK. 2010. Landscape epidemiology of vector-borne diseases. Annual Review of Entomology 55:461–483.

Ritzkowski S. 1997. K-Ar-Altersbestimmungen der bernsteinführenden Sedimente des Samlandes (Paläogen, Bezirk Kaliningrad). Metalla Bochum 66:19-23.

Robe JL, Machado S, Bolzan AR, dos Santos JPJ, Valer FB, dos Santos AP, Blauth ML, Gottschalk MS. 2014. Comparative ecological niche modeling and evolutionary ecology of Neotropical mycophagous Drosophilidae (Diptera) species. Studies on Neotropical Fauna and Environment 49:79-94.

Rose H, Wall R. 2011. Modelling the impact of climate change on spatial patterns of disease risk: Sheep blowfly strike by *Lucilia sericata* in Great Britain. International Journal for Parasitology 41:739-746.

Rosenzweig ML. 1973. Habitat selection experiments with a pair of coexisting heteromyid rodents species. Ecology 54:111–117.

Rothschild M, Clay T. 1952. *Fleas, Flukes and Cuckoos: A Study of Bird Parasites*. 3rd edition London: Collins. In: Krasnov BR, (Ed.). 2008. *Functional and Evolutionary Ecology of Fleas: A Model for Ecological Parasitology*. New York: Cambridge University Press.

Rózsa L. 2000. Spite, xenophobia, and collaboration between hosts and parasites.Oikos 91:14438-14444.

Russo IM, Chimimba CT, Bloomer P. 2010. Bioregion heterogeneity correlates with extensive mitochondrial DNA diversity in the Namaqua rock mouse, *Micaelamys namaquensis* (Rodentia: Muridae) from southern Africa – evidence for a species complex. BMC Evolutionary Biology 10:307.

Russo IRM, Chimimba CT, Bloomer P. 2006. Mitochondrial DNA differentiation between two species of *Aethomys* (Rodentia: Muridae) from southern Africa. Journal of Mammalogy 87:545-553.

Stellenbosch University https://scholar.sun.ac.za

Scott L, Anderson HM, Anderson JM. 1997. Vegetation history. In: Cowling RM, Richardson DM, Pierce SM, (Eds.), *Vegetation of southern Africa*. Cambridge: Cambridge University Press, pp. 62-84.

Scwan TG. 1986. Seasonal abundance of fleas (Siphonaptera) on grassland rodents in Lake Nakuru National Park, Kenya, and potential of plague transmission. Bulletin of Entomological Research 76:633-648.

Segerman J. 1995. *Siphonaptera of Southern Africa. Handbook for the identification of fleas. Publications of The South African Institute for Medical Research No. 57*. South African Institute for Medical Research, Johannesburg, South Africa.

Shenbrot G, Kransov B, Khokhlova I, Demidova T, Fielden L. 2002. Habitat-dependence difference in architecture and microclimate of the burrows of Sundevall's jird (*Merione crassus*) (*Rodentia: Gerbillinae*) in the Negev Desert, Israel. Journal of Arid Environments 51:265-279.

Shenbrot G, Krasnov B, Lu L. 2007. Geographical range size and host specificity in ectoparasites: a case study with *Amphipsylla* fleas and rodent hosts. Journal of Biogeography 34:1679-1690.

Sheperd AJ, Hummitzsch DE, Leman PA, Hartwig EK. 1983. Studies on plague in the Eastern Cape Province of South Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene 77:800-808.

Sheperd AJ, Leman PA. 1983. Plague in South African rodents 1972 – 1981. Transactions of the Royal Society of Tropical Medicine and Hygiene 77:208-211.

Singleton G, Krebs CJ, Davis S, Chambers L, Brown P. 2001. Reproductive changes in fluctuating house Mouse populations in south eastern Australia. Proceedings of the Royal Society London B 268:1741–8.

Sire C, Durand P, Pointier J-P, Theron A. 2001. Genetic diversity of *Schistosoma mansoni* within and among individual hosts (*Rattus rattus*): infrapopulation differentiation at microspatial scale. International Journal of Parasitology 31:1609-1616.

Stellenbosch University https://scholar.sun.ac.za

Smit HA, Robinson TJ, van Vuuren BJ. 2007. Coalescence methods reveal the impact of vacariance on the spatial genetic structure of *Elephantulus edwardii* (Afrotheria, Macroscelidae). Molecular Ecology 16:2680-2692.

Spencer R-J, Cavanough VC, Baxter GS, Kennedy MS. 2005. Adult free zones in small mammal populations: response of Australian native rodents to reduced cover. Austral Ecology 30:876–84.

Stark HE. 2002. Population dynamics of adult fleas (Siphonaptera) on hosts and in nests of the California vole. Journal of Medical Entomology 39:818-824.

Štefka, J, Hoeck PE, Keller LF, Smith VS. 2011. A hitchhikers guide to the Galapagos: co-phylogeography of Galapagos mockingbirds and their parasites. BMC Evolutionary Biology 2011, 11:284.

Stenseth, Nils Chr, Bakyt B. Atshabar, Mike Begon, Steven R. Belmain, Eric Bertherat, Elisabeth Carniel, Kenneth L. Gage, Herwig Leirs, and Lila Rahalison. 2008. Plague: Past, Present, and Future. PLoS Medicine 2008, 5:1.

Stephens M, Scheet P. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. American Journal of Human Genetics 76:449-462.

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

Stockwell DRB, Peterson AT. 2002. Effects of sample size on accuracy of species distribution models. Ecological Modelling 148:1-13.

Swofford DL. 2000. *PAUP: Phylogenetic Analysis Using Parsimony (and other methods)*. Sunderland: Sinauer Associates.

Till WM. 1963. Ethiopian mites of the genus Androlaelaps Berlese s. lat. (Acari: Mesostigmata). Bulletin of the Natural History Museum. Zoology 10:1-104.

Stellenbosch University https://scholar.sun.ac.za

Tipton VJ. 1960. The genus *Laelaps*. With review of the Laelaptinae and a new subfamily Alphalaelaptinae (Acarina: Laelaptinae). University of California Publications in Entomology 16:233-356.

Tolley KA, Chase BM, Forest F. 2008. Speciation and radiations track climate transitions since the Miocene climatic optimum: a case study of southern African chameleons. Journal of Biogeography 35:1402-1414.

Tolley KA, Makokha JS, Houniet DT, Swart BL, Matthee CA. 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.

Tripet F, Christe P, Møller AP. 2002. The importance of host spatial distribution for parasite specialization and speciation: a comparative study of bird fleas (Siphonaptera: Ceratophyllidae). Journal of Animal Ecology 71:735-748.

Tyson PD, Partridge TC. 2000. The evolution of Cenozoic climates. In. Partridge TC, Maud RR, eds., *The Cenozoic of Southern Africa*. New York: Oxford University Press.

van der Mescht L, le Roux PC, Matthee S. 2013. Remnant fragments within an agricultural matrix enhance conditions for a rodent host and its fleas. Parasitology 140:368-377.

van der Mescht L, Matthee S, Matthee CA. 2015a. Comparative phylogeography between two generalist flea species reveal a complex interaction between parasite life history and host vicariance: Parasite-host intimacy matters. BMC Evolutionary Biology 15:105.

van der Mescht L, Matthee S, Matthee CA. 2015b. A genetic perspective on the taxonomy and evolution of the medically important flea, *Dinopsyllus ellobius* (Siphonaptera: Dinopsyllinae), and the resurrection of *Dinopsyllus abaris*, Biological Journal of the Linnean Society 116:541-557.

van der Mescht L. 2011. Ectoparasite assemblage of the four-striped mouse, *Rhabdomys pumilio*: the effect of anthropogenic habitat transformation and temporal variation. MsC thesis. Stellenbosch University, South Africa.

Stellenbosch University https://scholar.sun.ac.za

Vashchenok VS. 1988. Fleas – vectors of pathogens causing diseases in humans and animals. Nuaka, in Russian. In: Krasnov BR, Shenbrot GI, Khokhlova IS, Stanko M, Morand S, Mouillot D. 2015. Assembly rules of ectoparasite communities across scales: combining patterns of abiotic factors, host composition, geographic space, phylogeny and traits. Ecography 38:184-197.

Venturi FP, Chimimba CT, van Aarde RJ, Fairall1 N. 2004. The distribution of two medically andagriculturally important cryptic rodent species, *Mastomys natalensis* and *M. coucha* (Rodentia: Muridae) in South Africa. African Zoology 39:235-245.

Wardhana AH, Cecchi G, Muharsini S, Cameron MM, Ready PD, Hall MJR. 2014. Environmental and phylogeographical determinants of the distribution of the Old World screwworm fly in Indonesia. Acta Tropica 138S:S62-S68.

Weitschat W, Wichard W. 2010. Baltic amber. In: Penney D, (Ed.). 2010. *Biodiversity of fossils in amber from the major world deposits*. Manchester: Siri Scientific Press.

Whiteman NK, Kimball RT, Parker PG. 2007. Co-phylogeography and comparative population genetics of the threatened Galápagos hawk and three ectoparasite species: ecology shapes population histories within parasite communities. Molecular Ecology 16:4759-4773.

Whiteman NK, Parker PG. 2005. Using parasites to infer host population history: a new rationale for parasite conservation. Animal Conservation 8:175-181.

Whiting MF, Whiting AS, Hastriter MW, de la Cruz KD. 2008. A molecular phylogeny of fleas (Insecta: Siphonaptera): origins and host associations. Cladistics 24:677-707.

Whiting MF. 2002. Mecoptera is paraphyletic: multiple genes and phylogeny of Mecoptera and Siphonaptera. Zoologica Scripta 31:93-104.

Whittaker R.J. Araújo MB, Jepson P, Ladle RJ, Watson JEM, Willis KJ. 2005. Conservation biogeography: assessment and prospect. Diversity and Distributions 11:3-23.

Stellenbosch University https://scholar.sun.ac.za

Willows-Munro S, Matthee CA. 2011. Linking lineage diversification to climate and habitat heterogeneity: phylogeography of the southern African shrew *Myosorex varius.* Journal of Biogeography 38:1976-1991.

Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A. NCEAS Predicting Species Distributions Working Group. 2008. Effects of sample size on the performance of species distribution models. Diversity and Distributions 14:763-773.

Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, Damgaard CF, Dormann CF, Forchhammer MC, Grytnes J-A, Guisan A, Heikkinen RK, Høye TT, Kühn I, Luoto M, Maiorano L, Nilsson M-C, Normand S, Öckinger E, Schmidt NM, Termansen M, Timmerman A, Wardle DA, Aastrup P, Svenning J-C. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biological Reviews 88:15-30.

Zhang Z, Pech R, Davis S, Shi D, Wan X, Zhong W. 2003. Extrinsic and intrinsic factors determine the eruptive dynamics of Brandt's voles *Microtus brandti* in Inner Mongolia, China. Oikos 100:299–310.

Zhu Q, Hastriter M, Whiting M, de la Cruz KD. 2015. Fleas (Siphonaptera) are Cretaceous, and evolved with Theria. Molecular Phylogenetics and Evolution 90:129-139.

Zumpt F. 1961. *The arthropod parasites of vertebrates in Africa south of the Sahara. Vol. I (Chelicerata)*. Publications of the South African Institute for Medical Research. South African Institute for Medical Research, Johannesburg, South Africa.

Stellenbosch University https://scholar.sun.ac.za

Appendix A

Figures

Stellenbosch University https://scholar.sun.ac.za

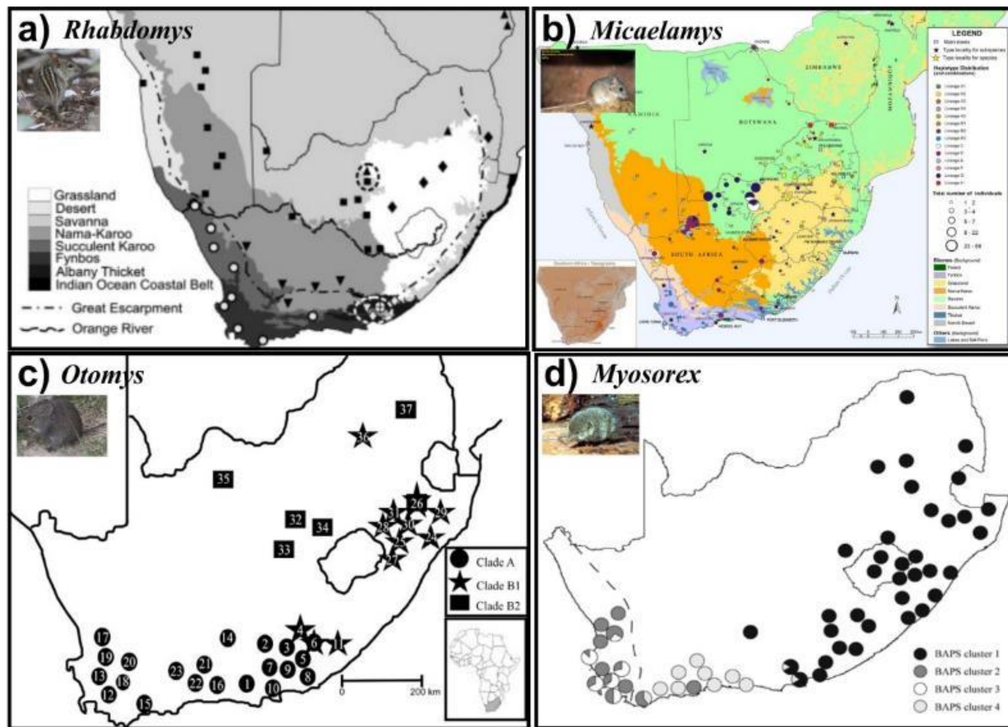


Figure A.1 Maps indicating published biogeographic patterns directly obtained for **a**) *Rhabdomys* (du Toit et al. 2012), **b**) *Micaelamys* (Russo et al. 2010), **c**) *Otomys* (Engelbrecht et al. 2011) and **d**) *Myosorex* (Willows-Munro & Matthee 2011).

Stellenbosch University https://scholar.sun.ac.za

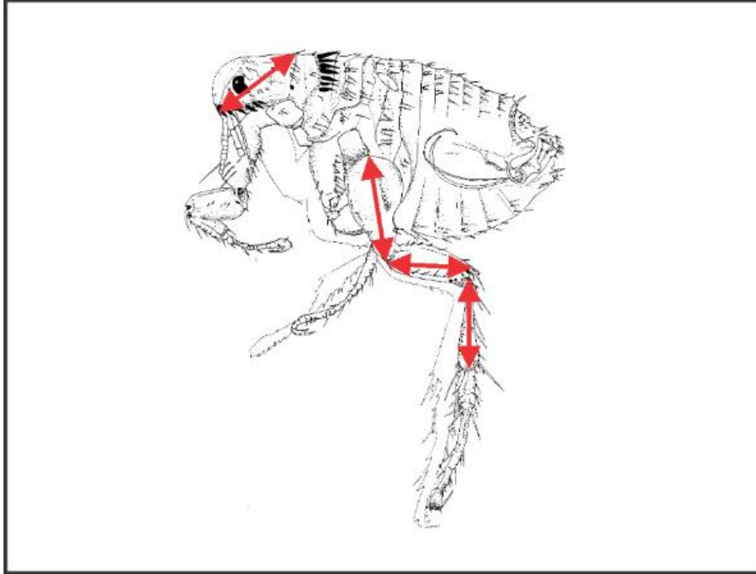


Figure A.2 Morphological measurements of size measured as length of head, coxa, femur and tibia of the left hind leg for males and females of both lineages.

Stellenbosch University https://scholar.sun.ac.za

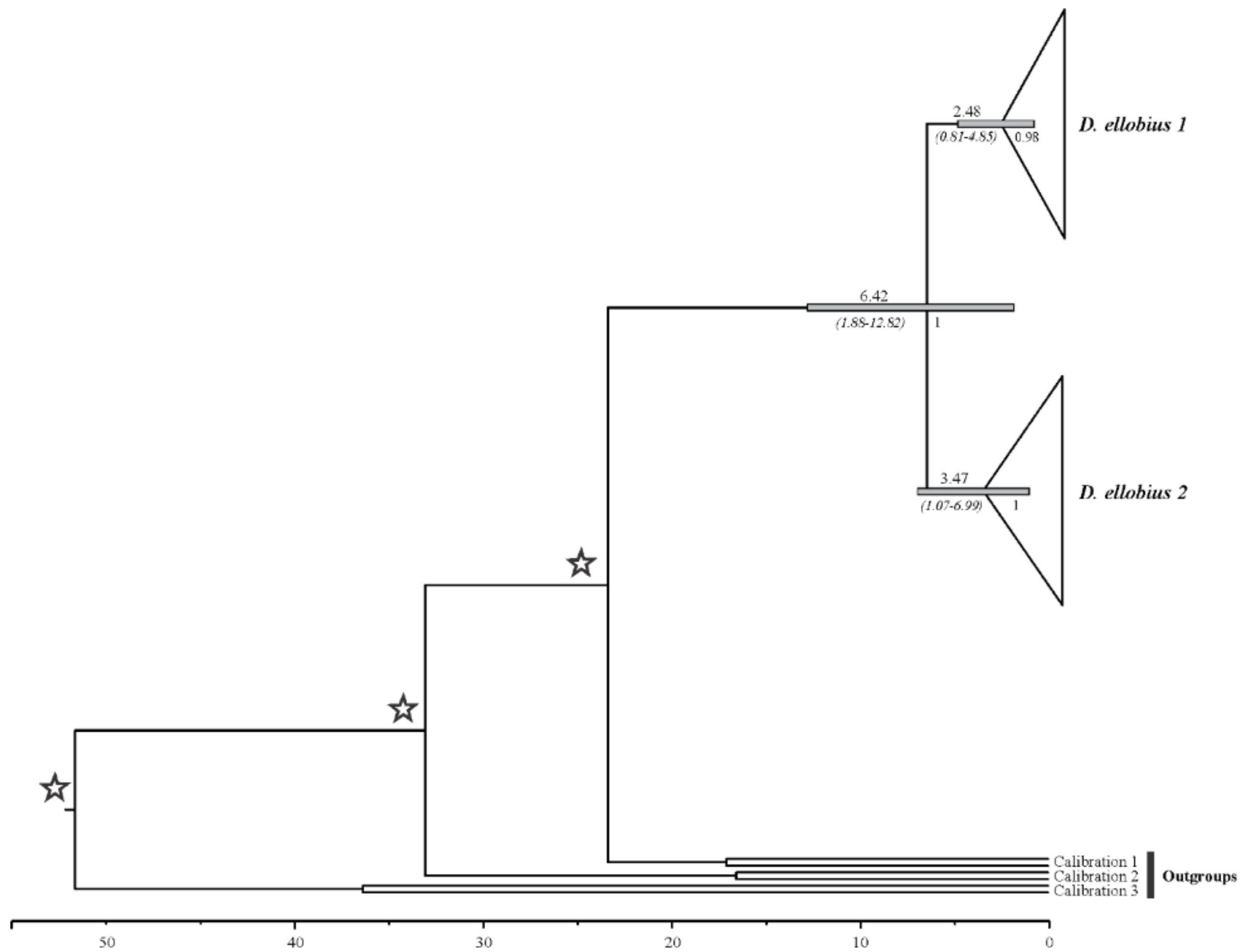


Figure A.3 Maximum lineage probability tree obtained from divergence dating in BEAST analysis. Posterior mean divergence dates in millions of years before present are indicated above nodes. The 95 % HPD credibility intervals are indicated by values below nodes and by grey shaded bars. Fossil calibrated nodes are indicated by stars (see Table B.5)**.**

Stellenbosch University https://scholar.sun.ac.za

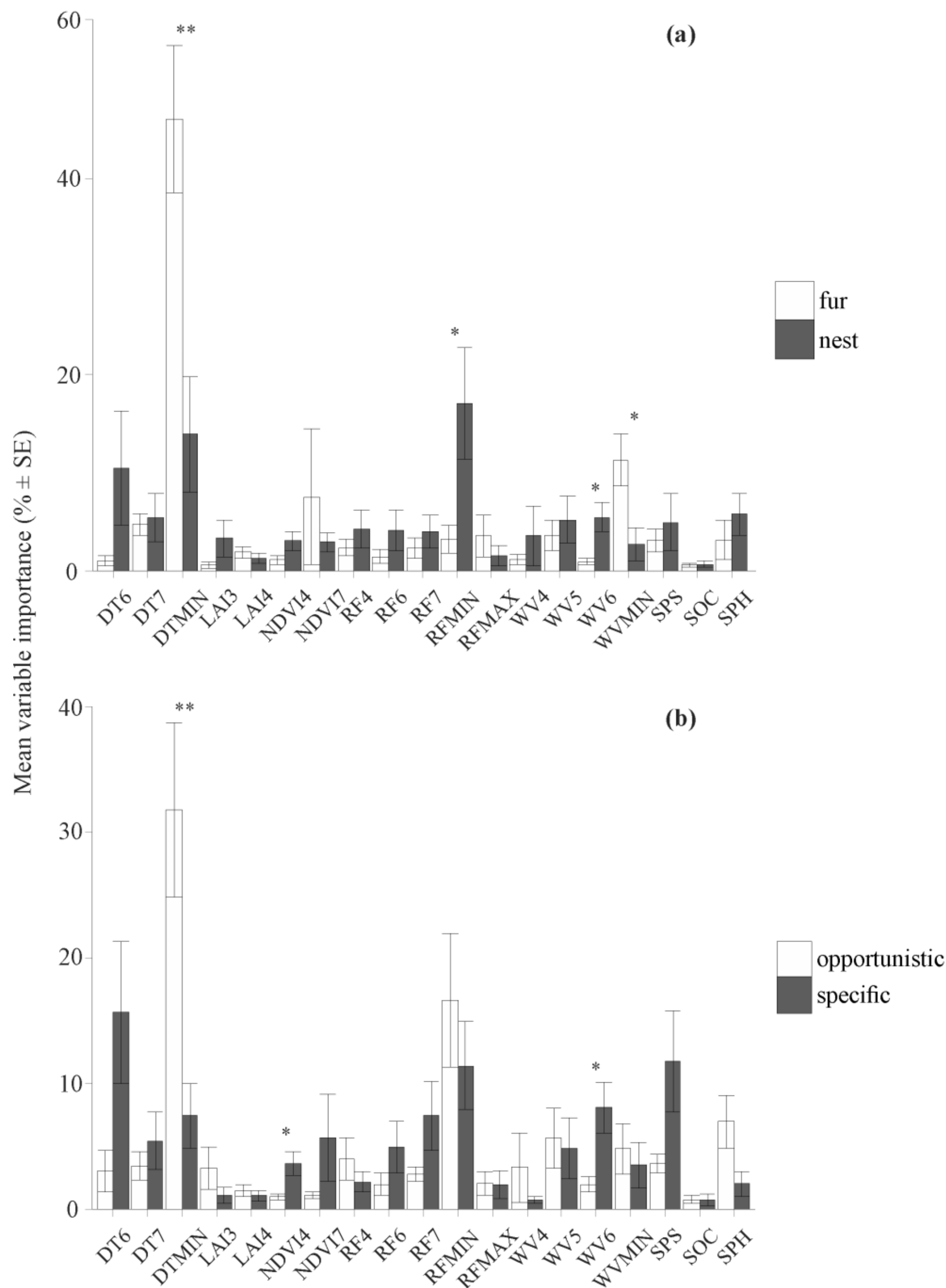


Figure A.4 Variable importance (i.e. percent relative predictor variable individual contribution) in MaxEnt models, averaged across flea species based on (a) microhabitat preference and (b) host specificity (see Table B.8 for variable reference code). Significant differences in the contribution of predictor variables between the two categories of species are indicated by asterisks: *** 0.001, ** 0.01, * 0.05.

Stellenbosch University https://scholar.sun.ac.za

Appendix B

Tables

Table B.1 Host identity and abundance, parasite prevalence and the number of specimens sequenced for each flea species per locality.

Province	Locality	Code	Geographic coordinates	Host	Host (*n*)	*Listropsylla agrippinae* Prevalence (%)	*Listropsylla agrippinae* Sequenced (n)	*Chiastopsylla rossi* Prevalence (%)	*Chiastopsylla rossi* Sequenced (n)
Western Cape	Anysberg	AB	-33.46 S 20.59 E	*Rhabdomys pumilio*	33	3	4	42	12
				Micealamys namaquensis	6	17	1	0	0
				Otomys irroratus	1	100	0	100	2
				Myotomys unisulcatus	10	0	0	0	0
				Elephantulus edwardii	1	0	0	0	0
				Total	51	6	5	29	14
	Beaufort West	BW	-32.22 S 22.80 E	*Rhabdomys intermedius*	33	6	10	36	10
				Micealamys namaquensis	6	33	2	0	0
				Otomys irroratus	4	100	1	75	4
				Myotomys unisulcatus	4	0	0	0	0
				Shrew spp.	2	0	0	0	0
				Total	49	12	13	31	14
	Buffeljagsrivier	BR	-34.05 S 20.53 E	*Rhabdomys pumilio*	27	19	6	59	8
				Otomys irroratus	8	0	0	25	2
				Shrew spp.	1	0	0	0	0
				Total	36	14	6	50	10
	Kanu	CS	-33.95 S 18.83 E	*Rhabdomys pumilio*	73	12	10	36	7
				Micealamys namaquensis	5	0	0	0	0
				Otomys irroratus	10	0	0	0	0
				Shrew spp.	6	0	0	0	0
				Mus minutoides	9	0	0	0	0
				Total	103	9	10	25	7

continued…

Province	Locality	Code	Geographic coordinates	Host	Host (*n*)	*Listropsylla agrippinae* Prevalence (%)	*Listropsylla agrippinae* Sequenced (n)	*Chiastopsylla rossi* Prevalence (%)	*Chiastopsylla rossi* Sequenced (n)
	Mooiplaas	MP	-33.92 S 18.75 E	*Rhabdomys pumilio*	35	11	4	34	9
				Otomys irroratus	1	0	0	0	0
				Elephantulus edwardii	1	0	0	100	2
				Myotomys unisulcatus	5	0	0	0	0
				Total	42	10	4	31	11
	Vanrhynsdorp	VR	-31.73 S 18.77 E	*Rhabdomys pumilio*	25	8	5	56	15
				Micealamys namaquensis	4	0	0	0	0
				Otomys irroratus	1	0	0	100	0
				Mus musculus	1	0	0	0	0
				Total	31	6	5	48	15
Northern Cape	Dronfield	DF	-28.74 S 24.77 E	*Rhabdomys intermedius*	32	9	0	34	4
				Tatera leucogaster	1	0	0	0	0
				Rattus rattus	1	0	0	0	0
				Mastomys coucha	2	0	0	0	0
				Total	36	8	0	31	4
	Garies	GS	-30.43 S 17.89 E	*Rhabdomys pumilio*	27	19	6	26	5
				Micealamys namaquensis	15	13	4	7	0
				Myotomys unisulcatus	2	0	0	0	0
				Elephantulus edwardii	2	0	0	0	0
				Gerbillurus paeba	1	0	0	0	0
				Total	47	15	10	17	5
	Loeriesfontein	LF	-30.95 S 19.44 E	*Rhabdomys intermedius*	31	48	12	48	15
				Micealamys namaquensis	16	31	1	0	0
continued…				*Myotomys unisulcatus*	5	40	2	40	1

Province	Locality	Code	Geographic coordinates	Host	Host (*n*)	*Listropsylla agrippinae* Prevalence (%)	*Listropsylla agrippinae* Sequenced (n)	*Chiastopsylla rossi* Prevalence (%)	*Chiastopsylla rossi* Sequenced (n)
				Elephantulus edwardii	3	0	0	0	0
				Total	55	40	15	31	16
	Springbok	GP	-29.70 S 18.03 E	*Rhabdomys pumilio*	56	23	12	2	1
				Micealamys namaquensis	31	0	0	6	0
				Myotomys unisulcatus	17	6	2	18	2
				Mus musculus	1	0	0	0	0
				Elephantulus edwardii	5	0	0	0	0
				Desmodillus auricularis	4	0	0	0	0
				Parotomys brantsii	2	0	0	0	0
				Total	116	12	14	5	3
	Sutherland	SL	-32.40 S 20.90 E	*Rhabdomys intermedius*	25	32	14	40	8
				Myotomys unisulcatus	9	33	0	67	0
				Gerbillurus paeba	1	0	0	0	0
				Total	35	31	14	46	8
	Three Sisters	TS	-31.89 S 23.15 E	*Rhabdomys intermedius*	5	20	1	20	0
				Micealamys namaquensis	6	17	2	0	0
				Myotomys unisulcatus	9	22	7	0	0
				Mus musculus	4	0	0	0	0
				Gerbillurus paeba	1	0	0	0	0
				Shrew spp.	2	0	0	0	0
				Total	27	15	10	4	0
Eastern Cape	Alice	AL	-32.79 S 26.85 E	*Rhabdomys dilectus*	6	0	0	17	3
				Otomys irroratus	8	0	0	13	1
continued…				*Rattus rattus*	2	0	0	0	0

Stellenbosch University https://scholar.sun.ac.za

Province	Locality	Code	Geographic coordinates	Host	Host (*n*)	*Listropsylla agrippinae* Prevalence (%)	*Listropsylla agrippinae* Sequenced (n)	*Chiastopsylla rossi* Prevalence (%)	*Chiastopsylla rossi* Sequenced (n)
				Mastomys natalensis	15	0	0	7	1
				Total	31	0	0	10	5
	Dohne	DE	-32.53 S 27.46 E	*Rhabdomys dilectus*	30	30	10	30	5
				Otomys irroratus	8	13	1	75	2
				Shrew spp.	6	0	0	17	3
				Total	44	23	11	36	10
	Fort Beaufort	FB	-32.78 S 26.63 E	*Rhabdomys pumilio*	8	25	3	38	4
				Rhabdomys dilectus	8	50	5	88	6
				Micealamys namaquensis	16	13	1	13	1
				Otomys irroratus	6	0	0	33	4
				Mus musculus	2	0	0	0	0
				Shrew spp.	2	0	0	50	1
				Total	42	19	9	36	16
	Hogsback	HB	-32.59 S 26.92 E	*Rhabdomys dilectus*	10	30	0	30	3
				Otomys irroratus	2	50	0	100	13
				Mus musculus	4	0	0	25	1
				Rattus rattus	1	0	0	0	0
				Shrew spp.	1	0	0	0	0
				Total	18	22	0	33	17
	The Croft	TC	-32.55 S 27.37 E	*Rhabdomys dilectus*	25	0	0	16	4
				Otomys irroratus	15	0	0	33	6
				Shrew spp.	3	0	0	33	1
				Total	43	0	0	23	11

continued…

Province	Locality	Code	Geographic coordinates	Host	Host (*n*)	*Listropsylla agrippinae* Prevalence (%)	*Listropsylla agrippinae* Sequenced (n)	*Chiastopsylla rossi* Prevalence (%)	*Chiastopsylla rossi* Sequenced (n)
Gauteng	Kaalplaas*	KP	-25.63 S 28.17 E	*Rhabdomys dilectus*	97	*	*	*	*
				Otomys irroratus	1	*	*	*	*
				Mus musculus	1	*	*	*	*
				Mastomys natalensis	44	*	*	*	*
				Shrew spp.	11	*	*	*	*
				Dasymys spp.	4	*	*	*	*
				Total	158	0	0	9	5
North West	Mooinooi*	MN	-25.47 S 27.33 E	*Micealamys chrysophilus*	8	*	*	*	*
				Mus musculus	7	*	*	*	*
				Mastomys natalensis	21	*	*	*	*
				Shrew spp.	3	*	*	*	*
				Lemniscomous rosalia	14	*	*	*	*
				Steatomys pratensis	10	*	*	*	*
				Total	63	0	0	13	8
KwaZulu-Natal	Albert Falls	AF	-29.47 S 30.40 E	*Mastomys natalensis*	26	0	0	15	5
				Otomys irroratus	2	0	0	100	4
				Shrew spp.	1	0	0	0	0
				Total	29	0	0	21	9

*Could not determine which host the samples were sequenced from.

Table B.2 Primers used for PCR amplification of mitochondrial and nuclear genes for the two flea species. COII amplification consisted of a denaturation cycle of 1 min at 95°C followed by a 10 cycle loop of 1 min at 95°C, 45°C, and 72°C, respectively. A 30 cycle loop was then performed with denaturation for 1min at 93°C followed by annealing for 1 min at the primer specific temperature, and 1 min extension at 72°C, followed by a final extension period of 5 min at 72°C. General PCR cycling conditions for the EF1-α region included an initial denaturation of 5 min at 94°C followed by 40 cycles of 30 s denaturation at 94°C, 45 s annealing at primer specific temperature, and 1 min extension at 72°C, followed by a final extension period of 7 min at 72°C. *Primers from Whiting (2002).

Taxon	Marker	Primer	F/R	Sequence (5'-3')	Annealing temperature (°C)
L agrippinae	COII	COII-2a*	F	ATAGAKCWTCYCCHTTAATAGAACA	50-52°C
		COII-9b*	R	GTACTTGCTTTCAGTCATCTWATG	50-52°C
	EF1-α	ForlagNEW	F	TTGGATGGCACCAAGTTGAC	59.6°C
		RevlagNEW	R	TGGCTTTCACTTTGGGAGTC	59.6°C
C rossi	COII	COII-2a*	F	ATAGAKCWTCYCCHTTAATAGAACA	50-52°C
		COII-9b*	R	GTACTTGCTTTCAGTCATCTWATG	50-52°C
	EF1-α	EF-1a M 44-1*	F	GCTGAGCGYGARCGTGGTATCAC	59.6°C
		EF-1a rcM 4.0*	R	ACAGVCACKGTYTGYCTCATRTC	59.6°C

Stellenbosch University https://scholar.sun.ac.za

Table B.3 Geographical coordinates of all localities sampled throughout South Africa.

Province	Locality	Code	Latitude	Longitude
Western Cape	Anysberg	AB	-33.97	18.97
	Beaufort West	BW	-32.22	22.80
	Blaauwberg	BB	-33.80	18.46
	Buffeljagsrivier	BR	-34.05	20.53
	Hottentotsholland	HT	-32.33	19.22
	Kanu	CS	-33.95	18.83
	Mooiplaas	MP	-33.92	18.75
	Vanrhynsdorp	VR	-31.73	18.77
Northern Cape	Dronfield	DF	-28.74	24.77
	Garies	GS	-30.43	17.89
	Loeriesfontein	LF	-30.95	19.44
	Springbok	GP	-29.70	18.03
	Sutherland	SL	-32.40	20.90
	Three Sisters	TS	-24.63	30.78
Eastern Cape	Alice	AL	-32.79	26.85
	Dohne	DE	-32.53	27.46
	East London	EL	-33.01	27.70
	Fort Beaufort	FB	-32.78	26.63
	Hogsback	HB	-32.59	26.92
	The Croft	TC	-32.55	27.37
Gauteng	Kaalplaas	KP	-25.63	28.17
	Rietvlei	RV	-25.70	26.67
North West	Mooinooi	MN	-25.47	27.33
	Woodside	WS	-26.09	25.36
KwaZulu-Natal	Albert Falls	AF	-29.47	30.40
	Chelmsford	CF	-27.96	29.93
	Inkunzi Lodge	IN	-28.56	31.24
	Mt Gilboa	MG	-29.94	30.29
	Mt Shannon	MS	-29.65	29.94
	Vernon Crookes	VC	-30.27	30.59
	Vryheid	VH	-27.75	30.80

Stellenbosch University https://scholar.sun.ac.za

Table B.4 Primers used for PCR amplification of the two gene fragments (see Whiting, 2002). Mitochondrial COII regions were amplified via a 'cold start' reaction consisting of a denaturation cycle of 1 min at 95°C followed by a 10 cycle loop of 1 min at 95, 45, and 72°C, respectively. A 30 cycle loop was then performed with denaturation for 1min at 93°C followed by annealing for 1 min at the primer specific temperature, and 1 min extension at 72°C, followed by a final extension period of 5 min at 72°C. General PCR cycling conditions for the nuclear EF1-α region included an initial denaturation of 5 min at 94°C followed by 40 cycles of 30 s denaturation at 94°C, 45 s annealing at primer specific temperature, and 1 min extension at 72°C, followed by a final extension period of 7 min at 72°C.

Gene	Primer	F/R	Sequence (5'-3')	Annealing temperature (°C)
COII	COII-2a	F	ATAGAKCWTCYCCHTTAATAGAACA	50-52°C
	COII-9b	R	GTACTTGCTTTCAGTCATCTWATG	50-52°C
EF1-α	EF-1a M 44-1	F	GCTGAGCGYGARCGTGGTATCAC	59.8°C
	EF-1a rcM 4.0	R	ACAGVCACKGTYTGYCTCATRTC	59.8°C

Stellenbosch University https://scholar.sun.ac.za

Table B.5 List indicating outgroups used for each calibration point in divergence dating with the respective GenBank accession numbers.

Calibration	**GenBank accession number**
Calibration 1: *Pulex*	
Pulex irritans	AF424041.1
Ctenocephalides canis	AF424040.1
Calibration 2: *Rhopalopsyllus*	
Tiamastus cavicola	EU335990.1
Parapsyllus longicornis	EU335985.1
Calibration 3: *Paleopsylla*	
Ctenophthalmus sanborni	EU335991.1
Ctenophthalmus formosanus	EU336024.1

Table B.6 A list of 10 characters assessed for each lineage with scores and morphological key from which it was extracted.

	Character		Key
Genal comb	(1) 5 conspicuous blunt spines	(0) 1 small inconspicuous spine	Segerman 1995
5th tarsal segment in males	(1) 4 pairs of lateral plantar bristles, 1 pair of pre-apical plantar spiniforms on fore- and mid 5th tarsi	(0) 5 pairs of lateral plantar bristles, at least 6 pre-apical plantar spiniforms on fore- and mid 5th tarsi	Segerman 1995
Frons	(1) Short, strongly rounded	(0) Longer, less strongly rounded	Segerman 1995
Vestigial tubercle	(1) About middle of frons	(0) Above middle of frons	Segerman 1995
Abdominal bristles	(1) Exceptionally long, main row reaching to or beyond bases of those in succeeding segment	(0) Shorter, main row not reaching bases of those in succeeding segement	Segerman 1995
Size	(1) Male less than 3 mm long, female less than 3.5 mm	(0) Usually considerably larger	Hopkins & Rothschild 1966
Male sternite VIII	(1) Emarginate dorso-posteriorly to form rounded-off right angle	(0) Wholly rounded except for small ventral sinus	Hopkins & Rothschild 1966
Male sternite IX	(1) Hardly expanded distally	(0) Clearly expanded distally	Segerman 1995
Female spermetheca	(1) Tail noticeably longer than bulga	(0) Shorter hilla	Segerman 1995
Male crochet phallosome	(1) Broad, smoothly rounded, truncate or slightly expanded apically	(0) More slender, tapered or sharply pointed	Segerman 1995
Pronotum	(1) 2 rows of bristles	(0) 3 rows of bristles	Segerman 1995

Table B.7 Hosts sampled, hosts sampled with *D. ellobius*, subsampled fleas, and the number of individuals of the two lineages from our subsample.

Locality	Code	Host species	Total host	Hosts with *D. ellobius*	Total *D. ellobius*	Subsampled fleas	D1 (*D. abaris*)	D2 (*D. ellobius*)
Anysberg	AB	*Rhabdomys pumilio*	14	0	0	0	0	0
		Micealamys namaquensis	8	0	0	0	0	0
		Otomys karoensis	4	0	0	0	0	0
		Myotomys unisulcatus	5	0	0	0	0	0
		Gerbillurus paeba	3	0	0	0	0	0
		Myosorex varius	1	1	7	6	6	0
			35	1	7	6	6	0
Bloubergstrand	BB	*Gerbillurus paeba*	3	3	8	8	0	8
			3	3	8	8	0	8
Hottentotsholland	HT	*Rhabdomys pumilio*	9	1	2	1	1	0
		Micealamys namaquensis	21	7	15	7	7	0
		Mus minutoides	1	0	0	0	0	0
		Otomys irroratus	1	0	0	0	0	0
			32	8	17	8	8	0
Vanrhynsdorp	VR	*Rhabdomys pumilio*	25	2	3	2	2	0
		Micealamys namaquensis	4	3	7	5	5	0
		Mus musculus	1	0	0	0	0	0
		Otomys irroratus	1	0	0	0	0	0
			31	5	10	7	7	0
Alice	AL	*Rhabdomys dilectus*	6	2	4	4	3	1
		Otomys irroratus	7	3	3	3	0	3
		Rattus rattus	2	0	0	0	0	0

Stellenbosch University https://scholar.sun.ac.za

Locality	Code	Host species	Total host	Hosts with *D. ellobius*	Total *D. ellobius*	Subsampled fleas	D1 (*D. abaris*)	D2 (*D. ellobius*)
continued…		*Mastomys natalensis*	14	6	6	5	0	5
			29	11	13	12	3	9
Dohne	DE	*Rhabdomys dilectus*	32	21	78	11	0	11
		Otomys irroratus	8	8	91	2	0	2
		Shrew spp.	6	1	1	1	0	1
			46	30	170	14	0	14
East London	EL	*Rhabdomys dilectus*	24	2	3	1	0	1
		Otomys irroratus	4	0	0	0	0	0
		Myotomys unisulcatus	1	0	0	0	0	0
		Mus musculus	4	0	0	0	0	0
		Mastomys natalensis	6	5	8	5	0	5
			39	7	11	6	0	6
Fort Beaufort	FB	*Rhabdomys pumilio*	8	1	2	2	2	0
		Rhabdomys dilectus	8	3	3	2	2	0
		Micealamys namaquensis	27	3	3	2	1	1
		Otomys irroratus	10	2	2	2	2	0
		Mus munitoides	1	0	0	0	0	0
		Saccostomys campestris	1	0	0	0	0	0
		Shrew spp.	2	0	0	0	0	0
			57	9	9	8	7	1
The Croft	TC	*Rhabdomys dilectus*	25	18	66	10	0	10
		Otomys irroratus	26	8	26	2	0	2
		Shrew spp.	3	1	1	1	0	1
			54	27	93	13	0	13

Locality	Code	Host species	Total host	Hosts with *D. ellobius*	Total *D. ellobius*	Subsampled fleas	D1 (*D. abaris*)	D2 (*D. ellobius*)
continued…								
Kaalplaas	KP	*Rhabdomys dilectus*	97	*	*	*	*	*
		Otomys irroratus	1	*	*	*	*	*
		Mus musculus	1	*	*	*	*	*
		Mastomys natalensis	44	*	*	*	*	*
		Shrew spp.	11	*	*	*	*	*
		Dasymys spp.	4	*	*	*	*	*
			158	10	19	3	0	3
Rietvlei	RV	*Rhabdomys dilectus*	24	1	1	1	0	1
		Otomys irroratus	2	0	0	0	0	0
		Mastomys natalensis	18	5	6	1	0	1
		Shrew spp.	9	0	0	0	0	0
			53	6	7	2	0	2
Mooinooi	MN	*Lemniscomys rosalia*	14	0	0	0	0	0
		Mastomys natalensis	21	2	4	3	0	3
		Aethomys chrysophilus	8	2	2	0	0	0
		Mus minutoides	7	2	2	2	0	2
		Steatomys pratensis	10	0	0	0	0	0
		Shrew spp.	3	0	0	0	0	0
			63	6	8	5	0	5
Woodside	WS	*Tatera brantsii*	20	6	12	8	8	0
		Mus indutus	4	0	0	0	0	0
			24	6	12	8	8	0
Chelmsford	CF	*Rhabdomys dilectus*	23	0	0	0	0	0

Stellenbosch University https://scholar.sun.ac.za

Locality	Code	Host species	Total host	Hosts with *D. ellobius*	Total *D. ellobius*	Subsampled fleas	D1 (*D. abaris*)	D2 (*D. ellobius*)
continued…		*Mastomys natalensis*	4	2	8	5	0	5
		Crocidura muriquensis	9	2	2	0	0	0
		Otomys angoniensis	3	0	0	0	0	0
			39	4	10	5	0	5
Inkunzi Lodge	IN	*Rhabdomys dilectus*	9	0	0	0	0	0
		Mastomys natalensis	9	4	6	6	0	6
		Aethomys chrysophilus	4	3	20	7	0	7
		Micealamys namaquensis	3	1	1	0	0	0
		Rattus rattus	3	0	0	0	0	0
		Myosorex varius	2	1	3	3	0	3
		Otomys irroratus	1	0	0	0	0	0
		Otomys angoniensis	1	0	0	0	0	0
			32	9	30	16	0	16
Mt Gilboa	MG	*Rhabdomys dilectus*	27	15	44	7	0	7
		Otomys irroratus	17	10	23	7	0	7
		Shrew spp.	6	3	7	1	0	1
			50	28	74	15	0	15
Mt Shannon	MS	*Rhabdomys dilectus*	30	6	10	5	0	5
		Otomys irroratus	12	0	0	0	0	0
		Shrew spp.	3	0	0	0	0	0
			45	6	10	5	0	5
Vernon Crookes	VC	*Rhabdomys dilectus*	16	3	3	3	0	3
		Mastomys natalensis	3	1	3	3	0	3
		Shrew spp.	7	0	0	0	0	0

Stellenbosch University https://scholar.sun.ac.za

Locality	Code	Host species	Total host	Hosts with *D. ellobius*	Total *D. ellobius*	Subsampled fleas	D1 (*D. abaris*)	D2 (*D. ellobius*)
continued…			26	4	6	6	0	6
Vryheid	VH	*Rhabdomys dilectus*	3	0	0	0	0	0
		Micealamys namaquensis	2	1	2	1	1	0
		Mastomys natalensis	4	1	5	3	3	0
		Lemniscomys rosalia	1	0	0	0	0	0
		Crocidura flavescens	1	0	0	0	0	0
		Mus musculus	1	0	0	0	0	0
		Grammomys spp.	1	0	0	0	0	0
		Shrew spp.	1	0	0	0	0	0
			14	2	7	4	4	0
19 localities		**23 species**	**830**	**182**	**521**	**151**	**43**	**108**

* Could not determine which host the samples were sequenced from.

Stellenbosch University https://scholar.sun.ac.za

Table B.8 Collinearity among the 19 final individual predictor variables. All final predictor variables had correlation values below 0.7 or above -0.7.

Predictor variable	DT6	DT7	DTMIN	LAI3	LAI4	NDVI4	NDVI7	RF4	RF6	RF7	RFMIN	RFMAX	WV4	WV5	WV6	WVMIN	SPS	SOC	SPH
DT6	1.00	0.49	-0.36	0.23	0.11	0.45	0.17	0.10	0.07	0.09	0.08	-0.14	-0.02	0.48	-0.25	0.01	-0.40	0.21	-0.13
DT7	0.49	1.00	0.13	0.22	0.42	0.43	0.26	-0.36	-0.31	0.22	-0.07	-0.17	0.32	-0.04	-0.34	0.14	0.00	-0.13	0.00
DTMIN	-0.36	0.13	1.00	0.01	0.23	0.11	0.12	-0.30	-0.24	0.03	-0.24	0.28	0.51	-0.28	0.02	0.15	0.33	-0.52	-0.09
LAI3	0.23	0.22	0.01	1.00	0.42	0.34	0.22	-0.09	-0.20	-0.07	-0.02	0.23	0.22	0.39	-0.39	0.11	-0.25	0.07	-0.24
LAI4	0.11	0.42	0.23	0.42	1.00	0.30	0.51	-0.39	-0.36	0.08	-0.01	0.07	0.38	-0.06	-0.32	0.16	0.02	-0.08	-0.16
NDVI4	0.45	0.43	0.11	0.34	0.30	1.00	0.09	-0.15	-0.19	-0.10	0.07	0.22	0.31	0.17	-0.39	0.39	-0.26	0.15	-0.31
NDVI7	0.17	0.26	0.12	0.22	0.51	0.09	1.00	-0.13	-0.41	0.13	0.00	0.08	0.15	0.13	-0.08	0.04	0.08	-0.04	-0.30
RF4	0.10	-0.36	-0.30	-0.09	-0.39	-0.15	-0.13	1.00	0.36	-0.15	0.02	-0.07	-0.45	0.31	0.26	-0.11	-0.06	0.09	0.03
RF6	0.07	-0.31	-0.24	-0.20	-0.36	-0.19	-0.41	0.36	1.00	-0.14	0.06	-0.24	-0.31	0.20	0.11	-0.30	-0.18	0.16	0.30
RF7	0.09	0.22	0.03	-0.07	0.08	-0.10	0.13	-0.15	-0.14	1.00	-0.05	-0.25	-0.11	-0.24	0.37	-0.19	0.09	-0.26	0.24
RFMIN	0.08	-0.07	-0.24	-0.02	-0.01	0.07	0.00	0.02	0.06	-0.05	1.00	-0.01	-0.16	0.03	0.01	0.15	-0.07	0.32	-0.08
RFMAX	-0.14	-0.17	0.28	0.23	0.07	0.22	0.08	-0.07	-0.24	-0.25	-0.01	1.00	0.24	0.10	-0.13	0.39	-0.20	0.14	-0.58
WV4	-0.02	0.32	0.51	0.22	0.38	0.31	0.15	-0.45	-0.31	-0.11	-0.16	0.24	1.00	-0.08	-0.38	0.16	0.01	-0.18	-0.24
WV5	0.48	-0.04	-0.28	0.39	-0.06	0.17	0.13	0.31	0.20	-0.24	0.03	0.10	-0.08	1.00	-0.37	-0.20	-0.25	0.26	-0.29
WV6	-0.25	-0.34	0.02	-0.39	-0.32	-0.39	-0.08	0.26	0.11	0.37	0.01	-0.13	-0.38	-0.37	1.00	-0.30	0.21	-0.25	0.23
WVMIN	0.01	0.14	0.15	0.11	0.16	0.39	0.04	-0.11	-0.30	-0.19	0.15	0.39	0.16	-0.20	-0.30	1.00	-0.13	0.25	-0.41
SPS	-0.40	0.00	0.33	-0.25	0.02	-0.26	0.08	-0.06	-0.18	0.09	-0.07	-0.20	0.01	-0.25	0.21	-0.13	1.00	-0.47	-0.04
SOC	0.21	-0.13	-0.52	0.07	-0.08	0.15	-0.04	0.09	0.16	-0.26	0.32	0.14	-0.18	0.26	-0.25	0.25	-0.47	1.00	-0.31
SPH	-0.13	0.00	-0.09	-0.24	-0.16	-0.31	-0.30	0.03	0.30	0.24	-0.08	-0.58	-0.24	-0.29	0.23	-0.41	-0.04	-0.31	1.00

Stellenbosch University https://scholar.sun.ac.za

Table B.9 List of predictor variables indicating their respective reference codes.

Predictor variable	**Reference code**
6th harmonic component of daytime land surface temperature	DT6
7th harmonic component of daytime land surface temperature	DT7
Minimum daytime land surface temperature	DTMIN
3rd harmonic component of daily LAI	LAI3
4th harmonic component of daily LAI	LAI4
4th harmonic component of daily NDVI	NDVI4
7th harmonic component of daily NDVI	NDVI7
4th harmonic component of daily rainfall	RF4
6th harmonic component of daily rainfall	RF6
7th harmonic component of daily rainfall	RF7
Minimum rainfall	RFMIN
Maximum rainfall	RFMAX
4th harmonic component of daily water vapour	WV4
5th harmonic component of daily water vapour	WV5
6th harmonic component of daily water vapour	WV6
Minimum water vapour	WVMIN
Soil percentage sand	SPS
Soil organic carbon	SOC
Soil pH	SPH