Testing the utility of DNA barcoding for the rapid assessment of Formicidae biodiversity in the eThekwini region

by

Sohana Singh

BSc. (*Hons*) Genetics

Submitted in fulfilment of the academic requirements for the degree of Master of Science in

The School of Life Sciences

University of KwaZulu-Natal

Pietermaritzburg

As the candidate's supervisor I have approved this dissertation for submission.

Signed: _____ Name: Dr S. Willows-Munro Date: _____

Preface

The experimental work described in this dissertation was carried out in the Discipline of Genetics, School of Life Sciences, University of KwaZulu-Natal, Pietermaritzburg, from May 2011 to April 2013 under the supervision of Dr S. Willows-Munro.

These studies represent original work by the author and have not otherwise been submitted in any form to another University. Where use has been made of the work by other authors it has been duly acknowledged in the text.

Name:

_____

Signed:

_____

Date:

_____

College of Science and Agriculture

Declaration of Plagiarism

I, Sohana Singh declare that:

1. The research reported in this thesis, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   a. Their words have been re-written but the general information attributed to them has been referenced
   b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed:

_____

Date:

_____

Declaration Plagiarism 22/05/08 FHDR Approved

Abstract

Testing the utility of DNA barcoding for the rapid assessment of Formicidae biodiversity in the eThekwini region.

Student: Sohana Singh

Supervisor: Dr S. Willows-Munro

The biodiversity of Durban (eThekwini municipality) in KwaZulu Natal is primarily threatened by urbanization although other factors such as climate change and the spread of invasive species also pose a significant threat. Knowledge of what species exist within the city is important for biodiversity surveillance, detecting invasive taxa and uncovering cryptic species. Conducting a comprehensive biodiversity inventory is a daunting task, especially for hyperdiverse groups such as terrestrial arthropods, where closely related species can often only be separated by subtle morphological characters. This study investigated whether the barcoding marker, Cytochrome Oxidase C Subunit 1 (COI) can be used to efficiently and accurately delineate species of ants (family Formicidae) in comparison to traditional taxonomic approaches. The feasibility of DNA barcoding for assembling biodiversity inventories for urban areas which could be useful in conservation planning was also evaluated. A total of 619 individuals were sequenced from 23 geographic localities within the eThekwini region and surrounding regions. DNA barcoding revealed 80 provisional species/ "barcode clusters" or monophyletic lineages which could represent distinct species, while morphology revealed 51 different morphospecies. Extrapolation measures of species richness indicated that as many as 153 species of ants could occur in the city. Phylogenetic and phylogeographic analyses were performed on co-distributed species belonging to the genera *Lepisiota*, *Camponotus*, *Pheidole* and *Pachycondyla* to better understand the spatial distribution of genetic variability in the eThekwini region. Nuclear markers 18S rDNA and 28S rDNA were also sequenced and compared for a subsample of individuals from *Camponotus* and *Pachycondyla*. There was genetic variation at COI and the nuclear markers for each of the species examined. In order to fully elucidate the population genetic patterns which could be expected in eThekwini and surrounding regions, further sampling across more localities is essential. The use of more nuclear markers could also assist in uncovering these unique patterns of genetic variation in an urban setting. In this study, the utility of COI as a species diagnostic tool in ants was confirmed. The barcoding library constructed showed promise in highlighting reserves that should be preserved and possible cryptic speciation for further investigation.

Acknowledgements

This dissertation is dedicated in memory of my late grandfather, Mr. J.L. Singh, who always encouraged me to further my education and to strive for the best.

A heartfelt thanks goes out to my dad, P.J. Singh, and grandmother Mrs L. Singh, who sacrificed a lot while providing unconditional love and support so that I could study towards my dreams. I would also like to thank my siblings, Pooja and Nikhil Singh who also showed love, support and encouragement.

I would like to thank my supervisor, Dr. Sandi Willows-Munro for her guidance, motivation and inspiration and for instilling in me the ethics of hard work. Thank you for sharing your knowledge and passion with me and I really appreciate the quick feedback, the ideas and encouragement.

I would like to acknowledge the National Research Foundation (NRF) and the University of KwaZulu-Natal for providing the funding for this project. I would also like to thank the University of Guelph in Canada for assistance with this project.

Few endeavours would be bearable without the support of good friends. Thanks goes out to my housemates, Sandipa Bhikraj, Ashnie Moolchand, Serisheena Nunnan, Kireshnie Naicker, Reansha Bissin, Kerisha Ahir and Dane Seunanden, for the late night chats, laughter and help with cooking. Thanks to my awesome friends for keeping me sane, for always being there for the amazing fun times we've had and for the down times… Laurencia Harris, Chavonne Chetty, Bilaal Ismail, Kerwin Moonsamy. All of you guys were basically my family away from home.

Life at campus would have been boring without these epic ladies – Letrisha Padayachee, Mary Ranketse, Melissa (Milly) Govender, Lefentse Mashamaite, Mampolelo Photolo. Thanks for the fun, crazy and sometimes enlightening chats at 'lunch time', and for supporting me during the conference in Cape Town where we had tons of fun and adventure, and for supporting me at the campus research day.

Thank you to the past and present members of the Conservation Genetics Lab – Ayanda Shelembe, Connor Stobie, Thina Duze, Qiniso and Sihle Mthetwa, Joro, Riel Coetzee and Tara Vorster, for making the lab a fun environment to be in and for thoughts and help that were needed.

Collection of data for this project would not have been easy without the help of the SANBI interns, Zuki, Kirby, Matabara, Patrick, Lorika and Vince, and their assistance with field work and sorting of the insects. I would also like to thank Pooja, Ashnie, Kerisha, Letrisha, Mary, Ayanda, Shravan Ramdial, Leeth Singh and Kapil Ramisur for assistance with field work.

Finally, I want to express my gratitude for the support and assistance of the staff of the Department of Genetics at the University of KwaZulu-Natal – Dr. Pillay, Mr. Dzomba, Dr. Meusel, Megan Brunkhorst and Goodman Zondi.

Table of Contents

List of tables

List of figures

Chapter One

DNA Barcoding of Invertebrates in Africa and its Prospects for Landscape Genetics and Phylogeography – A Case Study Using the Ants of Madagascar

Abstract

Maintaining biodiversity is essential for ecosystems to retain their resilience to the pressures imposed on them by mankind and natural disasters. Biodiversity is threatened by habitat degradation, land-use change, pollution, climate change and loss of species and genetic diversity. In order to protect and preserve biodiversity, there must be prior knowledge of what biodiversity there is to begin with. With the rise of molecular technologies, it has now become possible to assess and quantify biodiversity faster and more efficiently. DNA barcoding is a diagnostic technique which uses a short, standardized segment of the Cytochrome C Oxidase I (COI) mitochondrial gene to identify animal species. Despite criticisms, it has been applied successfully to many studies involving arthropods. In this introductory chapter, I will provide a summary of DNA barcoding on the African continent. Traditionally, DNA barcode data is used above the species level in order to delimit species but in this chapter I tested if the sequence data could be informative below the species level. Using data from the ants of Madagascar DNA barcode project as a case study, the COI data was used in landscape and phylogeographic analyses to determine the spatial distribution of genetic variation of *Anochetus madagascarensis* across Madagascar. A total of 120 COI sequences of *A. madagascarensis* from 39 localities within Madagascar and a neighbouring island were downloaded from the Barcode of Life Database (BOLD). Landscape genetics and phylogeographical analyses revealed that there were three distinct populations; a group from Northern Madagascar, a southern group and an offshore group collected on the neighbouring island of Mayotte.

1.1 Introduction

1.1.1 Protecting our biodiversity

Biodiversity describes the wide range of the variety of living organisms on Earth, the genetic variation between them and the different ecosystems, communities and habitats in which organisms live. With the global biodiversity crisis upon us, it has now become extremely important to catalogue species in order to determine the geographic areas that contain the most biological diversity and which would qualify for elevated levels of protection and can be classified as 'biodiversity hotspots'. Biodiversity hotspots may be defined as species rich areas which usually harbour rare or threatened species and a large number of endemic species found in small areas which are characterized by a high degree of habitat loss or habitat fragmentation (Reid, 1998). Biodiversity hotspots are fundamental to conservation planning. Demographic analyses using estimates from population density and growth in human populations within biodiversity hotspots indicate that approximately 20% of the world's population are settled within biodiversity hotspots (Cincotta et al., 2000). This puts pressure on natural resources and in turn is a major threat to global biodiversity (Cox and Moore, 2010).

Africa has a wealth of biodiversity. To date Africa is home to 1,229 of the world's 4,700 mammal species, more than 2,000 species of birds and approximately 950 species of amphibians (McGinley and Hogan, 2011). Africa also contains regions of exceptionally high levels of endemism for example, the Cape Floral Kingdom in Southern Africa contains over 9000 vascular plant species of which approximately 69% are endemic (McGinley and Hogan, 2011). Africa's coastline also boasts a high marine biodiversity (McGinley and Hogan, 2011). The island of Madagascar is well known for its high levels of endemicity, there are over 12,000 plant species, 81% of which are native (Myers, 1988).

Biodiversity in Africa is still in a relatively good condition when compared to the rest of the world. There are over 2 million km^2 of protected areas. Despite this, the biodiversity is still vulnerable with about half of the terrestrial eco-regions in Africa losing approximately 50% of their area to anthropogenic factors such as urbanization, habitat degradation and cultivation (McGinley and Hogan, 2011). Perhaps one of the most serious threats to

biodiversity loss is habitat destruction and conversion of natural habitats to other uses resulting in the loss of open space (Croucamp, 2009; Heywood, 1995). Protecting our biodiversity is essential because an ecosystem that has a rich biodiversity will be more resilient to natural events, such as fires and storms, and human-induced environmental changes (Croucamp, 2009).

Unfortunately, the little biodiversity data available for African ecosystems is biased towards vertebrates, despite most biodiversity involving arthropods and other invertebrate taxa. The arthropods of Africa are rich and diverse and there are at least 580 families and approximately 100,000 species that are currently described (McGinley and Hogan, 2011). However, biodiversity information is lacking and patchy for many organisms, especially invertebrates which potentially constitute about 95% of all species on the continent (Myers et al., 2000). The description of biodiversity in Africa is hampered by the lack of taxonomic expertise and a lack of funding. Birds, plants and mammals are often used as indicator groups for biodiversity assessments because they are well-studied (Reid, 1998) . However, the use of such indicator groups may not be reliable as studies have indicated that species richness patterns and endemism often do not correspond across all taxa when fine scale geographic resolutions are considered (Reid, 1998). This is due to species having different habitat requirements and life history traits. Accurate biodiversity assessments and inventories are critical to conservation planning. Traditional taxonomy has long since been the backbone of species identifications, classification, biodiversity surveys and numerous other biological studies. The need for new technology and methods of collecting data for biodiversity inventories has become more important in the face of elevated levels of species extinction. This is worrying for conservationists because many species could already be extinct even before being described, and due to this, effective protection measures cannot be put into place (Swartz et al., 2008).

1.1.2 DNA barcoding

In 2003, Hebert and colleagues from the University of Guelph in Canada, coined the term ‘DNA barcoding’ after they used a 648 bp region of the mitochondrial cytochrome oxidase subunit 1 gene (COI) to identify animals from seven phyla and eight of the largest insect

orders (Hebert et al., 2003a). The main principle underpinning DNA barcoding is to identify an organism by comparing a standardized segment of a gene from the organism of interest to the same segment from a library of related organisms (Dasmahapatra and Mallet, 2006; Hajibabaei et al., 2006; Hajibabaei et al., 2007; Hebert et al., 2004; Lahaye et al., 2008; Valentini et al., 2008b). DNA barcoding rests on the assumption that genetic variation amongst species is greater than genetic variation within a species. By using a short, standardized sequence and a clustering algorithm, DNA barcodes have been suggested to correctly cluster together individuals belonging to the same species and to possibly discover new species (Hajibabaei et al., 2007; Moritz and Cicero, 2004).

The Consortium for the Barcode of Life (CBOL) was formed in May 2004 (Hajibabaei et al., 2007; Ratnasingham and Hebert, 2007). The CBOL is a group of international research organisations supporting the expansion of DNA barcoding as the international standard for identifying species (Hajibabaei et al., 2007). The 'Barcode of Life Initiative' is a project that was launched by CBOL to promote barcoding and identify the 10 million estimated species on Earth in a way that is fast and relatively cheap (Dasmahapatra and Mallet, 2006; Savolainen et al., 2005). An online database called the Barcode of Life Data System (BOLD; www.barcodinglife.org) was created to acquire, store, analyse and manage DNA barcode records (Hajibabaei *et al*., 2007; Ratnasingham and Hebert, 2007). All researchers doing DNA barcoding upload their data to the BOLD database. The BOLD database is an invaluable resource to biodiversity researchers and allows African scientists access to large amounts of data that was not available previously.

During the International Year of Biodiversity in 2010, the Barcode of Life launched the International Barcode of Life Project (iBOL; http://www.ibolproject.org), to coordinate global efforts and bring together 26 countries to collaborate. This collaborative effort was undertaken to broaden and strengthen DNA barcoding research by extending the focus towards developing countries in which a vast majority of the Earth's biodiversity is located (Vernooy et al., 2010). The partnering nations of iBOL include the central nodes - Canada, China, France, Finland, Germany, Netherlands, Portugal, United Kingdom, United States, and the regional nodes - Argentina, Australia, Brazil, India, Mexico, New Zealand, Norway, Russia, Saudi Arabia and South Africa. There are a number of global barcode campaigns

underway with goals to barcode every organism on Earth, from bacteria to mammals (Table 1.1). As of August 2013, the number of formally described species with barcodes were 138, 870 animals, 51,341 plants and 3,443 fungi and other organisms. The total number of barcode records were 2,651,524 (www.barcodinglife.org). Much of the progress made to date has been for the order Lepidoptera and marine life but many of the other global initiatives are also gaining momentum.

Table 1.1 Progress (in terms of number of species and specimens barcoded) made by on-going global DNA barcoding campaigns accessed in August 2013 from links provided on the iBOL website (www.ibol.org).

Campaign	**Specimens barcoded**	**Species**
Formicidae Barcode of Life	15409	1298
All Birds Barcoding Initiative (ABBI)	26573	4019
Trichoptera Barcode of Life	18549	2636
FISHBOL	88755	9769
Lepidoptera Barcode of Life	793964	669991
Mammal Barcode of Life	19862	858
Marine Barcode of Life	37182	6199
Polar Barcode of Life	30871	23953
All Fungi Barcoding	*	*
Coral Reef Barcode of Life	*	*
HealthBOL	*	*
Mosquito Barcoding Initiative	*	*
Shark Barcode of Life	*	*
Sponge Barcoding Project	*	*
BeeBOL	*	*

* No data available at the time of writing. Links to these campaigns are still under construction.

1.1.2.2 DNA barcoding methodology, advantages and criticisms

The methodology behind DNA barcoding is straightforward; DNA is retrieved from a tissue sample taken from the organism under study and universal polymerase chain reaction (PCR) primers are used to amplify a 658 bp fragment of the COI mitochondrial gene (used in most vertebrate and invertebrate animal groups). The COI gene was the ideal candidate for use in DNA barcoding because of its highly constrained amino acid sequence which in turn ensures that primers will be broadly applicable and results in straightforward sequence alignment (Moritz and Cicero, 2004; Packer et al., 2009). The COI sequence is variable enough to differentiate between species but is less variable in individuals belonging to the same species which means that individuals belonging to the same species will cluster closely together in a phylogeny (Stoeckle and Hebert, 2008; Valentini et al., 2008a).

A suit of primers have been designed to amplify the COI gene across a wide range of insects (Folmer et al., 1994; Lunt et al., 1996). Folmer *et al*. (1994) designed universal COI primers that are able to amplify across a diverse range of metazoan invertebrates. An important part of DNA barcoding is the construction of a DNA reference library for a particular group. The reference library usually includes sequences taken from specimens which have been identified by taxonomic experts (Hajibabaei et al., 2007). The matching of sequences from related organisms is done using a phylogenetic tree approach (Dasmahapatra and Mallet, 2006). The most widely used method in many DNA barcoding studies is the use of the Kimura 2-parameter (K2P) model to calculate genetic distances, these distances are then used to build a neighbor-joining tree. The query sequences are assigned the species name of the closest matching group it clusters with (Frezal and Leblois, 2008). These barcode clusters are known as Barcode Index Numbers (BINs).

If there are no matches within the barcode library, this could mean that the barcode sequence could represent a new species or a geographical variant of the species (Hajibabaei et al., 2007). Successful species identification relies on the “gap” that exists when interspecific genetic variation exceeds intraspecific genetic variation – this is termed the “barcoding gap”.

In early DNA barcoding studies, the standard threshold for delimiting mammals was ≤ 2% and for invertebrates it was ≤ 3% (Hebert et al., 2003a; Hebert et al., 2003b) The BOLD analytical workbench (www.barcodinglife.org, Ratnasingham and Hebert, 2007) offers a variety of functions to analyse COI data, including generation of a K2P neighbour-joining tree, visualization of the barcode gap, diagnostic character analysis and a clustering algorithm which assigns sequences to Operational Taxonomic Units (OTUs).

Even though the benefits of DNA barcoding are numerous and its role in biodiversity research is now well entrenched in the literature, there are those who contest and oppose it as a tool for species identification and discovery. One of the major criticisms is that DNA barcoding sometimes detects new species that are not real (false positives) or it doesn't detect differences in species that can be discriminated by other methods such as morphology (false negatives). This can be due to using only a single gene identification system and because the COI gene does not have the same mutation rate in all organisms (Hickerson *et al*., 2006). The rate heterogeneity can introduce bias towards the discovery of recent species and increase the rate of false negatives in some taxa that have faster mutation rates (Hickerson et al., 2006). The mitochondrial genome is maternally inherited, as such; its patterns of genetic relationships in species could differ from nuclear DNA patterns and from the "true" species tree. An example of this was shown in the DNA barcoding study on the ithomiine butterfly genus *Mechanitis*; the barcode results suggested that there were four new species but the nuclear AFLP data only supported one of these four new species (Dasmahapatra et al., 2010).

The recognition of species boundaries by mitochondrial and nuclear DNA may differ for the same set of species due to hybridization and incomplete lineage sorting (Collins and Cruickshank, 2012; Frezal and Leblois, 2008; Rubinoff, 2006). Identification using mitochondrial DNA can also be compromised by nuclear mitochondrial DNAs (NUMTs), which are the nuclear copies of mitochondrial DNA genes that are present in the nuclear genome (Frezal and Leblois, 2008).

Researchers have questioned whether the DNA barcoding gap exists or whether it is an actually an artefact of incomplete sampling across taxa, as was demonstrated in a study on

blue butterflies (Wiemers and Fiedler, 2007). Overlap between interspecific and intraspecific genetic distances could be due to a species having elevated levels of genetic diversity below the species level. High rates of intraspecific divergence can be attributed to populations that are geographically isolated, therefore geographic, intra-specific sampling is important (Bergsten et al., 2012; Frezal and Leblois, 2008; Hickerson et al., 2006). Another limiting factor to consider is paraphyly or polyphyly because of incomplete lineage sorting of mitochondrial DNA and introgression. The coverage (multiple individuals of a species and adequate geographical sampling) and reliability of DNA barcode libraries is a major factor on which the accuracy of species identifications is dependent (Bergsten et al., 2012; Ekrem et al., 2007; Jinbo et al., 2011). The lack of adequate reference data could sometimes be responsible for false positive and false negative identifications and underestimation of intraspecific genetic variation. Incongruence between the DNA barcode data and the traditional morphological definition of species, especially in poorly studied groups, can also result in overlap. In some instances, the intraspecific distance for one species could be higher than the interspecific distance of a different species without affecting the identification success (Collins and Cruickshank, 2012). It was suggested by Collins and Cruickshank (2012) that a more feasible representation of the barcode gap would be to plot the distance to the furthest conspecific against the distance to the nearest nonconspecific and the point where the difference between the two is zero indicates that there is no barcode gap. A study by Meier *et al*. (2006) suggested that when interpreting the barcode gap, the smallest interspecific distance should be considered rather than the mean interspecific distance.

Another argument is that DNA barcoding is intellectually poor and simplistic when compared to the rich legacy of traditional taxonomy using morphological approaches (Jinbo et al., 2011; Packer et al., 2009; Rubinoff, 2006). However, studies have shown that DNA barcoding outperforms traditional taxonomy when it comes to identifying specimens at different life stages (Davis et al., 2011; Greenstone et al., 2005; Kubisz et al., 2012; Schilthuizen et al., 2011), and species complexes that are hard to identify, for example the nine closely related butterfly species from Romania whose species statuses were resolved using DNA barcoding when external morphology is not informative (Dinca et al., 2010). DNA barcoding is also more useful than traditional taxonomy when degraded or very old samples such as museum samples are involved as shown by Hajibabaei *et al*. (2006) when they used a 100 bp fragment of COI to successfully identify over 90% of museum specimens of moths and wasps.

Critics have alluded that all DNA barcoding provides is information instead of knowledge and lacks a broader context (Fitzhugh, 2006; Rubinoff, 2006). But it can be argued that DNA barcoding in itself tests the hypothesis of its utility to identify species and can generate hypotheses about species that can be tested using other techniques (Waugh, 2007). In addition, DNA barcoding offers independent data that is free from personal bias that is often associated with identifications made by traditional taxonomy (Packer et al., 2009).

An issue that has been raised by numerous DNA barcode critics is the quality control of sequences that are submitted to DNA barcode libraries as misidentified sequences could be problematic. But for a DNA barcode to be valid, it should have taxonomic information, voucher specimen data such as photographs, locality information (GPS co-ordinates, country, state/province), date of collection, where the specimen has been deposited and the PCR primers and trace files (Ratnasingham and Hebert, 2007). This information makes it possible for taxonomic experts to identify the voucher specimen or to validate the identification at a later stage (Jinbo et al., 2011).

There has also been dispute over how the DNA barcode data is analysed. There has been much debate over whether the neighbour-joining tree using the K2P distance method is accurate and appropriate for delimiting species, particularly in cases where there is incomplete lineage sorting and paraphyly at the species level (Collins and Cruickshank, 2012; Jinbo *et al*., 2011). Incorrect model specification can bias genetic distances and identification results could be misleading, ambiguous or incorrect (Collins and Cruickshank, 2012; Jinbo et al., 2011).

Due to these shortcomings, new criteria and algorithms are being developed and applied as an alternative or a compliment to tree-based identifications. Distance-based criteria such as ‘Best Match’ and ‘Best Close Match’ were developed by Meier *et al* (2006). The ‘Best Match’ criteria assigns the query the species name of the barcode that best matches it, without considering the degree of similarity between the query and barcode sequence. However, this method is prone to misidentifications when the query sequence does not have a representative in the database. These misidentifications can be avoided by the ‘Best Close Match’ criteria

which employs the same principal as ‘Best Match’, but only assigns the species name if there is a high degree of similarity, otherwise the query sequence will be classified as unidentified and will require additional taxonomic intervention (Meier et al., 2006).

The Nearest Neighbour criteria is equivalent to the ‘Best Match’ criteria and has also been used for identification purposes in DNA barcode studies (van Velzen et al., 2012). Character-based identification criteria is considered to be more accurate as nucleotide variation in each base position is used as a diagnostic character (Collins and Cruickshank, 2012; Jinbo et al., 2011; van Velzen et al., 2012). This diagnostic method is implemented in programs such as DNA-BAR (Dasgupta, 2005) and BLOG (Weitschek et al., 2013).

Statistical methods such as Bayesian Inference or maximum likelihood (with appropriate models of evolution) can be used to estimate measures of confidence for species identification with DNA barcodes. Implementation of these methods in DNA barcoding studies is crucial as they incorporate explicit population genetic or phylogenetic models which fit the data better than the K2P model which is currently implemented (Jinbo et al., 2011; van Velzen et al., 2012). It is important to stress that no known method is without problems and limitations. Instead of discrediting techniques and methods, efforts should be focussed towards adapting and integrating different methods in order to achieve more reliable, and meaningful results.

1.1.3 Importance of applying DNA barcoding to arthropods

More than 90% of species on our planet are arthropods. Assembling an inventory for arthropod biodiversity and understanding the effects of fragmentation, climate change and urbanization on arthropod biodiversity is essential because they have important roles in the ecosystem as biological indicators due to their short life cycles and sensitivity to habitat disturbances (Bolger et al., 2000).

There have been a great many DNA barcoding studies done on arthropods to date. Apart from simply using DNA barcoding as a species diagnostic tool, its utility has been demonstrated in many other areas of study. Studies have used DNA barcoding to carry out biodiversity assessments using ants (Smith and Fisher, 2009; Smith et al., 2005), moths (deWaard et al., 2011; Janzen et al., 2005),and flies (Webb et al., 2012; Zhou et al., 2009; Zhou et al., 2011). DNA barcoding has been applied to studies focussed on arthropods of economic importance, investigating host specificity and vector-host relationships and identifying candidates for biocontrol agents. For example, DNA barcoding was used to identify bloodmeal from the tick *Ixodes scapularis*, in order to understand vector-host relationships and identify hosts (Gariepy et al., 2012). It was also used to investigate host specificity in a group of economically important biocontrol agents, the *Anicetus* wasps, which are employed to control agricultural pests such as *Ceroplastes spp*. (Zhang et al., 2011).

Many arthropods are difficult to identify morphologically at different developmental stages, and in these cases DNA has proven to be an effective taxonomic tool. It has been used to distinguish between four *Crioceris* leaf beetles, two of which were economically important invasive plant pests and two which are rare species and hard to distinguish them from each other at the egg and larval stages (Kubisz et al., 2012). DNA barcoding was successful in identifying eggs and larvae of closely related carabids; and spiders which are predators that are important for the biological control of agricultural pests (Greenstone et al., 2005). Distinguishing *Laricobius rubidus*, endemic to western North America, from *Laricobius nigrinus*, an introduced species used as a biocontrol agent, is another example of a study applying DNA barcoding to identify species which have morphologically indistinguishable larvae (Davis et al., 2011).

Data from DNA barcoding has also been used in ecological studies such as a study involving the bees of Canada in which sex-associations between cleptoparasitic species were revealed (Sheffield et al., 2009). In another study it was shown that a group of necrobiont beetles (often found on vertebrate cadavers) from the family Cholevidae can be used in forensic entomology, even though they previously weren't considered useful and due to their small size and cryptic adult and larval stages (Schilthuizen et al., 2011).

Given their small size some arthropods are not able to move great distances and often exhibit fine-scale patterns of diversity. The DNA barcoding marker COI has also been used in population-level studies, for example, Hebert *et al*. (2004) used DNA barcoding to uncover cryptic species of skipper butterfly (*Astraptes fulgerator*) and also found that the COI gene was useful in uncovering genetic structuring below the species level.

1.1.4 Below the species level – landscape genetics

Landscape genetics is a relatively new research area which incorporates other disciplines, namely; landscape ecology, population genetics and spatial statistics (Storfer et al., 2007). The goal of landscape genetics is to infer how landscape features influence selection, gene flow and genetic drift (Holderegger et al., 2006; Holderegger and Wagner, 2006). The term 'landscape genetics' was first used in a paper by Manel *et al*. (2003) where they stated that landscape genetics will enable understanding of how genetic variation is structured by environmental and geographical characteristics.

Landscape genetics is particularly useful for discovering boundaries to gene flow that are not obvious – these cryptic boundaries could be discontinuities in the gene flow across a population without any apparent cause i.e. there is no clear physical barrier to gene flow. By examining the history of these cryptic boundaries, the sequence data can be used to infer if secondary contact is being made among populations that were previously isolated (Manel et al., 2003). There are two important steps involved in landscape genetics; the identification of genetic discontinuities in a population and the subsequent correlation of the genetic discontinuities with landscape and environmental variables such as physical barriers. These barriers could be in the form of mountains, catchments, motorways, forests, valleys and urban areas (Holderegger and Wagner, 2008; Manel et al., 2003).

Landscape genetics can explore how the fragmentation of habitats, loss of habitat and spatial isolation effects species dynamics across landscapes, and hence the constraints imposed on the distribution of plants and animals as well as their genes (Holderegger and Wagner, 2008). Understanding these processes is important for the management of the genetic diversity of

species and, at a broader scale, ecosystems (Manel et al., 2003). In this chapter I provide a review of the current status of DNA barcoding in Africa. In addition, data from one of the major African Formicidae barcoding projects to date, was considered as a case study for testing the utility of the barcoding marker COI, below the species level.

1.2 Materials and methods

1.2.1 DNA barcoding in Africa

In order to better understand the scope of DNA barcoding within the African context, an online survey was conducted to ascertain the amount of sequence data available for invertebrates on BOLD. The focus was on COI sequence data available for specimens collected in the top 10 barcoding African countries. These include Madagascar, South Africa, Kenya, Tanzania, Gabon, Democratic Republic of Congo, Cameroon, Ghana, Ethiopia and Zambia. In particular, I examined the total number of arthropod taxa sequenced as well as the number of insects, Hymenoptera and Formicidae COI data available.

1.2.2 Selection of data for landscape genetics analyses

Madagascar is the ideal candidate for biodiversity surveys and studies of speciation and landscape genetics due its unique biota. It is also a biodiversity hotspot which is threatened by habitat degradation and destruction. Approximately 80% of the original vegetation in Madagascar has been destroyed and now consists of secondary grassland which does not harbour many species. Among the threatened are terrestrial arthropods. There are around 1000 species of ants in Madagascar but as much as 75% are still to be described (Smith et al., 2005).

Madagascar has launched conservation planning programs but the difficulty of planning such programs is identifying areas that qualify for protection. Compiling species inventories has proven to be beneficial to conservation planning (Smith et al., 2005). Two of the biggest

African barcoding studies on ants to date were from Madagascar and Mauritius. Both of these studies used ants as a means to perform rapid biodiversity assessments.

In the DNA barcoding study of the ants of Madagascar, subtle genetic structuring at the population level was noted for *Anochetus madagascarensis* (Smith et al., 2005). Individuals from localities separated by the highest mountain in Madagascar exhibited an average sequence divergence of 1.5%. It was hypothesized that there may be cryptic speciation due to the development of mountain ranges which can isolate populations, resulting in unique lineages and species in these region. Climatic shifts in Madagascar have also had an effect on vegetation structure in forests and thus, ant habitats (Smith et al., 2005).

These hypotheses in Smith *et al*. (2005) have not been tested using a phylogeographical or landscape genetics approach. In this chapter I will use as a case study, the COI sequence data for *A. madagascarensis*. Sequence data for this species was downloaded from the BOLD database and aligned using the BOLD sequence alignment option (Amino Acid-based HMM). BioEdit (Hall, 1999) was used to manually optimize the sequence alignments.

Sequence's containing too many gaps, were too short or contained too much missing data (many N's) were excluded from the alignments. A total of 118 sequences from 39 locations in Madagascar and three locations in the nearby island of Mayotte were used for further analysis. A GIS-referenced distribution map was constructed for *A. madagascarensis* with DIVA 7.5 (Hijmans et al., 2012) using GPS co-ordinates available on BOLD for each specimen collected (Figure 1.1).

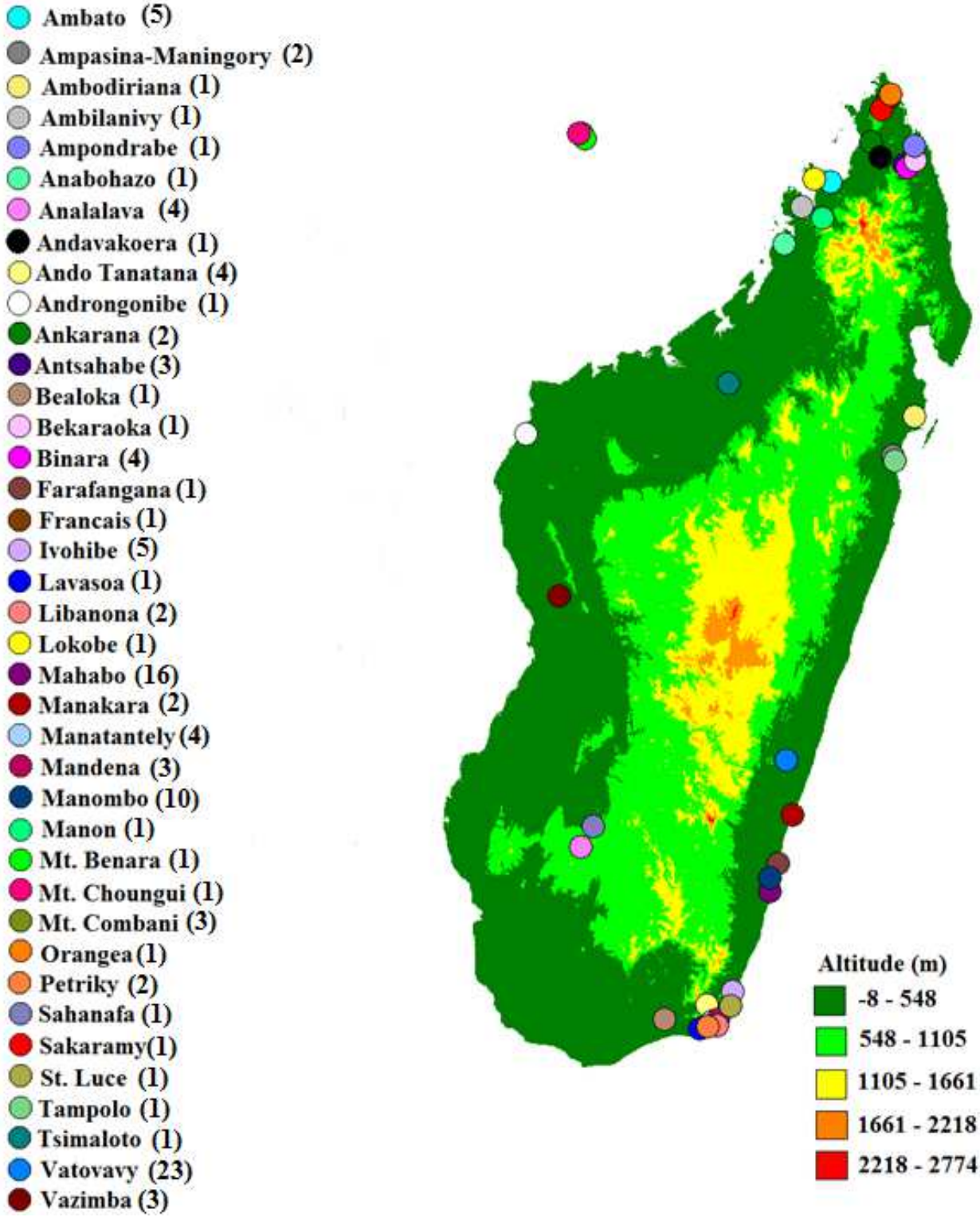


Figure 1.1 A map of Madagascar with the altitude layer and the number of *A. madagascarensis* individuals in each of the 39 sampling locations.

1.2.3 Diversity indices and population differentiation

Summary sequence statistics such as nucleotide diversity, haplotypic diversity and their standard deviations, the number of haplotypes, variable and parsimony informative characters were calculated for the COI sequence data using Arlequin 3.1 (Excoffier et al., 2005) and MEGA 5 (Tamura et al., 2011).

The finer scale population genetic structure of *A. madagascarensis* was also examined by employing a Bayesian clustering approach. The program BAPS 5.2 (Bayesian analysis of population structure; (Corander and Marttinen, 2006; Corander et al., 2003) was used. An analysis was run for each species with the number of clusters, K, set from 1 to 20 with five replicates of each K and two independent runs.

A median-joining network was constructed for *A. madagascarensis* haplotypes using the program Network 4.6.1.1, available from http://fluxus-engineering.com). These networks were used to explore the reticulate relationships among haplotypes and for a qualitative assessment of demographic history. Recent population expansions are characterized by star-burst pattern networks whereas structured networks could be indicative of populations that are stable (Walton et al., 2000).

In order to assess the significance of population-level genetic structure, an analysis of molecular variance (AMOVA), using Arlequin 3 was performed. To determine the scale at which genetic variation occurred, AMOVA was analysed on three levels; among groups (Φ_{CT}), among populations within groups (Φ_{SC}) and within populations (Φ_{ST}). The populations were grouped as follows; a North group of individuals including 30 individuals from 19 localities in Madagascar, a South group of 83 individuals that were sampled from 17 localities in the South of Madagascar, and an offshore group which consists of the five individuals that were sampled from Mayotte (Figure 1.1). The effect of geographical distance on the genetic divergence of populations was assessed using a Mantel test implemented in the Alleles in Space (AIS) software (Miller, 2005).

To visualize how patterns of genetic differentiation change over geographical distance, the AIS software was used to compute raw genetic and geographical distances between points. For this analysis, GPS co-ordinates of all the sampling localities had to be converted to Universal Transverse Mercator (UTM) co-ordinates. This was done using the online application available from http://www.whoi.edu/marine/ndsf/cgi bin/NDSFutility.cgi?form=0&from=LatLon&to=UTM. A connectivity network of the UTM co-ordinates of the sampling localities were created. Next, interindividual genetic distances were assigned to landscape co-ordinates at midpoints of the connectivity network edges. An inverse distance-weighted interpolation procedure was used to measure genetic distances at locations on a grid which overlays the whole sampled landscape. This produced genetic landscape shape interpolation plots which allows for the identification of sampling localities which are the most differentiated (genetically different). Interpolation parameters for the Genetic Landscape Shape surface plots were set at the defaults; number of bins for the X and Y axis = 50 , distance weight value a = 0.5.

1.3 Results and Discussion

1.3.1 Summary data for DNA barcoding in Africa

Out of the 54 African countries only 32 are represented in the global DNA barcoding database BOLD. As of August 2013 Madagascar leads the barcoding initiative in Africa with 28943 arthropod specimens barcoded representing 5445 species clusters (BINs), followed by South Africa with 26156 arthropod specimens barcoded representing 7637 species clusters (BINs) and Kenya with 18624 arthropod specimens barcoded representing 3723 species clusters (BINs) in August 2013 (Figure 1.2). Madagascar also has the highest number of specimens barcoded from the family Formicidae.

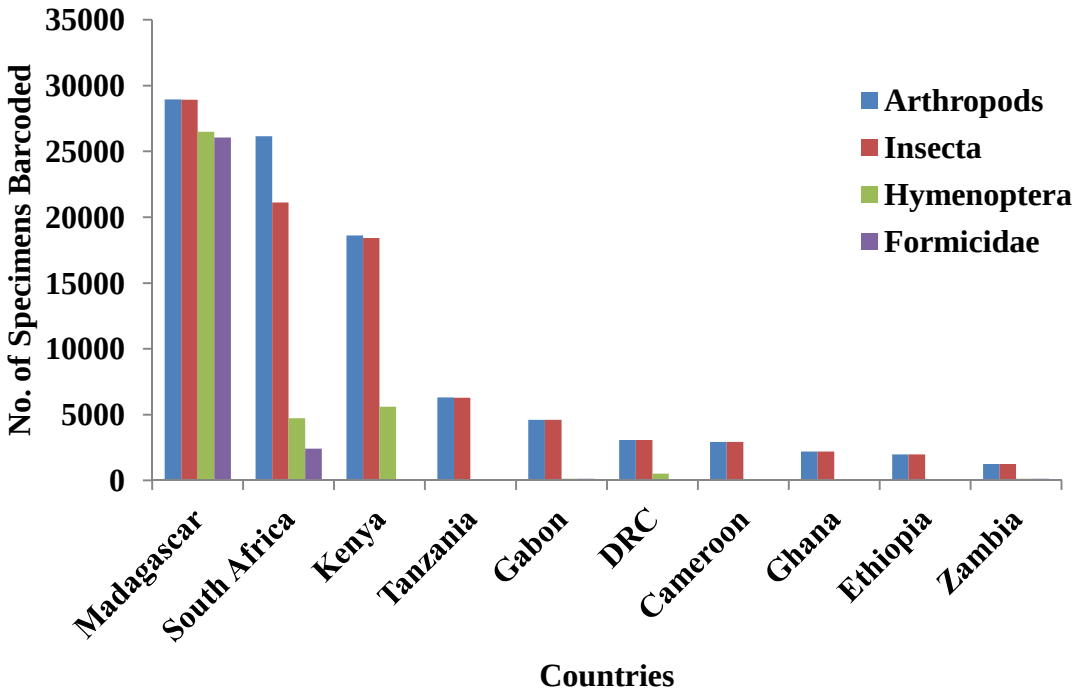


Figure 1.2 The progress of DNA barcoding for invertebrates in Africa (as of August 2013).

1.3.2 Summary sequence statistics

In total, 616 bp of COI from 118 individuals of *A. madagascarensis* were used for the analyses. The final alignment includes 59 variable sites, 39 of which were parsimony informative. The 118 sequences produced 45 unique haplotypes. A total of 41 haplotypes were private, occurring at a single locality and four were shared between localities. There were moderate levels of haplotype diversity ($h = 0.485 \pm 0.053$) and nucleotide diversity ($\pi = 0.00755 \pm 0.00107$) exhibited by COI. The values of Tajima's D -1.23, $P > 0.10$) and Fu's Fs -5.13, $P > 0.10$) were not significantly negative. Negative values would indicate that there has been an expansion in population size (Simonsen et al., 1995). Summary sequence statistics were also computed separately for the North, South and Offshore groups. The South group had the highest haplotype (0.982 ± 0.016) and nucleotide diversity (0.014 ± 0.001), followed by the Offshore group which had a haplotype diversity of 0.700 ± 0.218 and a nucleotide diversity of 0.009 ± 0.003 (Table 1.2). There was low haplotype diversity observed for the

North group (0.286 ± 0.064) and low nucleotide diversity (0.002 ± 0.001). The North and Offshore groups showed a significantly negative value for Tajima's D statistic (-2.26 and -1.49 respectively, Table 1.2) which points to a possible recent expansion in population size in these populations of *A. madagascarensis*.

Table 1.2 Summary sequence statistics for the *A. madagascarensis* dataset as a whole and for each of the groups separately. N is the total number of individuals, L is the number of localities, V is the number of variable characters, PI is the number of parsimony informative characters, Hd is the haplotype diversity, π is the nucleotide diversity and Sd is the standard deviation. Values in bold indicate results that were statistically significant ($P < 0.05$).

	N	**L**	**V**	**PI**	**Hd**	**Sd**	**π**	**Sd**	**Fu's Fs**	**Tajima's D**
A. madagascarensis	118	39	59	39	0.485	0.053	0.008	0.001	-5.13	-1.23
North Group	83	17	26	22	0.286	0.064	0.002	0.001	-0.53	**-2.26**
South Group	30	19	48	22	0.982	0.016	0.014	0.001	-13.45	-1.38
Offshore Group	5	3	8	0	0.700	0.218	0.009	0.003	-0.324	**-1.49**

1.3.3 Patterns revealed by the BAPS and haplotype network

The BAPS analysis revealed three genetically distinct clusters with a log (marginal) likelihood of optimal partition = -1055.20 and a probability of $P = 1.00$. These three clusters suggest that A. madagascarensis is geographically partitioned into three main groups; a northern group, southern group and an offshore group on the neighbouring island of Mayotte. The northern group includes the localities: Vatovavy, Analalava, Mahabo, Farafangana, Manombo, Manakara, Ando Tanatana, Manatantely, Ivohibe, Mandena, Libanona, Lavasoa, Petriky, St. Luce, Bealoka, Sahanafa and Vazimba; the southern group: Antsahabe, Binara, Ambato, Francais, Manon, Sakaramy, Orangea, Ambilanivy, Ankarana, Anabohazo, Lokobe, Andavakoera, Bekaraoka, Ampondrabe, Ambodiriana, Tampolo, Tsimaloto, Androngonibe, Ampasina-Maningory; and the offshore group of specimens collected from the neighbouring island of Mayotte: Mt. Benara, Mt. Combani and Mt. Choungui.

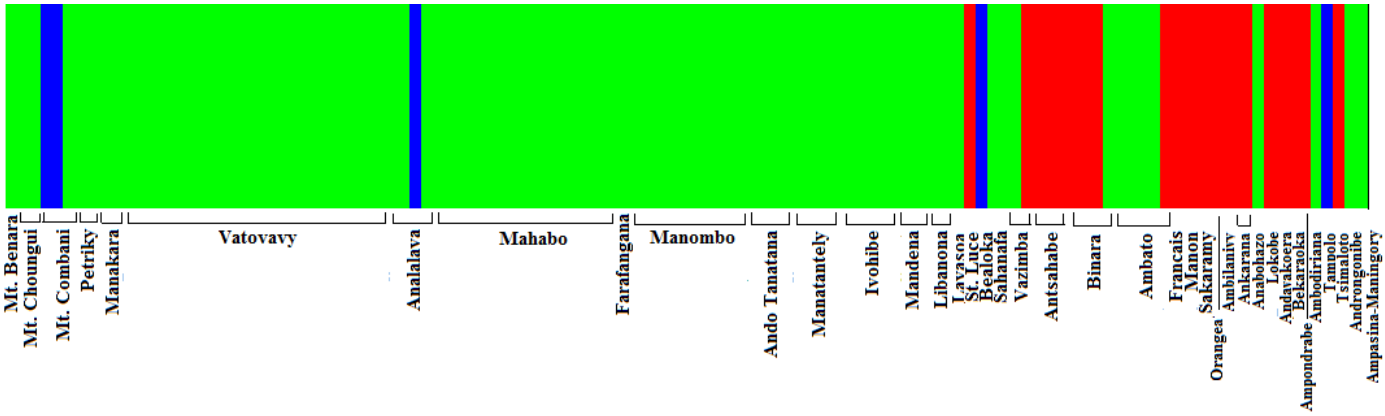


Figure 1.3 The three BAPS clusters which were inferred from the 118 *A. madagascarensis* sequences. The sequences were ordered from Offshore to North to South.

The median-joining haplotype network (Figure 1.4) supported the finding of the BAPS analysis (Figure 1.3), with some exceptions which are discussed below. The northern and southern group are closely associated. More than 50 mutational steps separated the offshore group from the southern group, and four mutational steps separated northern group from the southern group. The most frequently occurring haplotype sequenced from the most individuals and was comprised of individuals mainly sampled from the 17 northern localities in Madagascar, with the exception of the individuals sampled from the southern localities of Anabohazo, Ampasina-Maningory and Tampolo (Figure 1.1). Individuals from northern localities Analalava, Bealoka and Sahanafa, shared haplotypes with individuals from the southern localities. The southern group, consisted of the haplotypes that branched off from the northern group. These haplotypes mainly belonged to the individuals sampled from the 19 southern localities in Madagascar. The offshore group were the individuals from three localities in Mayotte (Figure 1.1). One individual from Mt. Combani in Mayotte shared a haplotype with individuals from the north group, it is unlikely that this is due to natural gene flow, and is probably the result of recent human-mediated introduction from the mainland. Haplotypes shared among the three groups could indicate contemporary gene flow among northern and southern populations or these could be the retention of ancestral haplotypes.

The southern Madagascan ant populations had the most number of unique haplotypes (24 private haplotypes). Populations from the southern group and offshore group seem to be more

genetically differentiated and this could be due to a barrier to gene flow or isolation by distance (Manel et al., 2003). The mean number of mutational steps separating haplotypes in the southern group was 2.4 and the mean number of mutational steps separating haplotypes in the northern group was 1.25.

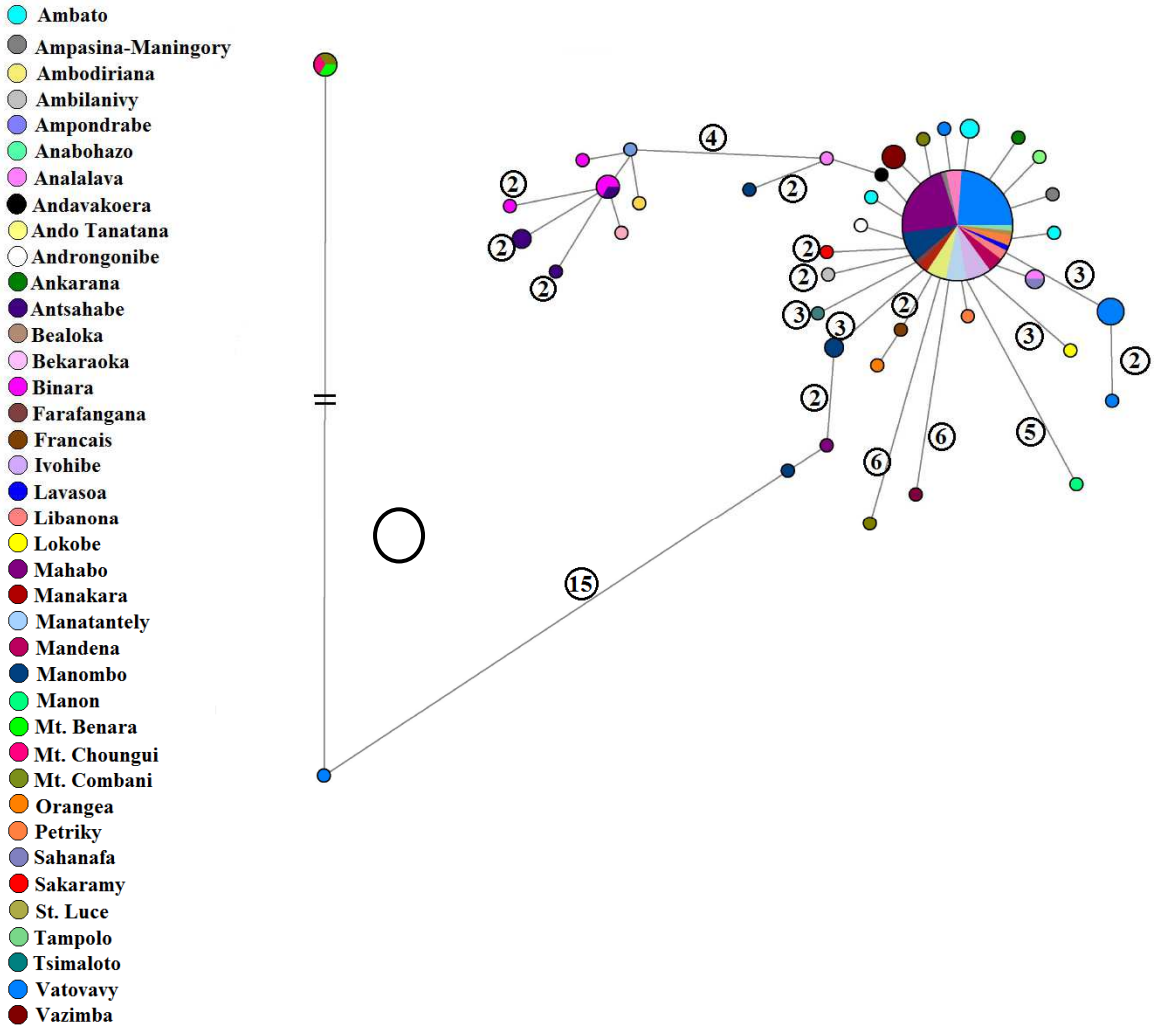


Figure 1.4 The median-joining haplotype network of *A. madagascarensis* haplotypes. The colours of localities on the map correspond to the colours on the haplotype network. The size of each haplotype is proportional to the number of individuals that share that haplotype. The number of mutations that are greater than one are indicated on the connection in the haplotype network.

1.3.4 Landscape genetic patterns

The genetic landscape shape interpolation surface plot revealed that the greatest genetic discontinuities (shown as peaks on the plot), are among populations located in the south-eastern edge of the plot representing the sampling localities (Figure 1.5). The flat surface in

the rest of the plot could indicate that these populations exhibit a high level of genetic connectivity (Figure 1.5). This finding was consistent with the finding of the Mantel test which showed a weak but significant positive correlation between genetic and geographic distances (1000 permutations, r = 0.21, probability of a correlation greater than or equal to observed <0.01). This low correlation coefficient could be due to sample size. However, when the Mantel test was performed for each group separately, it was found that there was a weak and insignificant positive correlation between genetic and geographic distance for the North group (r = 0.0021, probability of a correlation greater than or equal to observed >0.01) and the South group (r = 0.079, probability of a correlation greater than or equal to observed >0.01). The offshore group showed a negative but insignificant correlation between genetic and geographic distance (r = -0.24, probability of a correlation greater than or equal to observed >0.01).

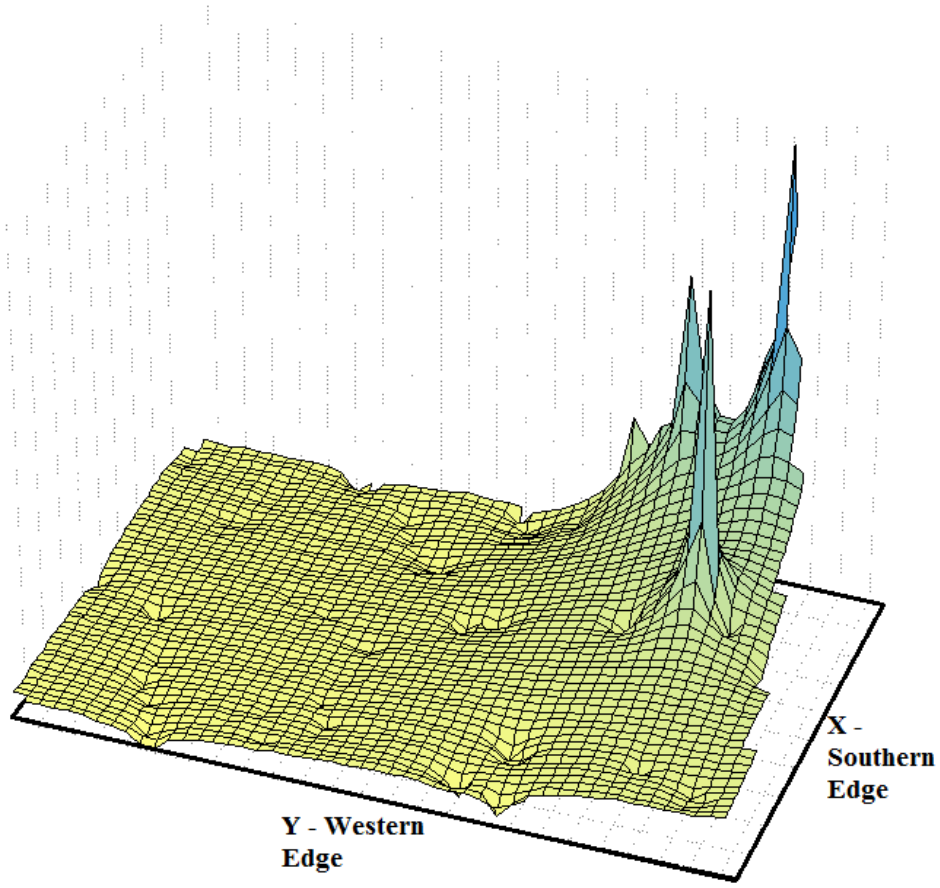


Figure 1.5. Genetic landscape shape interpolation plot for *A. madagascarensis*. The x and y axes indicate the geographic locations within the Delaunay triangulation network and the height of the surface plots show the average between interindividual genetic distances. The blue peaks show areas of high genetic variation and the flat yellow surface indicates areas of little or no genetic variation.

1.3.5 Analysis of molecular variance

Most of the genetic variation for *A. madagascarensis* was observed among groups (72.27%, $P<0.001$, Table 1.3). The values for all of the fixation indices were statistically significant at $\alpha = 0.05$. The least amount of genetic variation was found among populations within groups (9.57%, $P<0.01$, Table 1.3). The within population genetic variation was found to be 18.16% ($P<0.01$, Table 1.3).

Table 1.3. AMOVA design and results based on the groups recovered by the BAPS and haplotype networks analyses.

Source of variation	Percentage of variation	Fixation indices	*P*-value
Among groups (Northern, southern and Mayotte island groups)	72.27	0.72	0.00
Among populations within groups (individuals assigned to sampling localities)	9.57	0.35	0.005
Within populations	18.16	0.82	0.004

1.3.6 The use of barcode data for landscape genetic analysis

This case study using populations of *A. madagascarensis* from Madagascar and the neighbouring island of Mayotte, highlights how the DNA barcoding marker COI can be used below the species level in landscape genetics and phylogeographical analyses to elucidate barriers to gene flow and what landscape features may be acting as barriers. The BAPS clustering approach highlighted the presence of three genetically distinct clusters partitioning *A. madagascarensis* into a north, south and offshore group. This was confirmed by the median-joining haplotype network and the AMOVA analysis. The among group variation accounted for as much as 72% of the genetic variation.

Unsurprisingly, the populations from the Mayotte island are genetically differentiated from the other Madagascan populations. The presence of a shared haplotype suggests that there must be some contemporary (probably human-mediated) migration from the mainland Madagascar population to the island (Tollenaere et al., 2010). The genetic structure observed between the north and south group could be partially due to isolation by distance or the elevated landscape (Figure 1.1) between the north and south localities that could be acting as a barrier to gene flow. From the map of sampling localities, it appears that sampling was restricted to the extreme north and south of Madagascar and sampling was sparse over the rest of the area, especially at higher altitudes. This could indicate that populations of *A.*

madagascarensis do not occur at those sites, or that the sites were omitted by Smith *et al*. during sampling and hence this could mean that haplotypes connecting the north and south populations were not sampled.

In order for patterns revealed by landscape genetics and phylogeography to be unbiased and significant, populations should be sampled across their ranges. In addition molecular markers of different classes should be used to confirm or provide a contrast to any findings. In terms of conservation efforts, it has been shown that using a multi-species comparative landscape genetics or phylogeographic approach is more valuable and informative than using just one species (Chatzimanolis and Caterino, 2008; James et al., 2011).

1.4 Conclusion – problems and prospects

Although DNA barcoding and landscape genetics are undoubtedly useful to research and provide insight into the diversity of life, there are drawbacks associated with these disciplines. The main issue with both DNA barcoding and landscape genetics is that it is unlikely that a single 'universal' marker will be useful in all species across the diversity of life (Nielson and Matz, 2006). The COI gene has been used successfully in many animal studies, but there are several problems that are encountered since it is a mitochondrial gene which provides only a single locus identification system (Valentini et al., 2008b).

Mitochondrial genes are inherited only through the maternal lineage in most animals, therefore only a limited part of the evolutionary history is revealed (Dasmahapatra and Mallet, 2006). Interspecific hybridization and the unintentional amplification of pseudogenes are also problems that are encountered when relying only on mitochondrial genes (Bermingham and Moritz, 1998; Zhang and Hewitt, 2003). In order to ensure reliability for identifying species and making inferences about evolution and movement of populations, more efforts should be made to find several nuclear markers that will supplement and confirm the findings made by COI data (Dasmahapatra and Mallet, 2006).

Landscape genetic studies have the potential to enhance ecological knowledge, and will enable conservationists to manage the status of populations, study the effects of fragmentation and hence protect the genetic diversity of populations that are endangered (Manel et al., 2003; Segelbacher et al., 2010). This knowledge can be used to promote landscape connectivity of these endangered species because it allows for empirically based conservation corridors to be predicted and implemented (Segelbacher et al., 2010).

Taxonomists have argued that DNA barcoding cannot replace traditional taxonomy but it must be stressed that DNA barcoding in no way aims to replace taxonomy. The aim of DNA barcoding is to aid taxonomists to build and someday complete a global inventory of the diversity of life (Hajibabaei et al., 2005). Hopeful specialists in the field of DNA barcoding propose that one day, handheld DNA sequencing technology will be invented that can be used in the field by children, biologists or any enquiring mind, to find out the names and facts about any species on the planet (Hajibabaei et al., 2005; Savolainen et al., 2005).

Chapter Two

Revealing the diversity of Formicidae in an urban environment through DNA barcoding

Abstract

The biodiversity of Durban (eThekwini municipality) in KwaZulu Natal is primarily threatened by urbanization although other factors such as climate change and the spread of invasive species also pose significant threats. Knowledge of what species exist within the city is important for biodiversity surveillance, detecting invasive taxa and uncovering cryptic species. Conducting a comprehensive biodiversity inventory is a daunting task, especially for hyperdiverse groups such as terrestrial arthropods, where closely related species can often only be separated by subtle morphological characters. This study will investigate whether the barcoding marker, Cytochrome Oxidase C Subunit 1 (COI) can be used to efficiently and accurately delineate species of ants in comparison to traditional taxonomic approaches. The feasibility of DNA barcoding for assembling biodiversity inventories for urban areas which could be useful in conservation planning was also evaluated. A total of 619 individuals were sequenced from 23 geographic localities within the eThekwini region and surrounding regions. DNA barcoding revealed 80 provisional species/ "barcode clusters" or monophyletic lineages which could represent distinct species, while morphology revealed 51 different morphospecies. The accuracy of DNA barcoding in identifying species was tested by determining the optimal threshold for delineating species and evaluating a range of thresholds using three identification criteria, Best Close Match, Nearest Neighbour and ThreshID. The neighbour-joining tree based on K2P distances showed that the morphologically distinct species formed well-differentiated clusters. Interspecific genetic distance was higher than intraspecific genetic distance except in 27 cases due to singletons in the dataset. This highlighted the importance of sampling multiple individuals of the same species from different geographical localities. In this study, the utility of COI as a species diagnostic tool in ants is confirmed. The barcoding library constructed showed promise in highlighting reserves that should be preserved and possible cryptic speciation for further investigation. The continuous addition of new, updated sequences to the library from wide geographic ranges will increase the power and efficacy of finding a correct match for a species.

2.1 Introduction

Urbanization is expanding globally at a rapid rate which in turn has an adverse effect on the natural biodiversity. The city of Durban (eThekwini municipality) in KwaZulu Natal is located in the 'Maputaland Pondoland Albany' region (MPAR) – a biodiversity hotspot comprising of both terrestrial and aquatic ecosystems (Croucamp, 2009).

The considerable biodiversity within the MPAR is threatened by land transformation (agriculture and urban), alien plant and animal invasions and climate change. The eThekwini municipality has recently launched, in collaboration with the University of KwaZulu-Natal, a state of biodiversity programme. This program aims to track and monitor the status of biodiversity and natural resources in the urban environment and provide tools that can be used to monitor how much progress is being made to reduce loss of biodiversity in urban areas. This information will then be used to inform the public, policy-makers, stakeholders about the status of our biodiversity and implement conservation plans and promote sustainable use in the future (Anonymous, 2010).

As is the case with all cities around the world, there are large areas (approximately 49.1%) in Durban that have been dramatically altered from their natural state by human activities (Diederichs et al., 2010). This ultimately results in the loss of endemic biodiversity. However, in order to assess how much biodiversity is being lost, there needs to be a clear idea of how much biodiversity there actually is.

Knowledge of what species exist within the city is important for biodiversity surveillance, the detection of invasive taxa and uncovering cryptic species and new species. Conducting a comprehensive biodiversity inventory is a daunting task, especially for hyperdiverse groups such as terrestrial arthropods, where closely related species can often only be separated by subtle morphological characters.

The Barcode of Life initiative proposed a standardized method for the identification of species by sequencing a 658 base pair region of the mitochondrial gene, COI (Hebert et al., 2003b). The principle of DNA barcoding is straightforward; the barcode sequence from an unidentified specimen collected from anywhere around the world can be compared to an online, globally accessible digital library of DNA barcodes in order to identify the specimen. The project aims to revolutionize species identification and discovery by accelerating a process which can be tedious, time-consuming and expensive by traditional morphology-based taxonomy alone.

DNA barcoding has the potential to aid taxonomists by helping to highlight specimens that are cannot be assigned, to species that have already been described (Savolainen et al., 2005). Along with identifying species and discovering new species, DNA barcoding can also supplement and support research programs in ecology and other biodiversity disciplines (Hajibabaei et al., 2007; Savolainen et al., 2005).

However, in order for the full capability, power and reliability of DNA barcoding to be reached, it is essential that comprehensive reference libraries be constructed so that the probability of finding a true match for an unknown specimen is increased (deWaard et al., 2011). Many projects have successfully established reference libraries for several groups, for example DNA barcode libraries have been established for the Looper Moths of British Colombia, Canada (deWaard et al., 2011), the Arctic Life (Ephemeroptera, Plecoptera and Trichoptera) of Manitoba, Canada (Zhou et al., 2009), North American Ephemeroptera (Webb et al., 2012), the butterflies of Romania (Dinca et al., 2010) and many more. Few reference libraries, however, are currently available for African taxa and most regional barcoding projects have focused on rural areas that are not heavily impacted by human development. Investing in and compiling species inventories or the assembly of DNA barcode libraries within cities could be invaluable to urban conservation planning, biodiversity surveys and detection of invasive species. Green areas within cities can act as reservoirs of regional diversity and most importantly urban areas often depend on natural areas within the city to provide key ecological services such as water and air purification, noise reduction and for aesthetic appeal (Blaustein, 2013).

The aim of this study was to build a DNA barcode library for the Formicidae in the eThekwini region. Found in virtually any habitat in all parts of the world, the ants (Hymenoptera: Formicidae) are remarkable ecological engineers that are abundant, morphologically variable and temperamentally variable. They hold a great deal of influence over other invertebrate species as they act as predators and are involved in symbiotic and parasitic relationships with other species, including plants and bacteria (Bolton, 1994; Holldobler and Wilson, 1990). They are part of the family of social insects and have iteroparous and perennial colonies (Holldobler and Wilson, 1990). Ants are considered indicator species and are useful in highlighting areas that are high in biodiversity, biogeographic zones and points of evolutionary radiation as well as for monitoring change in the environment (Schoeman and Foord, 2012).

Despite their prominence in many ecosystems, the taxonomy of ants is poorly understood (Seppa, 2008). According to AntBase (Data accessed: 15 January 2013) 12649 species of ants have been described out of the 20 000 that are estimated to exist today. This leaves over 30% that still need to be discovered and described (Seppa, 2008). Thus, it is likely that there are more species that could be endangered or are already extinct. Ants are of conservation concern as 149 species are considered vulnerable on the IUCN Red List (Data accessed: 15 January 2013).

Surprisingly, studies involving ants are limited and do not feature as strongly as other taxonomic groups, such as mammals, and other arthropods such as Lepidoptera in scientific literature. This could be due to the ants being a taxonomically complex group (Bolton, 1994). Molecular data has proven to be instrumental in accelerating the rate of species discovery and description of ants. For example, DNA barcode studies have already been carried out on the ants of Madagascar and Mauritius (Smith and Fisher, 2009; Smith et al., 2005). These studies discovered high sequence divergence within some taxa that warrant further investigation (see chapter 1) in terms of their morphology, genetics, behaviour and life-history.

This study aims to begin the assembly of a barcode library for Formicidae of the eThekwini region. By adding barcode data generated in this study to the global initiative, the proportion

of ant species that matched BOLD reference entries was assessed in order to test the utility of DNA barcoding as an invasive species early detection and monitoring tool. The DNA barcoding method was tested for its efficiency and accuracy in delineating species of ants in comparison to traditional taxonomic approaches. By including individuals of the same species from multiple geographic localities, the effect of geographic scale of sampling on DNA barcode assignment was examined. This study highlights the feasibility of DNA barcoding for assembling biodiversity inventories for urban areas.

2.2 Materials and methods

2.2.1 Study area & sampling strategy

The eThekwini region is a metropolitan municipality covering a land area of 2297 km^2 including the city of Durban and surrounding areas. Located on the east coast of South Africa, in the province of KwaZulu-Natal, it is a sub-tropical coastal region, characterized by high temperatures, humidity and summer rainfall, and is influenced by the Indian Ocean and warm Agulhas current (Fairbanks et al., 2001). Although the region is largely urbanized, Durban is located within the Maputuland-Pondoland-Albany Region (MPAR), a globally important biodiversity hotspot (www.iucn.org). There are more than 2000 described species of plants, 82 terrestrial mammal species, 380 species of birds, 69 species of reptiles, 37 species of frogs and 25 endemic invertebrate groups found in the region (Croucamp, 2009). But this only represents a fraction of the biodiversity found in the city. According to the 2009-2010 State of Biodiversity Report by the eThekwini municipality, Durban has three of South Africa's eight terrestrial biomes (savannah, grassland and forest) ten distinct vegetation types (Eastern Valley Bushveld, KwaZulu-Natal Coastal Belt, KwaZulu-Natal Hinterland Thornveld, KwaZulu-Natal Sandstone Sourveld, Mangrove Forest, Ngongoni Veld, Northern Coastal Forest, Scarp Forest, Subtropical Coastal Lagoons, Subtropical Seashore Vegetation), and 97 km of coastline with a variety of beach types (Figure 2.1). The Durban Metropolitan Open Space System (D'MOSS) is a network of open spaces in the city that include nature reserves, large underdeveloped pieces of privately and municipality managed land, rural landscapes in upper catchments and riverine and coastal corridors. In total, D'MOSS consists

of about 74 711 ha of underdeveloped habitats, and integrates areas of high biodiversity linked together in the eThekwini municipal area.

Ants were collected from 18 green and developed areas within the city of Durban and from five surrounding localities (Pietermaritzburg, Port Shepstone and Northern KwaZulu Natal; Table 2.1). Geographic coordinates for all sampling localities included. At each site, ants were collected from all vegetation types present (Table 2.1). In most cases a site was only visited once, but some areas were sampled more than once to determine if sampling effort influenced the number of ant species recovered. Field work was conducted between July 2011 and April 2012; this period covered both the summer and winter seasons. Ants were collected using a variety of methods depending on vegetation type and size of the sampling site. Ants were hand collected from vegetation and litter in grassland and forest habitats for 10 minutes in a 1 m^2 quadrat placed at 10 different localities within a vegetation type. In forests, leaf-litter adjacent to the quadrat sampled was collected and specimens were then extracted using a Berlese funnel directly into 95% ethanol (EtOH). Yellow and blue pan traps containing sugar water were laid out in all habitats for the duration of the sampling period.

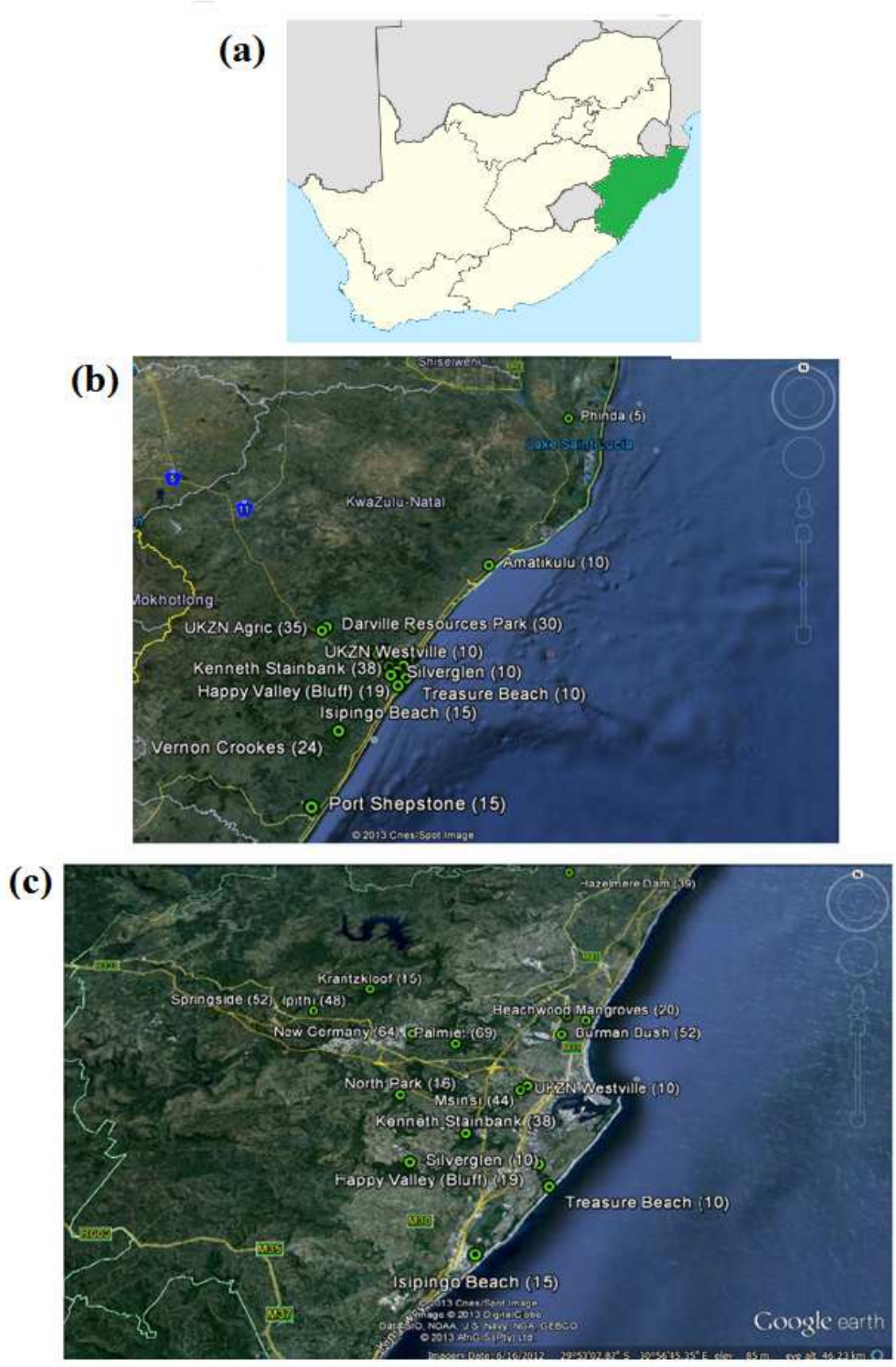


Figure 2.1. (a) Map of South Africa highlighting KwaZulu Natal. (b) A Google Earth image of the sampling localities within eThekwini and surrounding areas. (c) A Google Earth image of sampling localities in eThekwini central. The number of specimens barcoded in each locality is indicated in brackets.

Table 2.1. Sampling site locations and characteristics. * Denotes that a reserve was only sampled once.

Site	Vegetation type	Latitude	Longitude	Replicates
Amatikulu	Forest, grassland	29°6'54.72" S	31°36'9.36" E	*
Beachwood Mangroves	Mangroves, swamp	29°47'43.32"S	31°2'30.53" E	*
Burman Bush	Coastal bush	29°48'39.96"S	31°1'8.76" E	*
Darville Resources Park	Grassland	29°35'54.33"S	30°26'4.76" E	*
Happy Valley (Bluff)	Grassland, wetland, costal forest	29°55'55.92"S	30°59'33.72" E	*
Hazelmere Dam	Grassland	29°36'0" S	31°2'29.76" E	*
Ipithi	Grassland, forest	29°47'26.93"S	30°47'59.26" E	3
Isipingo Beach	Beach	29°59'59.28"S	30°56'42" E	*
Kenneth Stainbank	Grassland, forest	29°54'25.56"S	30°56'12.48" E	*
Krantzkloof	Forest	29°45'52.56"S	30°50'48.84" E	*
UKZN Agric PMB	Botanical garden	29°37'32.82"S	30°24'13.96" E	2
Msinsi	Grassland, forest	29°51'48.98"S	30°59'13.81" E	3
New Germany	Grassland, forest	29°48'45" S	30°53'19.32" E	2
North Park	Coastal bush	29°52'23.88"S	30°52'57" E	*
Palmiet	Grassland, riverine forest	29°49.35' S	30°55.58' E	3
Phinda	Bushveld	27°41.742'S	32°21.369'E	*
Port Shepstone	Urban garden	30°43'34"S	30°25'16" E	*
Seaton Park	Costal forest	29°47'28.54"S	31°1'30.71" E	*
Silverglen	Coastal grassland, bush clump mosaic	29°55'53.43"S	30°53'40.39" E	*
Springside	Grassland, forest	29°46'49" S	30°46'23" E	3
Treasure Beach	Dune forest, mangrove swamps, wetlands	29°57'0" S	31°0'0" E	*
UKZN Westville	Urban	29°52'3" S	30°58'50.88" E	*
Vernon Crookes	Grassland, coastal forest	30°17'17.52"S	30°33'43.56" E	2

2.2.2 Definition of morphospecies and taxonomic identification

Specimens were preserved in 95% ethanol and stored at 4 °C prior to sorting. Specimens were sorted to the lowest taxonomic level possible based on morphology (Bolton, 1994; Holldobler and Wilson, 1990). Taxonomic assignment was confirmed after consultation with the reference collection at Iziko South African Museum and collaborating taxonomist, Hamish Robertson. Because the identification of barcode clusters or BINs is based on the comparision of inter vs. intraspecific sequence divergence, multiple individuals per morphospecies were used in initial analyses. A maximum of five individuals per sampling location were selected for DNA barcoding to incorporate spatially-correlated variation within species (Bergsten et al., 2012). Each individual was given a unique sample identification number and photographed using a Leica Montage Stereo Light Microscope at the Microscopy and Microanalysis Unit (MMU) at the University of KwaZulu Natal. Collection details, taxonomic assignment and photographs of each specimen submitted for sequencing was uploaded onto the Barcode of Life Database (BOLD).

2.2.3 DNA extraction and amplification

A leg was excised from each individual and placed in a well of a 96-well microplate containing 50 µl of 95% EtOH. An entire individual was sampled in cases when they were too small for a leg to be excised. Voucher specimens are stored in a designated storage facility at -80 °C, at the University of KwaZulu-Natal or at the Iziko South African Museum. DNA extraction, PCR amplification and sequencing of the barcode region of the COI mitochondrial gene was performed at the Canadian Centre for DNA Barcoding (CCDB) at the University of Guelph, Canada, using standard protocols (Hajibabaei *et al*., 2005).

The COI sequences generated were aligned using the BOLD aligner option in the BOLD informatics workbench, which uses an amino acid based Hidden Markov Model (HMM) algorithm. Using this probabilistic model, the multiple sequence alignment is converted into a position-specific scoring system. This enables databases to be searched for homologous sequences (Eddy, 1998). The alignments were downloaded and inspected manually in

BioEdit 7.0.9.0 (Hall, 1999) for the presence of gaps and stop codons. All sequences and associated information (GPS coordinates, images and trace files) are available through BOLD. Of the 665 specimens submitted for barcoding, 619 were barcode compliant (sequence length more than 500 bp and not flagged as misidentification or contamination).

2.2.4 Sequence analysis and tree construction

The COI sequences from 624 specimens were used to construct a neighbour-joining (NJ) tree using the Kimura-2-Parameter (Kimura, 1980) model, as implemented in the 'sequence analysis' module of BOLD (Appendix 1). The specimens were assigned to BINs (Barcode Index Numbers) or 'barcode clusters' on the tree using a clustering algorithm implemented in the bioinformatics workbench of BOLD. The clustering algorithm creates operational taxonomic units (OTUs) and assigns specimens to putative species using sequence data. The Barcode Index Number System is useful when taxonomic information, particularly at the species-level, is unavailable. The system is reliable for species verification and uncovering diversity rapidly because the barcode sequence clusters show high concordance with species (Ratnasingham and Hebert, 2007).

Each sequence was matched against the library of COI barcode sequences in BOLD using the BOLD identification search engine and the database for all the barcode records (including records that have not yet been identified to the species level). When the match success was above 95%, the species name from the best top match was allocated to that morphospecies. Provisional genus-level identification was allocated to those sequences where no species-level match was available. The mean sequence composition was computed using the sequence alignment composition option on BOLD.

2.2.5 Analysing the DNA barcode data using Spider

Further analyses of the DNA barcode data were carried out in R statistical software (http://www.r-project.org) using the Spider library (Brown et al., 2012). Spider (Species

Identity and Evolution in R) contains various beneficial analyses for DNA barcoding, particularly for studies focused on the delimitation of species and speciation.

2.2.5.1 Measures of identification accuracy

The 'threshOpt' function was used to find the optimal threshold for identification given the data. The function computes the number of true positive, false negative, false positive and true negative identifications, including the cumulative error which is the sum of the false negative and false positive identifications. The analysis was performed with a range of threshold values from 0.1% to 5%.

Three functions in Spider (Best Close Match, Nearest Neighbour and Thresh ID) were used to test the accuracy of the identifications made by BOLD species identification engine. Each individual sequence is considered as an unknown and compared to the rest of the DNA sequences in the alignment.

The Best Close Match (Meier et al., 2006), returns the closest individual to the query as a correct match. If the result is "incorrect", then the closest match is different to the name of the query. A result of "ambiguous" indicates that more than one species is the closest match to the query while a result of "no ID" indicates that the query has no closest match. The Nearest Neighbour algorithm also finds the closest match to the query and if the name matches with that of its closest match, the result is returned as "TRUE" or "FALSE" if there is no match. The Thresh ID function carries out an analysis that is threshold based and similar to the specimen identification tool on BOLD. The results are interpreted similarly to that of the Best Close Match function, however, the Thresh ID function is more inclusive and includes all the results within the threshold rather than a single nearest-neighbour match. The Best Close Match and Thresh ID tests were carried out from 1% to 5% threshold levels since the threshold for delimiting majority of insect species is ≥3% (Stoeckle and Hebert, 2008) and because the threshOpt function indicated that the optimal threshold is 1.1% for the current data.

2.2.5.2 The barcode gap

To explore the existence of the 'barcode gap', intra and inter-specific distances were computed using the K2P distance measure. To further visualize the barcode gap, the furthest intraspecific value for an individual amongst members of its own species and the closest interspecific distance was calculated. The maximum intraspecific distance was subtracted from the minimum interspecific distance to check for the absence of the barcode gap.

2.2.6 Phylogenetic analysis

To determine the relationships among MOTUs, maximum likelihood (ML) and Bayesian analyses were conducted using one representative of each barcode cluster (n = 80). These two methods are model-based and the software jModeltest 0.1 was used to find the best-fit model for nucleotide substitution using the corrected Akaike Information Criterion (AICc). The model selected was TPM2uf + I + G (proportion of invariable sites = 0.3970 and gamma shape parameter = 0.4760). Since this particular model was not available in MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001), the GTR + I + G model was chosen as the closest approximation and implemented in maximum likelihood and Bayesian analyses. Maximum likelihood analysis was performed using the program Garli 0.96 (Zwickl, 2006). Branch support was assessed by a 100 nonparametric bootstrap replicates.

Bayesian Inference was performed using two independent runs with four parallel chains and five million Markov Chain Monte Carlo (MCMC) generations were run. Trees were sampled every 300^{th} generation. Priors were set to nst = 6, gamma and invariant sites. Tracer 1.5 was employed to inspect the convergence between estimated values of model parameters from the independent runs and their effective sampling sizes (ESS). The first 499980 trees were discarded as burnin.

2.3 Results

2.3.1 Sequence analysis

DNA was extracted from 665 individuals, and 624 individuals (94%) were successfully sequenced for COI. The final alignment consisted of 619 sequences as five sequences were not barcode compliant, i.e. sequence length less than 200 bp. Nucleotide composition values (Table 2.2) show that the sequences were AT rich which is expected in insect mtDNA (Crozier and Crozier, 1993).

Table 2.2. Sequence composition statistics for the complete alignment of the Formicidae of eThekwini (624 sequences).

	Minimum	Mean	Maximum	Standard Error
G%	9.73	12.07	23.10	0.041
C%	11.09	18.86	28.98	0.106
A%	24.23	29.43	34.95	0.059
T%	31.61	38.99	44.68	0.098
GC%	22.04	30.94	41.49	0.121
GC% Codon Pos 1	27.86	37.86	51.52	0.136
GC% Codon Pos 2	33.21	37.75	43.31	0.041
GC% Codon Pos 3	0.46	17.19	34.92	0.259

There was high nucleotide diversity and haplotype diversity for the Formicidae of eThekwini highlighting the high levels of biodiversity found in the area (Table 2.3). Although the sorting of specimens using morphology had recovered 51 morphospecies, analysis of the DNA data resulted in 80 DNA barcode clusters (BINs) as indicated on the neighbour-joining tree of K2P distances generated with BOLD (Appendix 1). With the assistance of a taxonomist, problematic individuals were formally identified to species or near species level.

Table 2.3. Summary sequence statistics for the complete alignment of the Formicidae of eThekwini excluding non-barcode compliant sequences (619 sequences) and the sequence alignment of one representative of each BIN.

Alignment	**n**	**V**	**PI**	**H**	**h**	**SD**	**π**	**SD**
Complete	619	373	362	92	0.97	0.003	0.234	0.002
BINs	80	369	347	80	1.00	0.002	0.214	0.003

2.3.2 Species identification

A total of 25 out of 80 (31%) barcode clusters (BINs) exhibited a 95% or higher match at the species-level when using the BOLD specimen identification tool. Seven out of 80 (9%) of the barcode clusters exhibited a 95% or higher match at the genus level and 48 out of 80 (60%) had a match below 95% at either the species or genus level. In total our sample includes representatives of 25 genera and six subfamilies (Table 2.4). The genus *Tetramorium* was the most well-represented in our dataset, followed by *Pheidole*, *Pachycondyla*, and *Camponotus*.

In most cases the DNA data recovered the same taxonomy as the morphological assignments, with some notable exceptions. Individuals belonging to the BIN cluster '532504' were identified as *Anochetus grandidieri* by the molecular data (with an 84% match) however, they were identified as *Microdaceton exornatum* using traditional taxonomy. Individuals belonging to the BIN cluster '483802' were identified as *Anochetus madagascarensis* by DNA data (with a 95.97% match) while traditional taxonomic methods identified them as *Anochetus bequaerti*. Individuals belonging to BIN cluster '133387' were morphologically identified as *Camponotus nr. cintellus*, while the BOLD identification showed a 98.92% match with *Camponotus AFRC*. Individuals identified as *Pachycondyla peringueyi* with a 83.33% match on BOLD, corresponding to the BIN cluster '499613', were identified as *Pachycondyla havilandi* by traditional taxonomy. An individual from the BIN cluster '469662' was identified to genus level on BOLD as *Pachycondyla*, with a match of 100%. Traditional morphology taxonomy was able to refine the taxonomy and identify the individual to species level as *Pachycondyla sculpturata*.

Table 2.4. Results of provisional species assignment of barcode clusters based on searches on BOLD.

Identification by BOLD	BIN no.	Provisional species name assigned	Subfamily	%Match
Aneuretus simoni	483797	*Aneuretus etk*	Aneuretinae	87.38
Anochetus grandidieri	532690	*Anochetus grandidieriD*	Ponerinae	94.26
Anochetus grandidieri	483804	*Anochetus grandidieri*	Ponerinae	92.91
Anochetus grandidieri	494812	*Anochetus grandidieriA*	Ponerinae	91.98
Anochetus grandidieri	532504	*Anochetus grandidieriB*	Ponerinae	84.65
Anochetus grandidieri	532569	*Anochetus grandidieriC*	Ponerinae	86.79
Anochetus madagascarensis	483802	*Anochetus madagascarensis*	Ponerinae	95.97
Anoplolepis custodiens	515500	*Anoplolepis custodiens*	Formicinae	97.38
Camponotus AFRC	133387	*Camponotus AFRC*	Formicinae	98.92
Camponotus AFRCSA	133907	*Camponotus AFRCSA*	Formicinae	99.03
Camponotus eugeniae	133898	*Camponotus eugeniae*	Formicinae	99.69
Camponotus maculatus	8304	*Camponotus maculatus*	Formicinae	99.50
Camponotus niveosetosus	133278	*Camponotus niveosetosus*	Formicinae	98.30
Camponotus petersii	500732	*Camponotus petersii*	Formicinae	97.53
Cataulacus	481213	*Cataulacus ETKF*	Myrmicinae	99.07
Cataulacus erbrardi	2981	*Cataulacus erbrardi*	Myrmicinae	97.30
Crematogaster	262235	*Crematogaster ETKB*	Myrmicinae	89.04
Crematogaster	262236	*Crematogaster ETKC*	Myrmicinae	84.26
Crematogaster castanea	17180	*Crematogaster castanea*	Myrmicinae	98.61
Hypoponera	483796	*Hypoponera ETKJ*	Ponerinae	97.53
Lepisiota AFRC	264251	*Lepisiota AFRC*	Formicinae	96.45
Lepisiota canescens	494813	*Lepisiota canescens*	Formicinae	95.37
Lepisiota crinita	515769	*Lepisiota crinita*	Formicinae	97.99
Lepisiota crinita	264252	*Lepisiota crinitaB*	Formicinae	94.44
Lepisiota incisa	264250	*Lepisiota incisa*	Formicinae	99.85
Leptogenys	262510	*Leptogenys ETKD*	Ponerinae	85.09
Leptogenys	483794	*Leptogenys ETKI*	Ponerinae	92.13
Leptogenys elongata	494811	*Leptogenys ETK12*	Ponerinae	83.62
Leptogenys AFRC	498867	*Leptogenys AFRC*	Ponerinae	88.20
Monomorium indet	532573	*Monomorium ETK13*	Myrmicinae	92.13
Monomorium termitobium	532688	*Monomorium ETK14*	Myrmicinae	86.85
Myrmicaria natalensis	511076	*Myrmicaria natalensis*	Myrmicinae	100
Oligomyrmex	471121	*Oligomyrmex ETKG*	Myrmicinae	95.83
Oligomyrmex	481216	*Oligomyrmex ETKH*	Myrmicinae	90.60
Oligomyrmex	498740	*Oligomyrmex ETKP*	Myrmicinae	89.00

Identification by BOLD	BIN no.	Provisional species name assigned	Subfamily	%Match
Oligomyrmex	501888	*Oligomyrmex ETKR*	Myrmicinae	99.54
Pachycondyla caffraria	262511	*Pachycondyla caffrariaA*	Ponerinae	98.17
Pachycondyla caffraria	483795	*Pachycondyla caffrariaC*	Ponerinae	97.69
Pachycondyla caffraria	483813	*Pachycondyla caffrariaB*	Ponerinae	97.71
Pachycondyla	469662	*Pachycondyla ETKF*	Ponerinae	100
Pachycondyla peringueyi	499613	*Pachycondyla ETK01*	Ponerinae	83.33
Pachycondyla tarsata	500346	*Pachycondyla tarsata*	Ponerinae	97.55
Pachycondyla villosa	500324	*Pachycondyla ETK02*	Ponerinae	86.50
Paratrechina	469861	*Paratrechina ETKF*	Formicinae	99.85
Paratrechina	532572	*Paratrechina ETKX*	Formicinae	90.57
Pheidole MG145	19139	*Pheidole megad*	Myrmicinae	99.39
Pheidole	261760	*Pheidole megab*	Myrmicinae	85.88
Pheidole	261761	*Pheidole megac*	Myrmicinae	100
Pheidole	261918	*Pheidole ETKA*	Myrmicinae	99.85
Pheidole	507075	*Pheidole ETKT*	Myrmicinae	89.91
Pheidole MG145	514931	*Pheidole MG*	Myrmicinae	98.31
Pheidole	532691	*Pheidole ETKY*	Myrmicinae	86.70
Plagiolepis indet	35746	*Plagiolepis indet*	Formicinae	99.83
Plagiolepis	483803	*Plagiolepis ETKK*	Formicinae	99.38
Plagiolepis	483814	*Plagiolepis ETKM*	Formicinae	99.23
Plagiolepis	532362	*Plagiolepis ETKV*	Formicinae	98.61
Plectroctena mandibularis	499292	*Plectroctena ETK03*	Ponerinae	86.70
Polyrachis shistacea	262703	*Polyrachis shistacea*	Formicinae	99.85
Ponera pennyslvanica	498868	*Ponera ETK04*	Ponerinae	84.74
Proformica	497931	*Proformica ETKO*	Formicinae	84.06
Pyramica brevicornis	532574	*Pyramica ETK05*	Myrmicinae	84.07
Solenopsis punctaticeps	506479	*Solenopsis punctaticeps*	Myrmicinae	98.77
Technomyrmex pallipes	261945	*Technomyrmex ETK06*	Dolichoderinae	94.65
Technomyrmex fisheri	481212	*Technomyrmex ETK07*	Dolichoderinae	90.54
Tetramorium humbloti	7716	*Tetramorium humbloti*	Myrmicinae	99.07
Tetramorium quadrispinosum	36449	*Tetramorium quadrispinosum*	Myrmicinae	98.77
Tetramorium grassii	154316	*Tetramorium grassii*	Myrmicinae	98.36
Tetramorium frigidum	261762	*Tetramorium frigidum*	Myrmicinae	96.60
Tetramorium	264778	*Tetramorium ETKE*	Myrmicinae	99.12
Tetramorium	481214	*Tetramorium ETKG*	Myrmicinae	87.86
Tetramorium simillimum	490123	*Tetramorium ETK08*	Myrmicinae	91.94

Identification by BOLD	BIN no.	Provisional species name assigned	Subfamily	%Match
Tetramorium proximum	532568	*Tetramorium ETK09*	Myrmicinae	89.66
Tetramorium erectum	532570	*Tetramorium ETK10*	Myrmicinae	94.44
Tetramorium	532571	*Tetramorium ETKW*	Myrmicinae	88.73
Tetraponera clypeata	483805	*Tetraponera ETK11*	Pseudomyrmicinae	92.17
Tetraponera	483812	*Tetraponera ETKL*	Pseudomyrmicinae	92.91
Tetraponera	497624	*Tetraponera ETKN*	Pseudomyrmicinae	97.70
Tetraponera	499152	*Tetraponera ETKQ*	Pseudomyrmicinae	91.75
Tetraponera	509309	*Tetraponera ETKU*	Pseudomyrmicinae	93.20

2.3.3 Threshold optimization & measures of identification accuracy

The lowest cumulative error (3 sequences) using the threshold optimization function, was at the 1.1% threshold level while the highest cumulative error (107 sequences) was observed for threshold values between 4.5 to 5% (Figure 2.2, Table 2.5). The optimal threshold for delimiting ant species (1.1% COI sequence divergence) was lower than the 3% COI sequence divergence considered to be the standard threshold for delimiting insect species (Hebert *et al*., 2003b). As the percentage threshold for delimiting species increases, the number of false negatives increases.

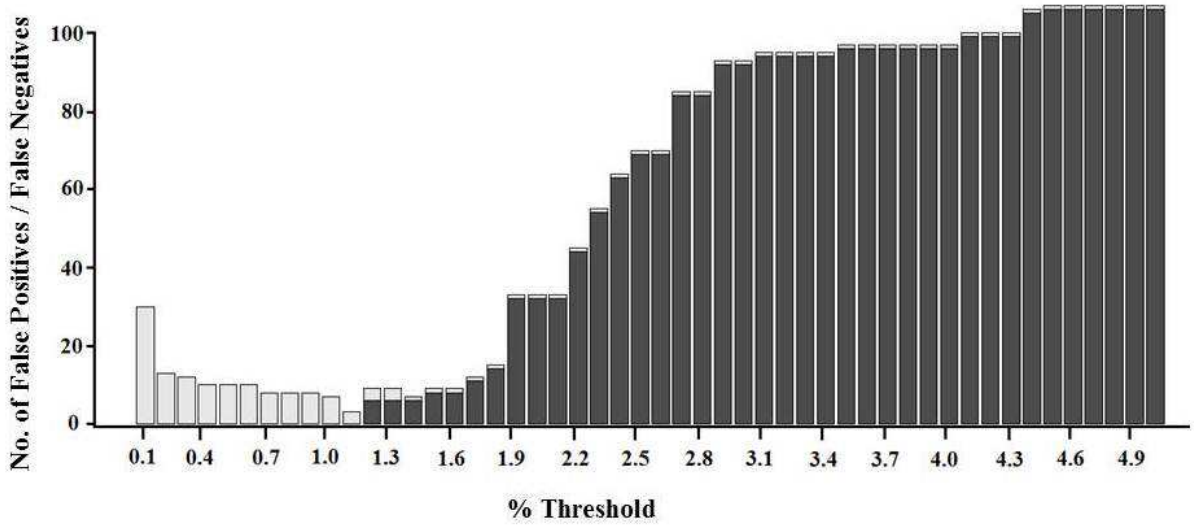


Figure 2.2. A barplot of the false positives (light grey) and false negative (dark grey) rate of identification for the Formicidae of eThekwini from a threshold of 0.1 to 5.0%.

The identification criteria analysis was done to evaluate what percentage threshold COI sequence divergence would be considered the optimal threshold for the DNA barcode library for the ants of eThekwini. The Nearest Neighbour criteria identified 598 “TRUE” matches (nearest species name has the same name as the query) and 21 “FALSE” identifications (no species name matches the query). The Best Close Match and ThreshID criteria yielded similar results. As the percentage threshold increased, there was an increase in the number of “Incorrect” and “No ID”. With the threshID criteria, there was also an increase in the number of “Ambiguous” results as the percentage threshold increased. The analysis suggests that the optimal threshold for species delimitation was 1.1% for the ants of eThekwini because 595 individuals were identified correctly and 24 individuals had no closest matches in the library (No ID result, Figure 2.3). Changes to the percentage threshold had no effect on species monophyly. The result of species having no closest matches was due to the presence of singletons in the dataset, i.e. only one representative from a species.

Table 2.5. The results of the threshold optimization analysis of thresholds from 1.0 – 5.0%.

% Threshold	**True negatives**	**True positives**	**False negative**	**False positives**	**Cumulative error**
1.0	21	591	0	7	7
1.1	21	595	0	3	3
2.0	20	566	32	1	33
3.0	18	508	92	1	93
4.0	18	504	96	1	97
5.0	16	496	106	1	107

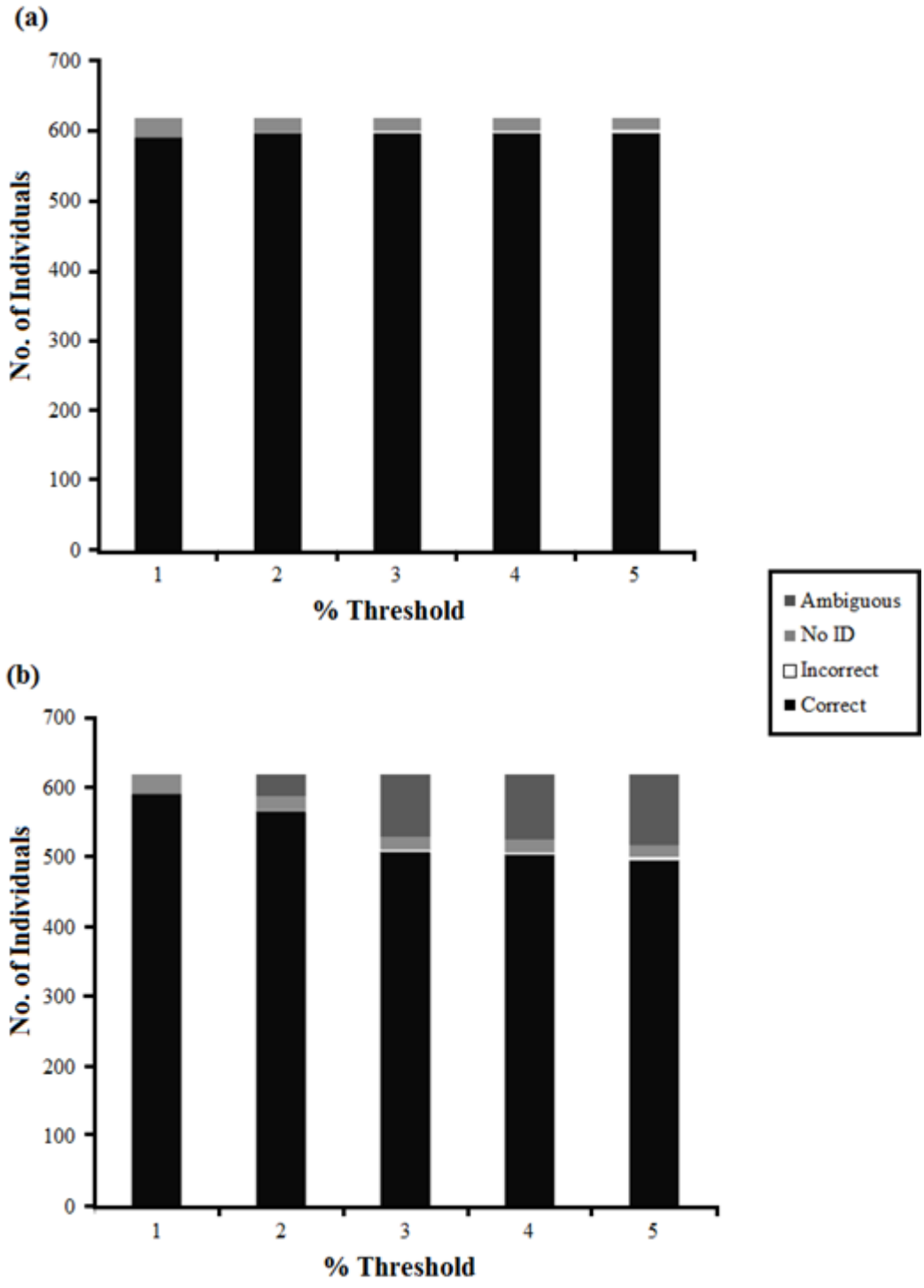


Figure 2.3. A barplot of the success rate of the (a) Best Close Match identification criteria and the (b) ThreshID identification criteria at 1 – 5 % thresholds.

2.3.4 The barcode gap

The COI Kimura two-parameter (K2P) sequence divergences were much higher between species than within species (Figure 2.4 and 2.5) and suggests that COI is a good marker to use for species delimitation within the eThekwini region. This finding meets one of the critical assumptions of DNA barcoding which dictates that interspecific genetic distance should be higher than intraspecific genetic distance (Hebert et al., 2004a; Meyer and Paulay, 2005). The average among-species COI sequence divergence value was 0.26. The minimum interspecific genetic distance was 0.04 while the maximum interspecific genetic distance was 0.4. The minimum intraspecific genetic distance was 0 and the maximum intraspecific genetic distance was 0.02 (Figure 2.4).

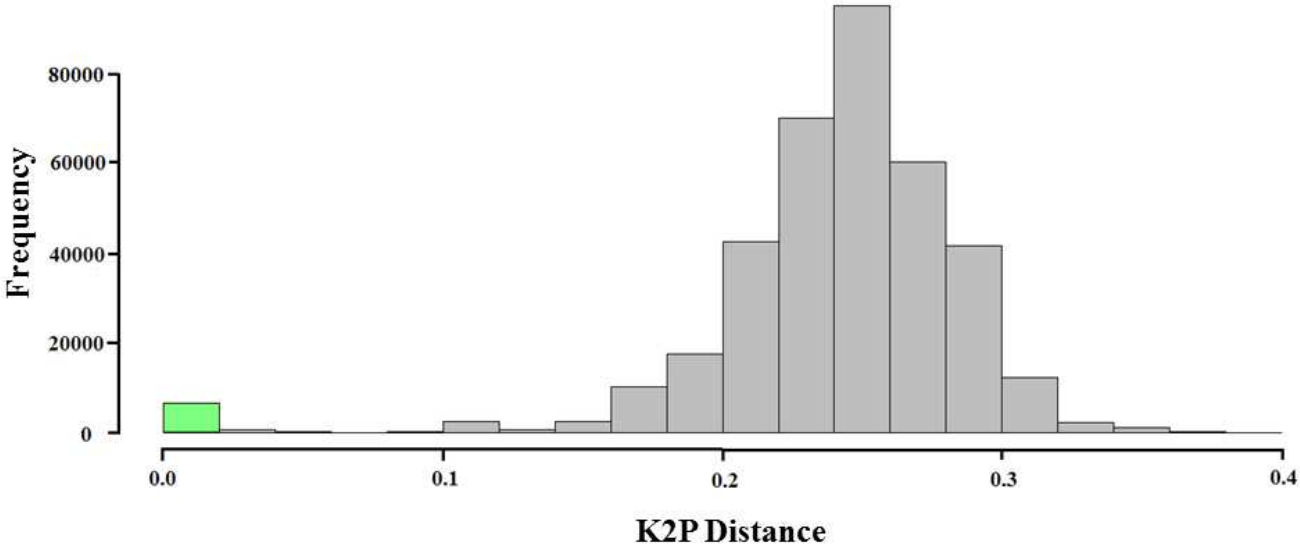


Figure 2.4. Distribution of pairwise distances calculated using the K2P model. The green bar indicates the intraspecific genetic distance and the grey bars indicate the interspecific genetic distances.

However, there were 27 individuals where there was a zero or negative difference when the maximum interspecific distance was subtracted from the minimum interspecific distance. This indicates a lack of the barcode gap in these cases and are shown in red in Figure 2.5. Again, this is due to singletons because of insufficient spatial sampling in the reference library.

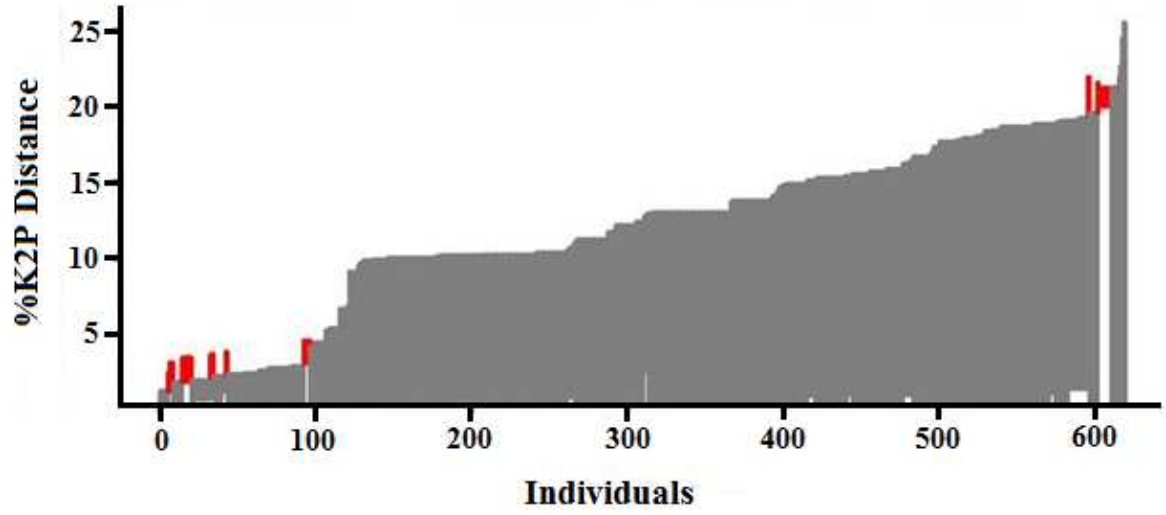


Figure 2.5. A line plot of the barcode gap for the Formicidae of eThekwini. For each individual in the dataset, the grey lines represent the furthest intraspecific distance (bottom of line value), and the closest interspecific distance (top of the line value). The red lines show where this relationship is reversed.

2.3.5 Phylogenetic analysis

The phylogenetic relationships between representatives of each barcode cluster (BINs) are illustrated in Figure 2.6. The maximum likelihood and Bayesian Inference consensus trees were congruent with each other and Figure 2.6 represents the maximum likelihood tree with bootstrap and posterior probability values annotated onto the branches. The tree was rooted at the midpoint. All values above 75% bootstrap and 0.95 posterior probability were considered significantly well-supported, while values below 50% bootstrap and 0.50 posterior probability were considered weakly supported and not shown on the tree. There were several genera that were not monophyletic; these included the genera *Pheidole*, *Tetraponera*, *Oligomyrmex*, *Leptogenys*, *Pachycondyla*, *Ponera*, *Myrmicaria*, *Monomorium*, *Pyramica* and *Anochetus*. Species in the genera *Pheidole*, *Lepisiota*, *Camponotus*, *Pachycondyla*, *Oligomyrmex*, *Anochetus*, *Tetraponera* and *Leptogenys* all contained multiple BINs. This could indicate cryptic speciation and species from four of these genera (*Pheidole*, *Lepisiota*, *Camponotus* and *Pachycondyla*) will be further examined in chapter 4 using a phylogeographic and landscape genetics approach.

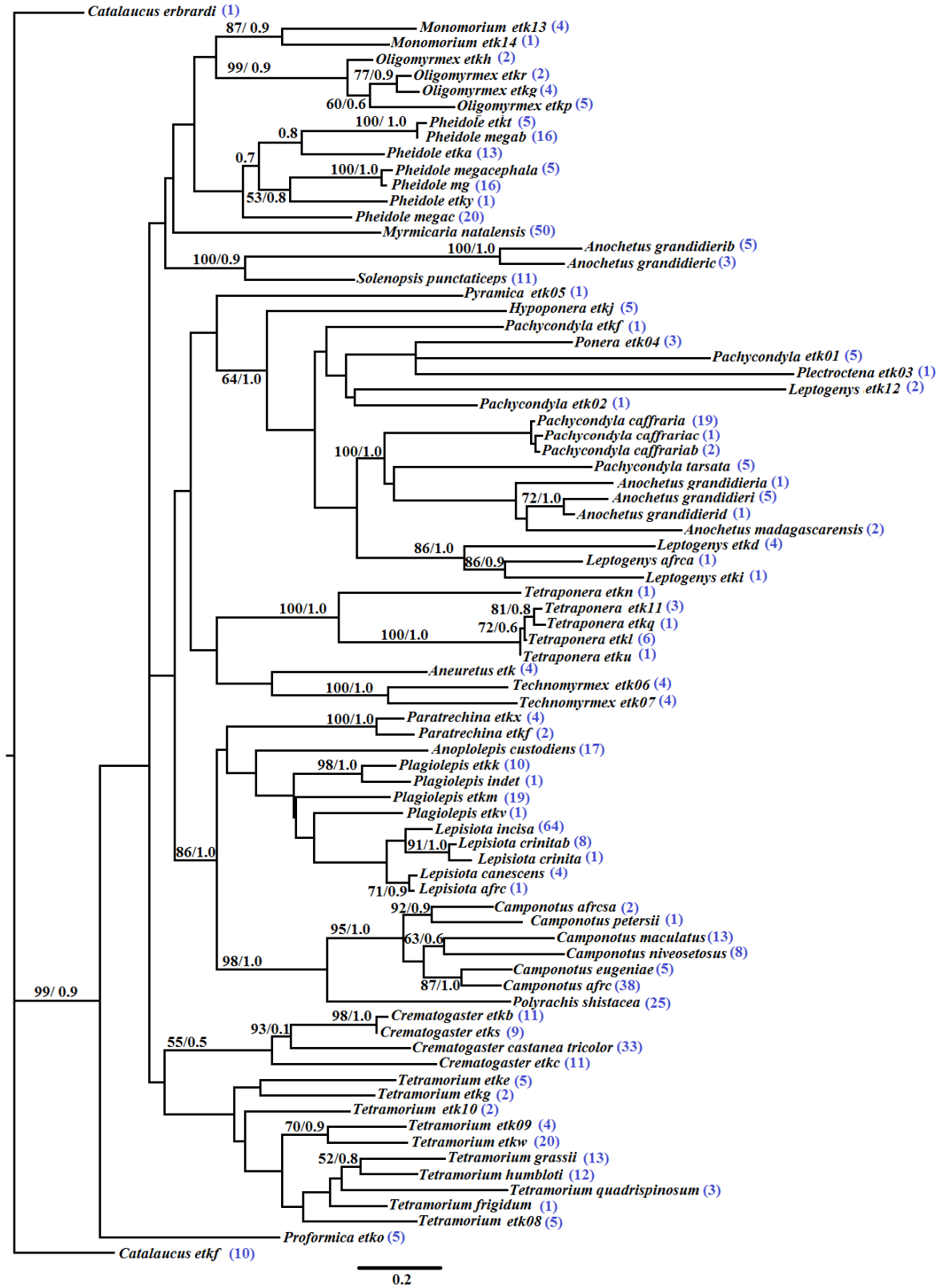


Figure 2.6 A maximum likelihood tree of one individual per BIN. Bootstrap values above 50 and Bayesian posterior probabilities above 0.5 are shown. The number of individuals per BIN are also indicated.

2.4 Discussion

2.4.1 A DNA barcode library for eThekwini Formicidae

The aim of this chapter was to test the utility of the DNA barcoding marker COI for the rapid assessment of Formicidae biodiversity in various nature reserves/ “green areas” in the eThekwini region. A total of 51 morphospecies were found using morphological characters alone, whereas 80 distinct barcode clusters or were found using COI sequences. The current DNA barcode library for the Formicidae of eThekwini consists of 619 sequences from 80 potential species from 26 genera, and six subfamilies. Spatial sampling included individuals, across 23 nature reserves/ “green areas”.

2.4.2 Identification success using DNA barcoding

Although there are potential shortfalls using either morphological taxonomy or DNA barcoding, when used in conjunction, they could be very effective in identifying species. An integrative approach could aid in achieving more sound, reliable and efficient species identifications. In a study addressing if DNA barcodes could correctly assign unknown organisms to higher taxa when a species-level match is unavailable in a DNA barcode library, it was shown that tree-based methods successfully assigned, with high accuracy, to higher taxonomic levels (Wilson et al., 2011). In contrast, another study demonstrated that tree-based methods ranked the worst for identifying species (van Velzen et al., 2012). Therefore, along with the neighbour-joining tree, other methods were also used in conjunction to test the ability of DNA barcoding to identify species of ants. One of the major criticisms of DNA barcoding is that the integrity of identifications are compromised by false positives which is overestimating the true number of species, and false negatives which is an underestimate of the true species number (Packer et al., 2009). This can be overcome by using a threshold optimized to have the least cumulative error (Meyer and Paulay, 2005). In the DNA barcode study of the ants of Madagascar, the optimal threshold of 2-3% was found to be suitable for the data (Smith et al., 2005). In the present study of ants within the eThekwini region, the BCM and ThreshID criteria were evaluated between the 1% (least cumulative error) and 5% (most cumulative error) thresholds (Figure 2.2 and 2.3, Table 2.5). Both criteria had the most

correct identifications at a threshold of 1.1% COI sequence divergence which is less than the optimal threshold range identified for the ants of Madagascar.

Using the global data available through BOLD, only 25 out of 80 (31%) of the barcode clusters could be identified to species level with a match of over 95%. This highlights the lack of comparative information available on African taxa in the database and justifies the need for the present study. Generally, the morphological delimitation underestimated the diversity of ant species in the city (51 morphospecies vs 80 molecular BINs). This was not always the case, for example, five barcode clusters were identified as *Anochetus grandidieri* despite individuals being morphologically quite diverse. Taxonomic identification identified one species as being part of an entirely different genera, *Microdaceton exornatum*. Two of the other species did belong to the Genus *Anochetus*, but one was taxonomically identified as *Anochetus bequaerti* and the other as *Anochetus nr. natalensis*. This suggests that a large amount of the specimens sampled in this study do not have representatives in the BOLD database, and are therefore unique. This highlights the usefulness of local barcode libraries.

2.4.3 Effects of sampling on inter and intraspecific diversity

Due to the sampling regime of this study, there were some species that were only represented by a single sequence. Obtaining multiple individuals of the same species is critical for the investigation of interspecific as well as intraspecific variation (Fisher, 1999). Using the Best Close Match and ThreshOpt identification criteria, there were 24 instances where there was a "NO ID" result (Figure 2.3). This was due to the query sequence having no closest match in the dataset because of the presence of singletons in the dataset.

Species with only a single representative were also responsible for a zero or negative difference when the maximum interspecific distance was subtracted from the minimum interspecific distance, and can obscure the barcode gap (Figure 2.5). This emphasizes the need to sample multiple representatives of the same species for inclusion in a DNA barcode library in order to increase the probability of a successful identification where an unknown query is concerned.

Geographic sampling is also crucial. In a study investigating the effect of the increase in geographic sampling on DNA barcoding; it was shown that interspecific, intraspecific and proportion of monophyletic species are significantly affected by the geographic scale of sampling (Bergsten et al., 2012). Measuring variation in a single, small region would underestimate genetic variation whereas sampling a species across its geographic range will unveil more genetic variation. Although this study only focused on a small regional sampling scale, by making the data available through BOLD, I hope that the impact of geographic sampling can be assessed at a national and global scale.

2.4.4 Cryptic species

Cryptic speciation presents a challenge to quantifying biodiversity. The estimation of species richness and endemism is crucial to identifying habitats which qualify for conservation. The discovery of cryptic speciation has been accelerated by the ease of obtaining DNA sequences from organisms. Investigating whether cryptic speciation is more common in a certain geographic area, habitat, biome or vegetation type could enhance our knowledge and speed up conservation efforts (Bickford et al., 2006). In my study, the DNA barcode data provides evidence for possible cryptic speciation for *Lepisiota sp*., *Camponotus sp*., *Oligomyrmex sp*., *Pachycondyla sp*., and *Pheidole sp.* Morphologically they were very difficult to differentiate but each morphospecies grouped into more than one barcode cluster (Figure 2.6). In chapter 4 some of these species will be investigated using a landscape genetics approach.

2.5 Conclusion

Due to the increased risk of species extinctions in urban areas due to natural resources being overexploited, the invasion of non-native species and the destruction of natural habitats, rapid access to biodiversity data is essential. Appropriate measures can only be undertaken if the biodiversity data is comprehensive, accurate and up to date. The DNA barcode library built in this study has the potential to provide such data and new, updated records can easily be added to the library. The DNA barcode library can be improved by the continuous addition of new sequences from new species, as there are many more species expected to be encountered. More importantly, the DNA barcode data including its complementary data such as

photographs, taxonomic and geographic information, can serve as “DNA-based maps of biodiversity” (Hajibabaei, In Press).

A major advantage associated with DNA-based identification is that there are no limitations to how the data can be analysed. Alternative models and methods can be implemented to improve accuracy (Wilson *et al*., 2011). In this study, DNA barcoding proved to be effective in delimiting species and discovering provisional cryptic speciation which warrants further investigation using approaches such as phylogeography or landscape genetics. The barcode data could then be used to assess the diversity of species in an ecosystem as well as within-species genetic variability. This information can be used by stakeholders to inform the public and implement conservation plans.

Chapter Three

Investigating species richness and species assemblages of ants in eThekwini

Abstract

The transformation of natural ecosystems into urban areas has led to the threat of habitat fragmentation and increase in invasive species. The development of nature reserves or "green" areas within the urban environment has the potential to link habitats and promote connectivity and could alleviate some of the negative effects of urbanization. DNA barcode data from the Formicidae of eThekwini project was used to investigate species richness and diversity of ants distributed in nature reserves, parks, beaches and urban gardens in and around the city of Durban. The effects of the size of green areas and their proximity from the city centre on the number of species (BINs) were also investigated. Extrapolation measures of species richness indicated that as many as 153 species of ants could occur in the city. Biodiversity estimation indices showed that Ipithi nature reserve was the most diverse while Phinda and an urban garden in Port Shepstone was the least diverse. Species assemblages varied between reserves but it was observed that grassland and forest habitats supported unique species assemblages. There was no significant effect of the size of the green areas or the distance from the city centre on the number of species found. This study demonstrated that open spaces within eThekwini and surrounding areas contribute to maintaining diversity within an urban environment. These findings may assist in planning future urban development with taking the biological diversity into consideration.

3.1 Introduction

Natural ecosystems have been transformed by urbanization. The natural native vegetation which supports high species diversities and richness, is reduced in urban settlements by habitat fragmentation (Heino et al., 2005). The resulting patches also promote the increase of invasive species, which often outcompete endemic species (Suarez et al., 1998). Strategies to curb the negative impacts of urbanization on biological diversity, include the preservation of natural vegetated areas within urban landscapes through the creation of nature reserves and the development of green corridors that link protected habitats and promote connectivity and ecological structure and function within urban landscapes (Hamaide et al., 2006; Linehan et al., 1995; Pacheco and Vasconcelos, 2007; Yasuda and Koike, 2009). Deciding which areas should be given protection is difficult. Along with compiling biodiversity inventories, it is also necessary to identify key habitats and vegetation types that support the persistence of the natural biodiversity. In this study, the species assemblages, diversity and species richness of ants (family Formicidae) were investigated among several nature reserves, parks, beaches and urban gardens in the city of Durban (eThekwini municipality) and surrounding towns.

Ants were selected for use in this study as bio-indicators because as terrestrial invertebrates, ant diversity provides a valuable data source, which can be used in conservation planning and the design and selection of nature reserves. They are easy and cost-effective to sample, process and inventory compared to vertebrates (Kremen et al., 1993). They are also found in most terrestrial ecosystems and are also successful in most urban environments (Clarke et al., 2008). They can delimit areas of high endemism and biogeographic zones and identify points of evolutionary radiation (Kremen et al., 1993). There are a few examples of studies that have focused on the effects of urbanization on ants, their persistence and species richness (Buczkowski and Richmond, 2012; Clarke et al., 2008; Heterick et al., 2013; Ives et al., 2013; Lessard and Buddle, 2005; Menke et al., 2010a; Menke et al., 2010b; Pacheco and Vasconcelos, 2007; Thompson and McLachlan, 2007; Yamaguchi, 2005) However, few studies have been conducted on the effects of urbanization on indigenous ants in Africa.

The aim of this chapter is to investigate the distribution pattern of species assemblages in the eThekwini region and surrounding areas. The main questions addressed in this chapter are: First, if ant species diversity and richness is correlated to the size of urban green areas (nature reserves, parks, urban gardens and beaches) and if the different vegetation types present in the city support unique ant assemblages. If urban green areas act as refugia for native taxa it was expected that larger open spaces conserve greater species richness and diversity. Second, this study will investigate how invasive ant species are distributed in the city. Given that the Durban Harbour is one of the busiest ports in Africa, it is expected that the distribution of invasive species would be closely associated with the port and trade hotspot of the city centre (Hulme, 2009)

3.2 Materials and methods

3.2.1 Data

Data from the DNA barcode library for the Formicidae of eThekwini was utilized for this study (see chapter 2). This molecular-based species inventory also included important information such as spatial data, the size (in hectares) of nature reserves and parks (open spaces) sampled, the distance of the sampling locations from the Durban city centre and harbour and the types of vegetation within each sampling locality (Table 2.1, Chapter 2).

3.2.2 Species richness

A preliminary measure of species richness in each open space was done by comparing the number of barcode clusters (BINs) to the number of morphospecies (see chapter 2). A species accumulation curve was generated for the entire sampling region using the accumulation curve option on the BOLD informatics workbench. The curve was based on the number of genetic BINs and the total number of specimens sampled. The species diversity in each open space was also assessed as a measure of haplotype diversity, by constructing haplotype accumulation curves using the Spider package in R (Brown et al., 2012). This approach was also used to evaluate the sampling effort. For each locality, the sequences were subsampled at

random and the mean accumulated number of haplotypes was calculated as well as the standard deviation. A 1000 random permutations of the data were performed. An abundance data file consisting of the number of individuals per BIN in each sampling locality was used as input into the statistical package, PRIMER 6 (PRIMER-E, Ltd, Lutton, Ivybridge, UK). The total species richness of each locality was estimated using three extrapolation methods namely, Chao 2, first order Jacknife (Jacknife 1) and the Michaelis-Menten richness estimator, each with a 1000 permutations to create species accumulation curves. These methods are very useful in estimating species richness when species inventories are not complete and when there are low sample sizes as in this case where a maximum of five individuals per sampling location were included (Chao et al., 2005, 2006; Magurran, 2004).

Using PRIMER 6, the biodiversity of ant species in each of the open spaces was quantified by using Margalef's diversity index *d* (Margalef, 1958), the Shannon diversity index *H'* (Shannon and Weaver, 1963), Simpson's diversity index *D* (Simpson, 1949) and Fisher's α (Fisher et al., 1943). These diversity indices are the most widely used indices in ecological studies (Collwell and Coddington, 1994) and were evaluated for their applicability in this study.

3.2.3 Spatial distribution of species richness

To investigate the effect of vegetation type on species richness, a parsimony analysis was conducted using the program 'Mix' in the Phylip 3.6 package (Felsenstein, 2005). The input matrix was constructed using presence and absence data with the localities and particular vegetation type within that locality assigned as the terminal taxa, and the presence and absence of morphospecies or BINs coded as characters. If a morphospecies was present in a locality or a vegetation type, it was scored as '1' and if it was absent, it was scored as '0'. Although there are 10 different vegetation types in the region (http://bgis.sanbi.org/municipalities/summaries.asp?muni=ETH) we classified the different sampling sites broadly into grassland, forest and beach/urban. In cases where there were two different vegetation types sampled within one open space, then the presence and absence of taxa were scored for each vegetation type. The similarity of species assemblages between

each sampling site was also assessed by utilising the Bray-Curtis (Bray and Curtis, 1957) similarity measure to construct a cluster analysis dendogram using PRIMER 6.

3.2.4 Effects of open space size and distance from city centre on species richness

The effects of open space area and distance from the Durban city centre on the number of species (BINs) was tested using linear regression in SigmaPlot 12 (Systat Software, San Jose, CA). A total of 16 open spaces for which both area and distance from the Durban city centre values were available were used in the analysis (Table 3.1). The UKZN Life Science and Agriculture Botanical Garden and Darville Resources Park were not included in this analysis because no data was available for the area of these open spaces. The values for locality area and the distance from city centre were $\log_{10}$ transformed.

Table 3.1. Area and distance from Durban city centre of open spaces.

Open space	**Area (ha)**	**$\log_{10}$**	**Distance from city centre (km)**	**$\log_{10}$**
Amatikulu	2100	3.32	105.24	2.02
Beachwood Mangroves	76	1.88	7.17	0.86
Burman Bush	50	1.70	5.04	0.70
Happy Valley	45	1.65	8.67	0.94
Ipithi	12	1.08	22.49	1.35
Kenneth Stainbank	250	2.40	9.68	0.99
Krantzkloof	532	2.73	19.52	1.29
New Germany	110	2.04	13.55	1.13
North Park	53	1.72	13.24	1.21
Palmiet	50	1.70	9.82	0.99
Phinda	17000	4.23	309	2.49
Seaton Park	6	0.78	7.48	0.87
Silverglen	220	2.34	14.61	1.16
Springside	21	1.32	25.34	1.40
Treasure Beach	600	2.78	10.50	1.02
Vernon Crookes	2189	3.34	64.45	1.81

3.3 Results

3.3.1 Species richness and diversity

Both the number of BINs (n = 80) and the number of morphospecies (n = 51) were used in this study to estimate the biodiversity of the ants in the 18 open spaces within the eThekwini region and five open spaces in surrounding areas. Figure 3.1 shows the number of barcode clusters (based on COI data) in comparison to the number of morphospecies per open space. The number of barcode clusters were generally similar to the number of morphospecies identified per locality. The number of morphospecies only exceeded the number of barcode clusters in Burman Bush, Springside Nature Reserve and Hazelmere Dam. Ipithi Nature Reserve had the highest number of BINs, followed by New Germany Nature Reserve and Palmiet Nature Reserve. The least number of species (n=1) were found in Phinda Game Reserve and an urban garden in Port Shepstone. The most commonly encountered species were *Lepisiota incisa* (BIN no. = 264250) and *Pheidole sp*. (BIN no. = 507075, 261760, 261918, 19139, 514931, 532691, 261761) which were found in 14 and 17 out of 23 sites respectively. These species are considered to be invasive species (Fournier et al., 2012; Sithole et al., 2009) and this could explain why they are more common, widespread and abundant compared to the native species in this study.

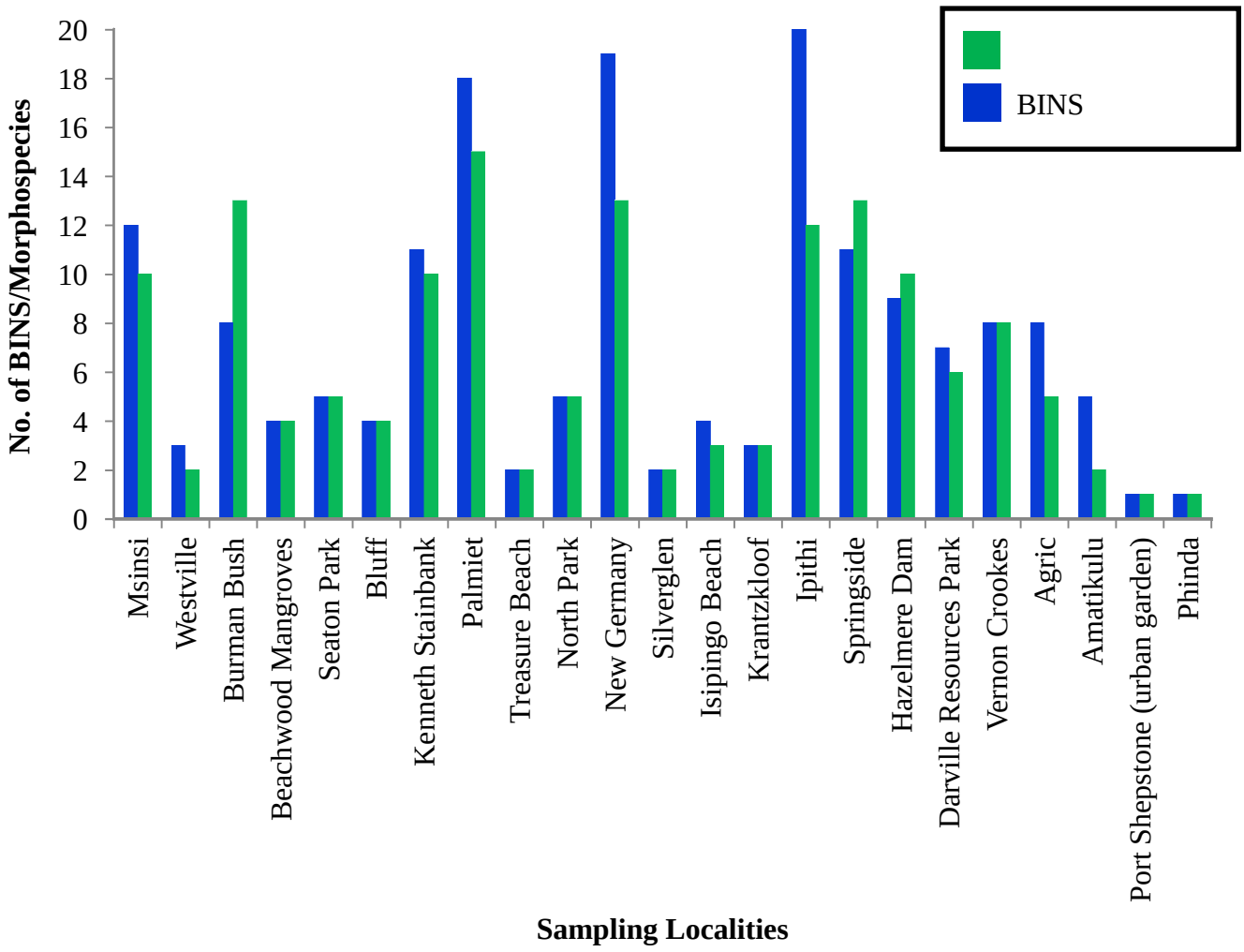


Figure 3.1. Graph of the number of barcode clusters (BINs) and morphospecies sampled per open space. The localities are ordered from the closest (Msinsi) to the furthest (Phinda) from the city centre and harbour.

The accumulation curve (Figure 3.2) generated using the BINs has not reached an asymptote, which indicates that there are still more potential species to be sampled within eThekwini and surrounding areas. This result is not unexpected given the narrow sampling period and highlights the high level of diversity present within the city.

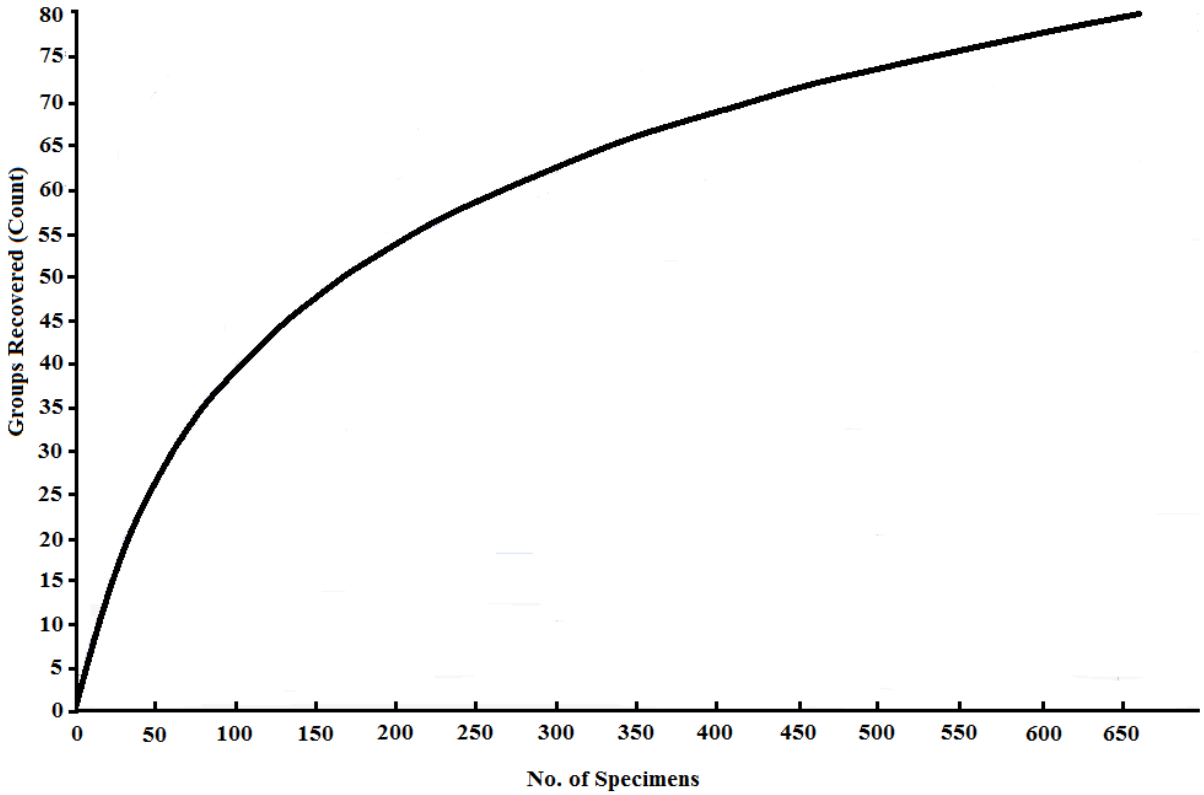


Figure 3.2. Accumulation curve for the entire sampling region generated with BOLD using the number of BINs.

Figures 3.3A-H shows the haplotype accumulation curves for each open space that was sampled more than once in the present study. The three most diverse localities in terms of haplotypic diversity were Ipithi, New Germany and Palmiet. Even though some of the curves (Burman Bush and Life Science and Agriculture Botanical Garden) have started to reach a plateau, it is clear that additional sampling would be needed for the complete inventory of ant biodiversity of the eThekwini region.

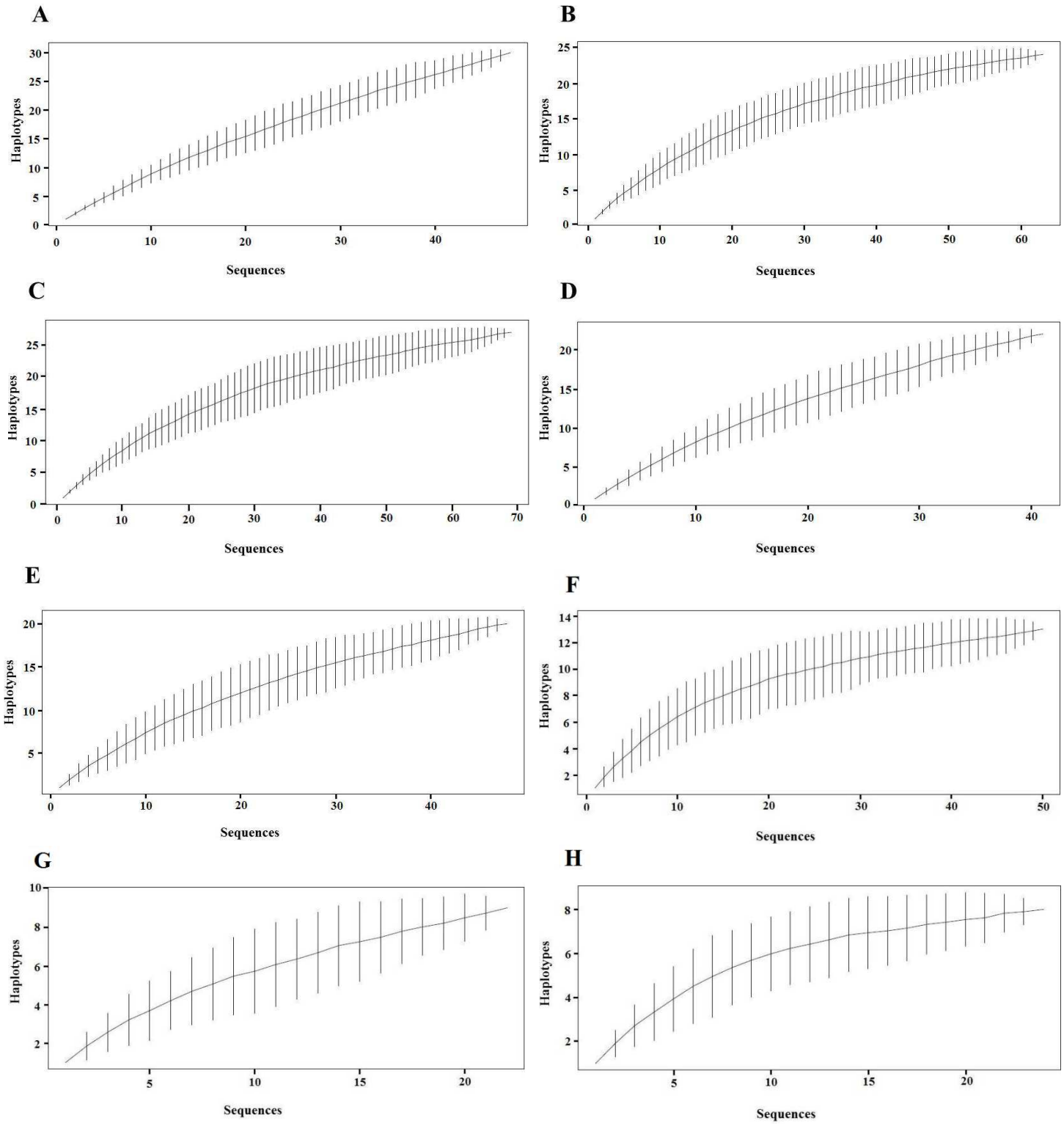


Figure 3.3. Haplotype accumulation curves for (A) Ipithi, (B) New Germany, (C) Palmiet, (D) Msinsi, (E) Springside, (F) Burman Bush, (G) Vernon Crookes, (H) Life Science and Agriculture Botanical Garden (PMB). The standard deviation is also shown. Ipithi nature reserve is the most diverse in terms of the number of haplotypes.

Extrapolation measures such as Chao 2, Jacknife 1 and Michaelis-Menton richness estimator were used to approximate the total species richness if sampling had continued (Figure 3.4). The Chao 2 measure estimated that at least 153 species could be encountered, while the Jacknife 1 measure was more conservative and suggested that at least 124 species could be encountered. The Michaelis-Menten estimator was the most conservative measure and indicated that at least 80 species should be encountered. Interestingly this is the same number of BINs recovered by the DNA barcode data.

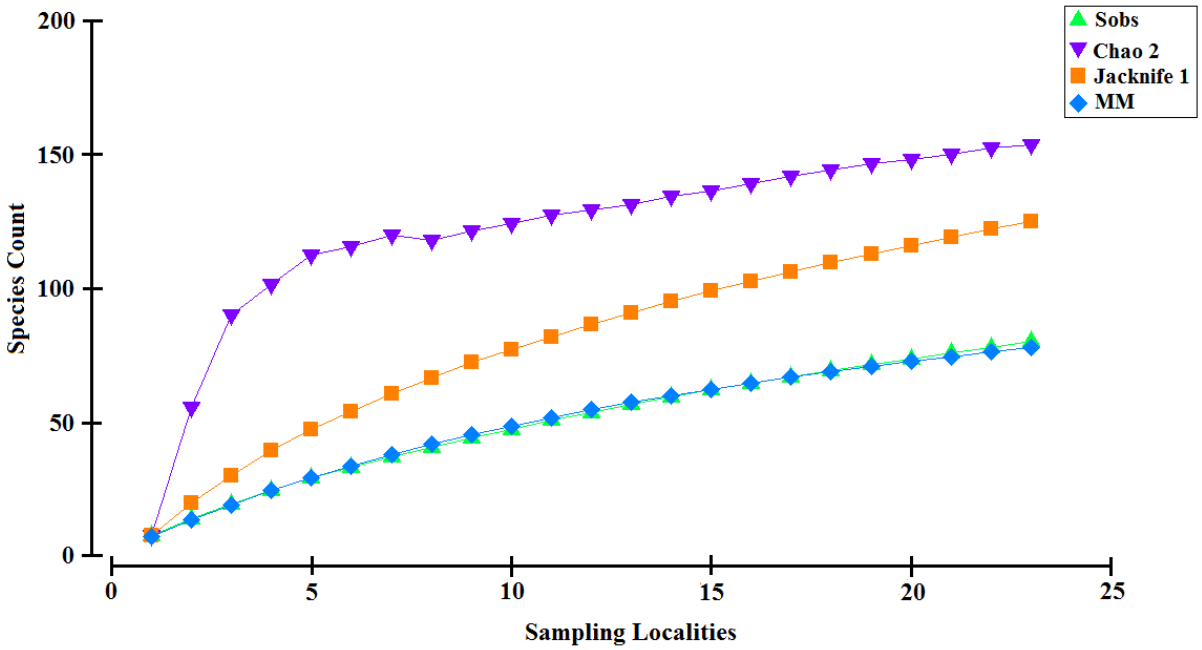


Figure 3.4. Species accumulation curve for the 23 sampling localities extrapolated using three extrapolation methods. S_{obs} is the observed number of species, while the Chao 2, Jacknife 1 and Michaelis-Menten (MM) estimators approximate the estimated number of species if sampling had continued.

Table 3.2. Biodiversity estimation indices where S is the total number of COI BINs, N is the total number of individuals, *d* is Margalef's diversity index, *H'* is the Shannon diversity index, *D* is the Simpson diversity index and α is Fisher's α.

Open space	**S**	**N**	***d***	***H'***	***D***	***α***
Amatikulu	5	10	1.74	0.62	0.80	3.97
Beachwood Mangroves	4	20	1.00	0.60	0.79	1.50
Burman Bush	9	49	2.06	0.85	0.85	3.24
Darville Resources Park	6	30	1.47	0.78	0.86	2.25
Happy Valley	4	19	1.02	0.60	0.79	1.55
Hazelmere Dam	9	12	2.31	0.86	0.87	4.16
Ipithi	20	49	4.88	1.22	0.95	12.61
Isipingo Beach	3	15	0.74	0.48	0.71	1.13
Kenneth Stainbank	12	37	3.05	0.99	0.91	6.17
Krantzkloof	3	15	0.74	0.48	0.71	1.13
Msinsi	11	40	2.71	0.98	0.91	5.01
Palmiet	18	69	4.02	1.18	0.94	7.92
New Germany	19	63	4.34	1.19	0.94	9.24
North Park	5	13	1.56	0.64	0.81	2.97
Seaton Park	5	25	1.24	0.70	0.83	1.88
Silverglen	2	10	0.43	0.30	0.56	0.75
Springside	11	49	2.57	0.96	0.89	4.41
Treasure Beach	2	10	0.43	0.30	0.55	0.75
UKZN Agric	8	24	2.20	0.84	0.88	4.20
UKZN Westville	2	9	0.46	0.30	0.55	0.80
Vernon Crookes	8	22	2.26	0.78	0.84	4.52

The Margalef's diversity index (*d*) is calculated from the total number of species and the total number of individuals present in the sampling area (Magurran, 2004; Margalef, 1958). The highest value for *d* was observed for Ipithi nature reserve (Table 3.2; $d = 4.88$) while the lowest value for *d* was obtained for Phinda and Port Shepstone (Table 3.2; $d = 0$) where only one species of ant was collected in each of these open spaces. The Shannon diversity index accounts for both species abundance and evenness in a given area (Magurran, 2004; Shannon

and Weaver, 1963). This measure of diversity also indicated that Ipithi nature reserve was the most diverse open space sampled and also indicates that the species are evenly distributed (Table 3.2; H' = 1.22). The Simpson's diversity index D is the probability that in an infinitely large community, any two individuals sampled will belong to the same species (Magurran, 2004; Simpson, 1949). The values for D were highest for Ipithi, Palmiet and New Germany and were 0.95, 0.94 and 0.94 respectively (Table 3.2). Fisher's α diversity index defines how individuals sampled are divided among species in the sampling area (Fisher et al., 1943; Magurran, 2004). The highest value for Fisher's α was 12.61 and was recorded for Ipithi nature reserve (Table 3.2).

The biodiversity indices were highest for Ipithi nature reserve followed by New Germany and Palmiet nature reserve indicating that these open spaces are high in ant biodiversity (Table 3.2). The lowest values for biodiversity indices were observed for Phinda game reserve and the urban garden in Port Shepstone where only one species was recorded.

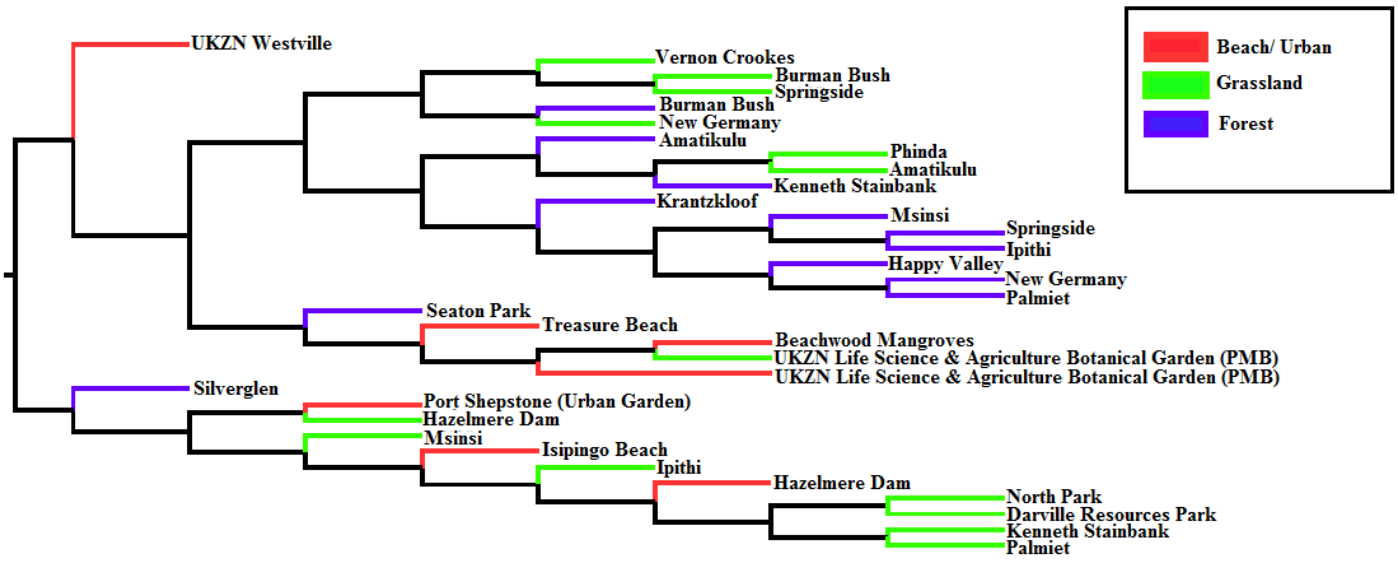
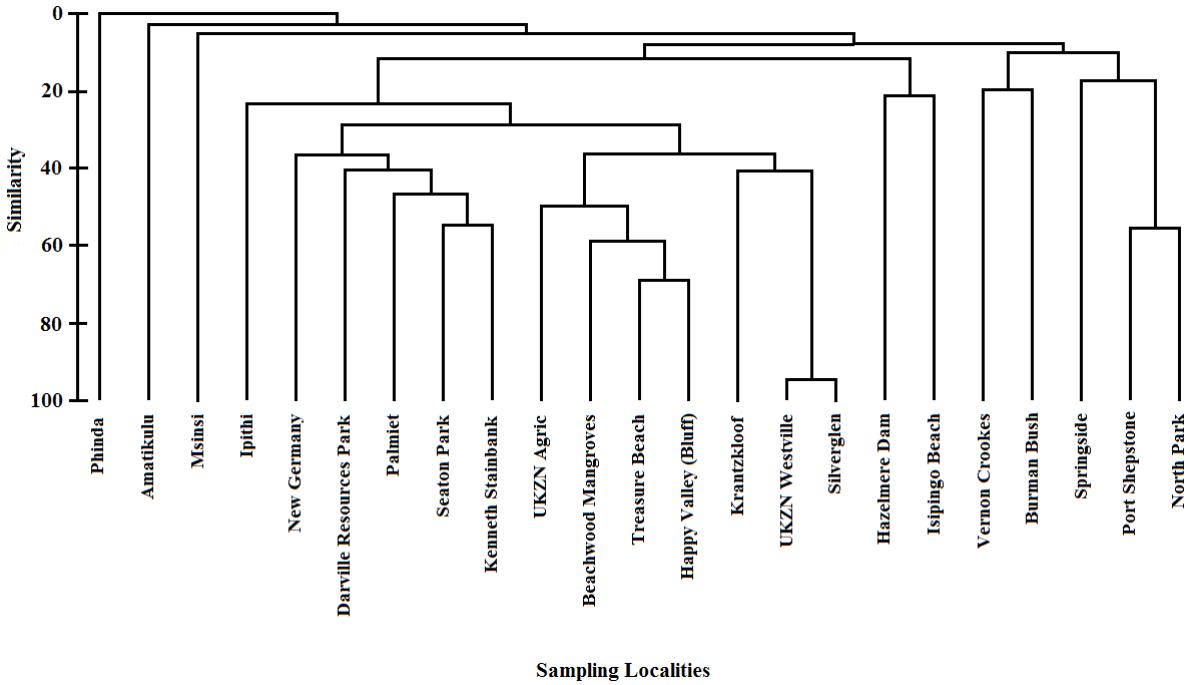


Figure 3.5. Parsimony tree of binary data based on presence or absence of a morphospecies within a reserve and a particular vegetation type.

3.3.2 Spatial distribution of species richness

In the investigation of species richness, two approaches were used. In the first approach, presence/absence data for each locality and vegetation type within the locality was used to construct a parsimony tree to assess if there was any similarity between species assemblages in different habitats. In this approach, ecosystems that share species will be drawn together in the tree. The topology in Figure 3.5 suggests that different vegetation types occurring in the same open space do not share a significant proportion of species (or BINs), instead similarity in vegetation type (forest, grassland and beach/urban) seems to be a more important determinant of species richness. In particular, the forest habitats of Msinsi, Springside, Ipithi, Happy Valley, New Germany and Palmiet seems to share a number of species. The grassland and beach/urban ecosystems do not form monophyletic lineages. This is probably because the classification system I used (grassland, forest and beach/urban) was too simplistic to capture the diversity of the ten different vegetation types found in the city. Nonetheless, what is clear from the clustering pattern observed is that geographic proximity is not a good indicator of species richness.



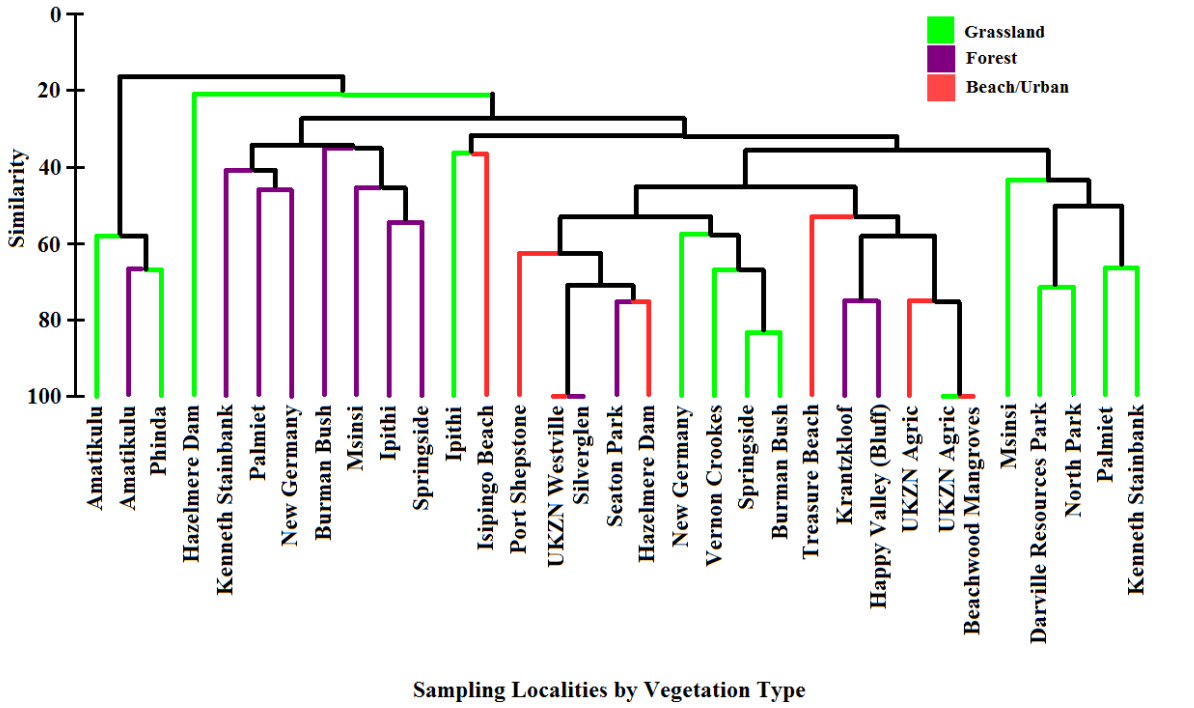


Figure 3.6. (a) Cluster analysis dendogram based on the Bray-Curtis similarity, showing the similarity between ant assemblages in the different sampling localities, (b) similarity between ant assemblages in the different sampling localities by vegetation type.

The second approach utilised the Bray-Curtis similarity index to construct a cluster analysis dendogram based on a resemblance matrix created from abundance data (Figure 6a) and the presence and absence data for each vegetation type within a locality (Figure 6b). The localities with the most similar species assemblages were UKZN Westville and Silverglen (94%). The next most similar assemblage, 69% similar, were the Happy Valley and Treasure Beach nature reserves. The Beachwood Mangroves nature reserve shared a 66% species assemblage similarity to Treasure Beach and Happy Valley sampling localities. The majority of the localities shared species assemblage similarities of approximately 20% to 60%. The single species collected from Phinda game reserve was not collected from any other locality (0%). The Bray-Curtis similarity analysis with localities divided into vegetation types recovered similar patterns to the parsimony analysis in Figure 5. Species assemblages in grassland types clustered together, for example the grassland in Burman Bush shared an 83% species assemblage similarity with the grassland of Springside. Forest vegetation types such as Krantzkloof and Happy Valley (Bluff) shared a 75% species assemblage similarity. The urban garden or beach species assemblages shared similarity with either forests or grasslands, for example, the urban species assemblage in UKZN Westville shared a 100% similarity to the forest in Silverglen and the urban species assemblage in the Beachwood Mangroves shared a 100% similarity to the grassland in UKZN Agric.

3.3.3 Effects of sampling locality size and distance from city centre and harbour on species richness

There was no significant relationship between the number of BINs recorded from an open space and the size of the urban green areas sampled ($r = 0.438$, $r^2 = 0.192$, Figure 3.7a). Similarly there was no correlation between ant diversity and the distance from the Durban city centre ($r = 0.178$, $r^2 = 0.0315$, Figure 3.7b).

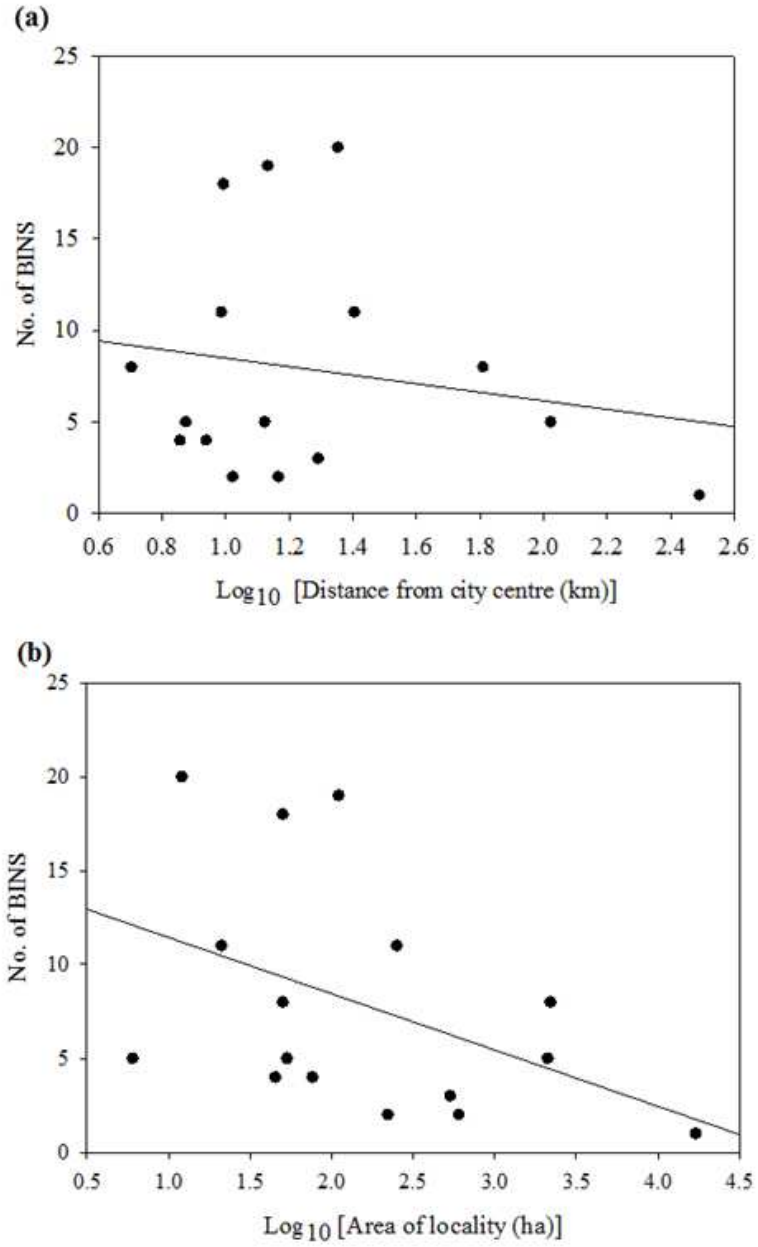


Figure 3.7. Relationship between (a) the number of species (BINS) recorded for the ants of eThekwini and the distance from the Durban city centre (n = 16) and (b) the area of each locality (n = 16).

3.4 Discussion

This study revealed a remarkable level of species richness and diversity within the ants of Durban. The number of potential species sampled in this study (80 species, 25 genera) can be compared favourably to other studies which examined the effects of urbanization on ants. The majority of the studies which investigated the effects of urban development on ants have been conducted in North America, one of the most urbanized regions in the world. In a survey of six urban and one forest land-use type in North Carolina, USA, 54 species of ants were found (Menke et al., 2010b). In San Francisco, USA, 24 protected natural areas within urban parks were surveyed and 15 species of ants were recorded (Clarke et al., 2008). A total of 40 species from 20 genera were collected from Fort Benning in Georgia, USA (Graham et al., 2009). There were 53 ant species from 24 genera collected from a network of 172 conserved lands in southern California (Mitrovich et al., 2010). In the streets of New York City, 13 species of ants were collected (Pecarevic et al., 2010).

The effects of urbanization on ants have also been evaluated in Asia. For example, 43 species were collected from 98 parks in Tokyo and Chiba City, Japan (Yamaguchi, 2005). In other regions of the world such as Australia, 60 morphospecies from 34 genera were encountered in 12 remnant riparian corridors in Sydney (Ives et al., 2013). In the threatened Cerrado Biome of Brazil, 12 public squares, two urban parks and three nature reserves were surveyed. A total of 143 species belonging to 39 genera were found (Pacheco and Vasconcelos, 2007). In all of these examples, it was found that urbanization did have an effect on ant species diversity and richness. Many of these studies have extensively sampled many localities while only 23 were sampled in this study. Despite this, many more species were found in this study which provides evidence for the support of Durban (eThekwini municipality) as a biodiversity hotspot. This study contributes towards a body of literature which highlights that ant species richness and biodiversity is greatest in the tropics and southern hemisphere, with diversity patterns of Formicidae being similar to that observed in other taxa (Blackburn and Gaston, 1996; Gaston, 2000; Willig et al., 2003).

In this study, only two known invasive species were encountered (*L. incisa* and *Pheidole sp*.). Although these two species were abundant and widely distributed in eThekwini and surrounding areas, their distribution was not correlated to the distance from the city centre and the harbour. It is unclear how the presence of these species effects the abundance of the native species of ants. A recent study suggested that urban areas may actually allow native species that are disappearing from their natural habitats to thrive in an urban setting despite the presence of invasive species (Menke et al., 2010b), but this will need to be studied in more detail in future studies.

Haplotype accumulation curves can highlight the rate of encountering unsampled genetic diversity. The haplotype accumulation curves recovered in this study (Figures 3.3A-H) indicate, for the most part, that there is still much more diversity to be sampled, even in open spaces that cover the smallest area and have the most number of sampled species such as Ipithi, the haplotype accumulation curves did not reach a plateau. The extrapolation measures estimate that as many as 73 more species could still be encountered in addition to the 80 species of ants that have been collected in the present study (Figure 3.4, Chao 2 measure). Species assemblage comparisons indicated that forest and grassland habitats each support unique species assemblages (Figure 3.5 and 3.6a-b). The size of the open space sampled was not a good predictor for species richness as smaller reserves such as Ipithi nature reserve (12 ha) was similar in terms of diversity to a larger reserve such as New Germany (110 ha). The proximity of the open space sampled to the heavily urbanized Durban city centre and harbour also did not explain variation in ant species richness (Figure 3.7).

In general, it appeared that nature reserves/ “green areas” with mainly grassland and forest habitats showed the most species diversity. From a conservation standpoint, this data indicates that these habitat types should be preserved to promote and protect biodiversity within a city.

3.5 Conclusion

It is thought that with the expansion of urbanization, the remaining fauna would be homogenized and there would be a lack of diversity and richness (McKinney, 2006). In this study, it was shown how the open/ “green” spaces within eThekwini and surrounding areas, can assist in maintaining the native diversity of ants in urban environments. Additionally, species assemblages within the open spaces were very diverse, and no two reserves had the same species assemblages. These findings may assist in planning future urban development with taking the biological diversity into consideration.

Chapter Four

Phylogeography and landscape genetics of four species of ants in eThekwini

Abstract

The inability of populations to maintain optimal effective population sizes due to habitat loss in urban areas leads to a decrease in genetic variation, difficulty in adapting to environmental change and increased risk of extinction. This can be prevented by increasing habitat connectivity in urban areas by propagating green areas and creating ecological or "green corridors" which enable gene flow among patches of open space. In order to assess habitat connectivity, the spatial distribution of genetic diversity should be investigated by using methods such as landscape genetics and phylogeography. Phylogenetic, landscape genetics and phylogeographic analyses were performed on co-distributed ant species belonging to the genera of *Lepisiota, Camponotus, Pheidole* and *Pachycondyla* to better understand the spatial distribution of genetic variability in the eThekwini region. The BINs from the genus *Pheidole* were considered to be *Pheidole megacephala* based on morphological observations. Phylogeographical and landscape genetic patterns were compared between the two invasive species, *Lepisiota incisa* and *P. megacephala*. Mitochondrial cytochrome oxidase I data and the nuclear markers 18S and 28S rDNA were also sequenced and compared for a subsample of *C. nr. cintellus*, *Pachycondyla caffraria* and *Pachycondyla havilandi* individuals. There was subtle genetic variation at COI and the nuclear markers for each of the species examined. Each species exhibited unique patterns of genetic variation which implies that the differences in evolutionary and life histories are more likely to have shaped population structure rather than the landscape alone. Nevertheless, the influence of landscape on population genetic structure cannot be ruled out. In order to fully elucidate the population structure patterns which could be expected in eThekwini and surrounding regions, further sampling across more localities is essential. The use of more nuclear markers could also assist in uncovering these unique patterns of genetic variation in an urban setting.

4.1 Introduction

Human activities have a profound effect on biodiversity, especially in urban areas where habitat loss and fragmentation negatively affect many taxa. Without sustainable resources, an area affected by habitat loss will be incapable of maintaining large population sizes (Frankham et al., 2010). This leads to a decrease in genetic variation, and the organisms affected can have difficulty in adapting to environmental changes which leads to an increase in the risk of extinction of vulnerable species or populations of species in a given area (Frankham et al., 2010). Such scenarios could be avoided by increasing habitat quality, area and connectivity through the propagation of green areas within cities and the establishment of nature reserves (Pacheco and Vasconcelos, 2007). Green corridors enable gene flow among disjunct habitat patches and could circumvent the effects of habitat fragmentation and reduce the decline in biodiversity (Menke et al., 2010b; Pacheco and Vasconcelos, 2007; Vergnes et al., 2012).

Another important facet of conservation is the identification of barriers to gene flow, such as landscape features which could possibly inhibit gene flow and lead to genetic structuring of populations. Understanding these barriers are an important consideration when establishing open spaces and linking corridors for conservation within the city. It is important to incorporate both phylogeography and landscape genetics in such investigations. Even though both these areas of study aim to understand the spatial distribution of genetic diversity, landscape genetics provides insight into contemporary processes which shape genetic diversity, while phylogeography aims to understand the ways in which historical processes affected genetic diversity (Avise et al., 1987; Manel et al., 2003; Wang, 2010).

Insects may provide a better indication of barriers to gene flow and the influence of habitat alteration than mammal species because they are very sensitive to ecosystem change (Bolger et al., 2000; Bromilow and Sperling, 2011). Ants are of particular interest because they are an important component of most ecosystems. Several phylogeographic studies have focused on ants. Examples include, *Platythyrea punctata* from the West Indies (Seal et al., 2011), *Myrmica rubra* and *M. ruginodis* across Europe (Leppanen et al., 2013) , *Solenopsis invicta*

in South America and *Oecophylla smaragdina* in Asia (Azuma et al., 2006). An example of a landscape genetics study that focussed on ants was the investigation of landscape features on the army ant, *Eciton burchellii*, which is a neotropical predator (Perez-Espona et al., 2012). Despite this growing body of literature very few studies have focussed on African ant taxa and in particular urban taxa.

Phylogeography seeks to describe the factors responsible for the phylogenetic associations and spatial distribution of species, under the assumption that strong vicariance events are expected to result in congruent genetic signatures among unrelated taxa (Avise et al., 1987). The aim of this study was to compare the landscape genetic and phylogeographical patterns among co-distributed species belonging to the genera *Camponotus*, *Lepisiota*, *Pachycondyla* and *Pheidole* in the eThekwini region. The city of Durban provides an interesting case study to examine the impact of human-mediated habitat fragmentation. The eThekwini municipality has established a network of open spaces such as rural landscapes, riverine and coastal corridors, nature reserves and privately owned land referred to as the Durban Metropolitan Open Space System (DMOSS). Mitochondrial COI barcode data from the Formicidae of eThekwini municipality (Durban, South Africa) barcode library (chapter 2) will be used in landscape and phylogeographical analyses. In addition, the nuclear markers 18S and 28S rDNA were sequenced for a subsample of species belonging to *Camponotus* and *Pachycondyla* in order to compare the phylogeographic and landscape genetic patterns obtained from a multimarker system.

Landscape features alone sometimes aren't responsible for population differentiation as species biology and life history also have an effect (Potter et al., 2012). Life history traits such as a higher reproductive output, high densities and changes in foraging are some of the differences exhibited by species in urban areas, and these traits are often characteristic of invasive species. In this study the patterns of genetic diversity shown by invasive ants *Lepisiota incisa* and BINs from the genus *Pheidole* considered to be *Pheidole megacephala*, to native ants *Camponotus cintellus* and *Pachycondyla caffraria* were compared. *P. megacephala* exhibits unicolonial social structure and ranks amongst the world's most harmful invasive species (Fournier et al., 2012). *Lepisiota incisa* is another invasive species which has been documented in southern Africa. They are highly abundant in homes, gardens

and parks which indicates that they have successfully colonized urban areas (Sithole et al., 2009).

4.2 Materials and methods

4.2.1 Selection of taxa for phylogeographic analysis

Co-distributed specimens of species from the genus, *Lepisiota*, *Pachycondyla* and *Pheidole* were selected for analysis using a phylogeographic approach. These taxa were selected from the DNA barcode library because there were a large number of individuals collected from five or more localities within eThekwini and surrounding areas. Analysis of the COI barcode data indicated that there was genetic structure below the species level, in some cases individuals belonging to the same species were clustered into more than one barcode cluster (Figure 4.1). This genetic structure could be indicative of cryptic speciation, or it could represent phylogeographic structure. In many cases, individuals belonging to different barcode clusters were morphologically indistinguishable even by an expert taxonomist (Dr. Hamish Robertson of Iziko Museum).

Table 4.1 Datasets used for phylogeographic and landscape genetic analyses. Species names in BOLD are those confirmed by a taxonomist. '-' denotes that no match was available.

Species Name	**BIN No.**	**No. of Individuals**	**GenBank BLAST Identification (%)**	**BOLD Identification (%)**
Lepisiota unknown	264251	1	95	96.45
Lepisiota canescens	494813	4	95	95.37
Lepisiota crinitaA	515769	1	91	97.99
Lepisiota crinitaB	264252	8	90	94.44
Lepisiota incisa	264250	64	91	99.85
Camponotus AFRCSA	133387	2	87	99.03

Camponotus eugeniae	133898	5	87	99.69
Camponotus niveosetosus	133278	8	89	98.30
Camponotus petersii	500732	1	86	97.53
Camponotus nr. cintellus	133907	38	87	98.92
Ponera pennsylvanica	498868	3	-	84.74
Pachycondyla etkf	469662	1	90	100
Pachycondyla tarsata	500346	5	-	97.55
Leptogenys etkd	262510	4	84	85.09
Pachycondyla caffrariaA	262511	19	-	98.17
Pachycondyla caffrariaB	483813	2	-	97.71
Pachycondyla caffrariaC	483795	1	-	97.69
Pheidole etkt	507075	5	85	89.91
Pheidole megab	261760	16	85	85.88
Pheidole mg	514931	16	98	98.31
Pheidole megac	261761	20	85	100
Pheidole megad	19139	5	99	99.39

Given the potential for cryptic speciation within these species, analyses were carried out using a tiered approach. Analyses were carried out for datasets consisting of all BINS within a particular morphospecies and were also carried out independently for datasets (BINs) containing individuals which were confirmed to species level by a taxonomist. In total, seven datasets were used, these are listed in Table 4.1. A dataset consisted of all BINs identified on BOLD as belonging to the genus *Lepisiota* (n = 70; BINs 264251, 494813, 515769, 264252, 264250), and a dataset consisting of only individuals confirmed to be *Lepisiota incisa* (n = 63, BIN number = 264250) was analysed. A dataset consisting of individuals from BINs identified on BOLD belonging to the genus *Pachycondyla* and one *Leptogenys* and one *Ponera* (n =36; BINs 262511, 483813, 483795, 469662, 499613, 500324). Even though the individuals from *Ponera* and *Leptogenys* were not from the genus *Pachycondyla*, they were morphologically difficult to differentiate from individuals that did belong to the genus and were thus included in the dataset. A dataset with individuals that were identified as *Pachycondyla caffraria* on BOLD (n = 22; BINs 262511, 483813, 483795) was analysed. A dataset containing all BINs that belonged to the genus *Pheidole* (n = 62; BINs 507075,

261760, 261918, 514931, 532691, 261761, 19139). This group is particularly difficult to diagnose morphologically but we suspect that the individuals from these BINs belong to the species *Pheidole megacephala*. A dataset which contained all BINs identified as the genus *Camponotus* on BOLD (n = 65; BINs 133387, 133907, 133898, 133278, 500732) and a dataset containing individuals that were confirmed to be *Camponotus nr. cintellus* (n = 49, BIN number = 133907) were analysed separately.

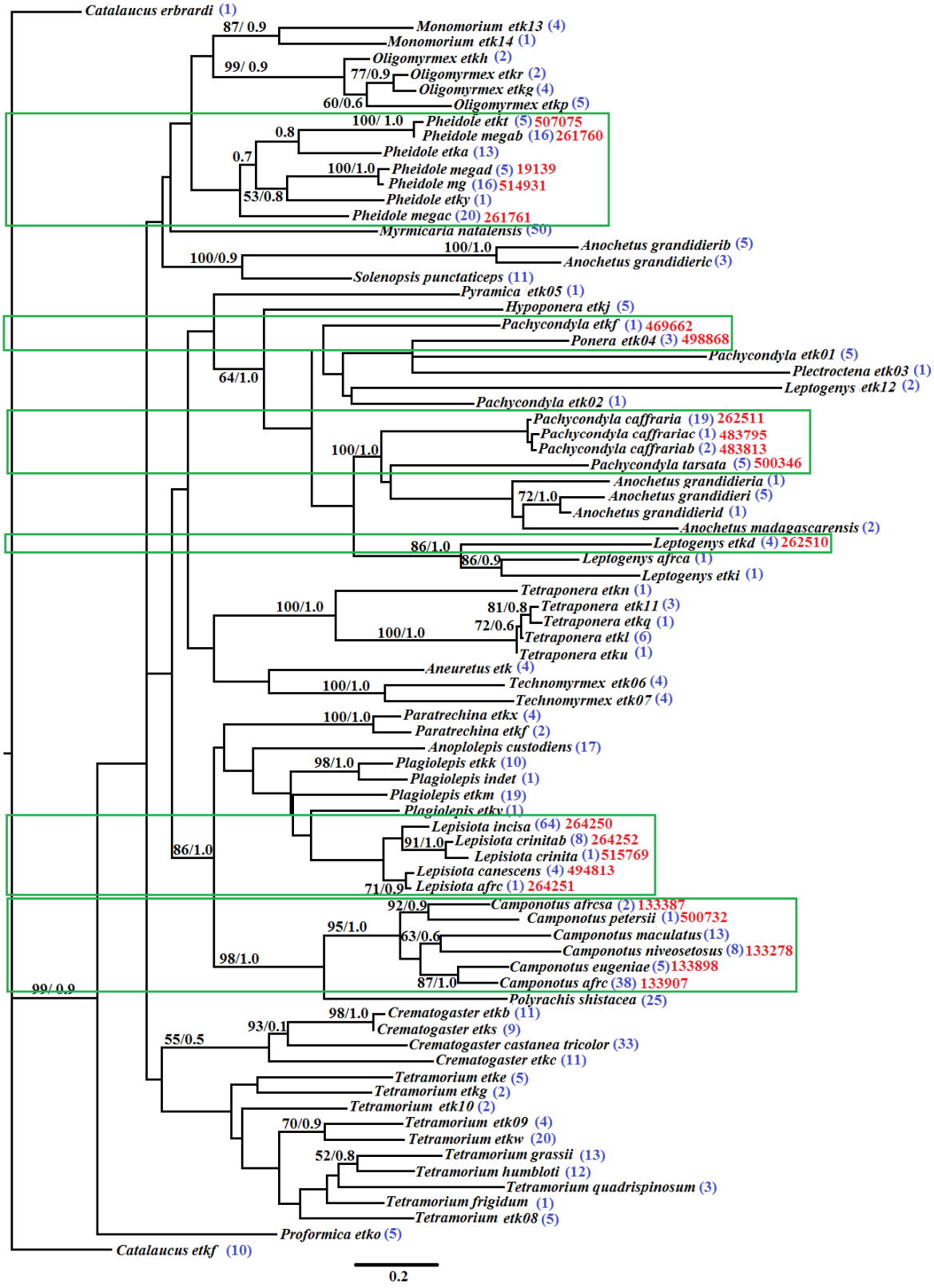


Figure 4.1 A maximum likelihood tree of one individual per BIN. Bootstrap values above 50 and Bayesian posterior probabilities above 0.5 are shown on the branches. The number of individuals per BIN are also indicated. BINs used in this phylogeographic study are highlighted in the green boxes and BIN numbers are also shown.

4.2.2 Subsampling for phylogeographic analysis using a multimarker system

Nuclear sequence data were collected from a subsample of 13 *Camponotus.* individuals (one individual per sampling locality, not barcoded) and 25 individuals of *Pachycondyla* from Palmiet, Msinsi, Happy Valley, Ipithi and Kenneth Stainbank nature reserves and four individuals of *P. havilandi* from Springside nature reserve (from DNA barcode library). DNA was extracted from leg tissue preserved in 99% ethanol. Tissue was crushed with liquid nitrogen and DNA was extracted using the Zymo Insect MiniPrep DNA extraction kit (Zymo Research) following the manufacturer's instructions. The polymerase chain reaction (PCR) was used to amplify the mitochondrial COI gene for the 13 individuals from *Camponotus*, using the primers (Hebert et al., 2004a): LepF1 (5' ATT CAA CCA ATC ATA AAG ATA TTG G 3') and LepR1 (5' TAA ACT TCT GGA TGT CCA AAA AAT CA 3'). The COI data available on BOLD was used for *Pachycondyla*. The amplifications were carried out using the following thermal cycling program: initial denaturation at 95 °C (3 min), and 35 cycles of 95 °C (30 sec), 45 °C (30 sec) and 72 °C (1 min), then a final extension of 72 °C (7 min).

Nuclear genes 18S and 28S ribosomal DNA genes were amplified for both *Pachycondyla* and *Camponotus* individuals using the following primer sequences (Saux et al., 2004); 18S-H17F (5' AAA TTA CCC ACT CCC GGC A 3') and 18S-H35R (5' TGG TGA GGT TTC CCG TGT T 3'), and for 28S, 28S-D2BF (5' GTC GGG TTG CTT GAG AGT GC 3') and 28S-D3AR (5' TCC GTG TTT CAA GAC GGG TC 3'). The PCR amplifications were performed under the following cycling conditions: initial denaturation at 95 °C (3 min), and 35 cycles of 95 °C (30 sec), 55 °C (30 sec) and 72 °C (1 min), then a final extension of 72 °C (7 min). All of the PCR reactions were performed in 25 µl reactions containing 0.1 µl (5U/ µl) DreamTaq (Thermo Scientific), 2 µl (10X) DreamTaq buffer containing 20 mM $MgCl_2$, 0.5 µl of each primer, 0.5 µl DNTPs and 2 µl (1-2 µg) of genomic DNA. A negative control was included with every PCR amplification to check for contamination. All amplified PCR products were verified using electrophoresis on a 1.0% agarose gel stained with ethidium bromide. The products were sequenced at the DNA sequencing facility at the University of Stellenbosch central analytical facility. The heterozygous positions for the nuclear genes were coded using the IUPAC system.

Multiple sequence alignment was carried out using ClustalX/W in BioEdit and thereafter refined manually to ensure homology (Hall, 1999). The nuclear haplotypes were reconstructed using PHASE software (Stephens and Donnelly, 2003) and SeqPhase (Flot, 2010) under the assumption that most ant species are diploid (Imai and Kubota, 1972). Highly variable regions in 18S and 28S that were difficult to align were excluded and trimmed out of the alignments for further analyses.

4.2.3 Molecular diversity indices

Nucleotide diversity (π), haplotype diversity (h), their standard deviations, the number of haplotypes (H), variable and parsimony informative characters, were calculated for each of the datasets using Arlequin 3 (Excoffier et al., 2005) and MEGA 5 (Tamura et al., 2011). The demographic histories of COI for all datasets were compared by calculating Fu's Fs (Fu, 1997) and Tajima's *D* (Tajima, 1989) statistics, were used to test for any departures from the neutral model of evolution as a result of population growth or positive selection (Simonsen et al., 1995). These statistics were computed using the program DnaSP 5 (Librado and Rozas, 2009) and were computed for both the mitochondrial and nuclear datasets.

4.2.4 Phylogenetic analysis

Median-joining haplotype networks were constructed for both the mitochondrial and nuclear datasets using Network 6 (http://www.fluxus-engineering.com). The best-fit models for nucleotide sequence evolution were selected for each dataset using the corrected Akaike information criterion (AICc) in jModeltest 2 (Darriba et al., 2012). Maximum likelihood and Bayesian analyses were performed as outlined in chapter 2, page 38.

4.2.5 Landscape genetics

To compare and contrast the patterns of genetic diversity over geographic space for each of the datasets, genetic landscape shape interpolation plots were constructed using Alleles in Space (Miller, 2005) as outlined in chapter 1, materials and methods. A spatial Bayesian

clustering algorithm in BAPS 5.2 (Corander and Marttinen, 2006; Corander et al., 2003) was used to investigate population structure in each of the datasets. The analyses were also performed for individuals of *C. nr. cintellus* and *Pachycondyla sp.* that were successfully sequenced for nuclear markers, 18S and 28S as well as for the combined datasets (COI + 18S +28S). Analysis of molecular variance (AMOVA) was used to partition the diversity of COI within and amongst sampling localities for all datasets and 18S and 28S for *C. nr cintellus* and *P. caffraria*. The four groupings investigated were populations from reserves with grassland and forest habitats (Ipithi, New Germany, Springside, Kenneth Stainbank, Msinsi, Palmiet, Amatikulu) grassland only (Phinda, Vernon Crookes, North Park, Darville Resources Park), forest only (Silverglen, Seaton Park, Krantzkloof, Happy Valley) and urban habitats (UKZN Westville, UKZN Agric, Port Shepstone urban garden, Isipingo beach, Treasure Beach, Beachwood Mangroves).

4.3 Results

4.3.1 Comparison of COI between taxa and multimarker data for *Camponotus* and *Pachycondyla*

The 658 bp of COI sequence data revealed that at the genus level, *Pachycondyla* was the most variable in terms of haplotype and nucleotide diversity ($h = 0.90 \pm 0.02$, $\pi = 0.12 \pm 0.02$; Table 4.2), followed by individuals from *Camponotus* ($h = 0.80 \pm 0.04$, $\pi = 0.07 \pm 0.02$; Table 4.2). The genus *Pheidole* had a haplotype diversity of 0.79 ± 0.03 and a nucleotide diversity of 0.11 ± 0.005 (Table 4.2). The least diverse genus was *Lepisiota*, with a haplotype diversity of 0.71 ± 0.05 and nucleotide diversity of 0.01 ± 0.006 (Table 4.2). As expected, the molecular diversity indices were lower for species in the genus *Lepisiota* and *Pheidole*, as these species are invasive. This is attributed to a loss of genetic diversity due to founder effects which are sometimes experienced by introduced species (Dlugosch and Parker, 2008). At the species level, *P. caffraria* had the highest haplotype and nucleotide diversity ($h = 0.80 \pm 0.05$, $\pi = 0.008 \pm 0.003$; Table 4.2) followed by *C. nr. cintellus* ($h = 0.66 \pm 0.05$, $\pi = 0.002 \pm 0.0003$; Table 4.2) and *L. incisa* ($h = 0.65 \pm 0.05$, $\pi = 0.002 \pm 0.0002$; Table 4.2). Significantly negative values for Tajima's D and Fu's Fs were only obtained for the dataset consisting of all individuals from *Camponotus* (Table 4.2) and for *Pheidole*. These negative values are indicative of a recent expansion in population size (Simonsen et al., 1995).

A total of 10 18S (793 bp) and 12 28S (363 bp) sequences were obtained for *Camponotus.* A total of 15 18S (731 bp) and 28 28S (496 bp) individuals of *Pachycondyla* were successfully sequenced. The concatenated dataset consisting of COI, 18S and 28S yielded 1770 bp for analysis of *Camponotus* and 1885 bp for analysis of *Pachycondyla.*

In both *Camponotus* and *Pachycondyla*, 18S was more variable than 28S. The haplotype and nucleotide diversities of 18S were also higher than 28S (Table 4.2). Negative values for Fu's Fs and Tajima's D were obtained for *Camponotus* (18S), *Pachycondyla* (18S and 28S), however, these values were not statistically significant and therefore do not point to a recent expansion in population size. The combined dataset (COI+18S+28S) indicated that there was a significant recent expansion in population size for *Camponotus* but not for *Pachycondyla.*

Table 4.2. Molecular diversity indices for the COI and nuclear (18S and 28S) datasets and combined (COI+18S+28S) datasets. n – number of individuals, L – number of localities individuals were collected from, V – variable sites, PI – parsimony informative sites, h = haplotype diversity, SD – standard deviation, π – nucleotide diversity. The values in bold indicate statistically significant results.

	n	**L**	**V**	**PI**	***h***	**SD**	***π***	**SD**	**Fu's Fs**	**Tajima's D**
Lepisiota	70	14	114	91	0.71	0.05	0.01	0.006	15.95	-1.65
L. incisa	63	14	4	4	0.65	0.05	0.002	0.0002	11.64	0.905
Pachycondyla	36	8	275	230	0.90	0.02	0.12	0.02	28.43	-0.30
P. caffraria	22	5	31	15	0.80	0.05	0.008	0.003	2.32	-1.36
Pheidole	53	15	20	20	0.79	0.03	0.11	0.005	12.01	2.17
Camponotus	66	13	317	105	0.80	0.04	0.07	0.02	16.72	-2.24
C. nr. cintellus	49	10	6	4	0.66	0.05	0.002	0.0003	-3.147	-0.71
Camponotus (18S)	24	12	76	58	0.97	0.02	0.02	0.002	-8.48	-1.22
Camponotus (28S)	20	10	15	15	0.89	0.04	0.01	0.001	0.59	0.75
Camponotus (18S+28S+COI)	12	12	500	38	1.00	0.03	0.06	0.04	-0.10	-2.21
Pachycondyla (18S)	30	6	101	66	0.97	0.02	0.02	0.004	-7.70	-1.90
Pachycondyla (28S)	56	7	79	65	0.91	0.03	0.02	0.002	-4.85	-1.07
Pachycondyla (18S+28S+COI)	28	7	406	260	0.97	0.02	0.08	0.01	2.92	-0.49

The AMOVA results for *Lepisiota* and *L. incisa* indicate that most of the genetic variation is within populations (Figure 4.2). This was also true for *P. caffraria*, *Camponotus sp.* and *C.nr. cintellus*, however, the result was not statistically significant for *C.nr. cintellus*. In *Pachycondyla* and *Pheidole*, most of the genetic variation was found among populations within groups. Among-group genetic variation was only observed for *Lepisiota*, *L. incisa* and *Pachycondyla* (Figure 4.2). This suggests that the vegetation-type grouping investigated do not play a significant role in the genetic structuring of ant species.

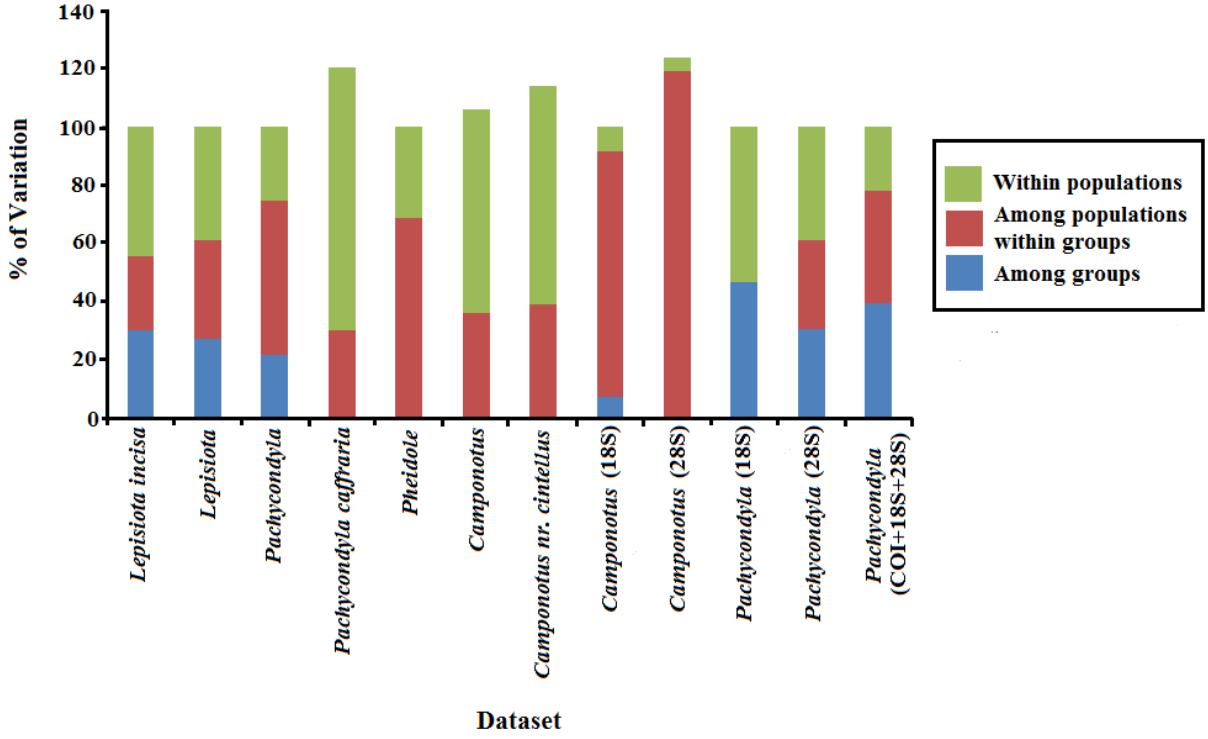


Figure 4.2 AMOVA analysis assessing the hierarchical genetic structure of the seven datasets for COI and nuclear genes 18S and 28S, based on vegetation type.

The nuclear AMOVA results for *Camponotus* (18 S and 28S) highlighted that most of the genetic variation was partitioned among populations within groups for both these markers (Figure 4.2). The AMOVA for the combined dataset for *Camponotus* could not be computed because of missing data for one or more localities for each of the markers. For *Pachycondyla* (18S and 28S), most of the genetic variation was found within populations. However, the

combined dataset for *Pachycondyla* showed that most of the genetic variation was found among groups but this result was not statistically significant (Figure 4.2, Table 4.3).

Table 4.3. Partitioning of genetic variation (analysis of molecular variance, AMOVA) among groups (AG), among populations within groups (APWG) and within populations (WP) for the ant taxa. Values in bold indicate results that are not statistically significant.

	Variance Component			% of variation		
	AG	APWG	WP	AG	APWG	WP
Lepisiota	0.299	0.364	0.554	29.95	25.48	44.58
L. incisa	0.269	0.464	0.609	26.99	33.89	39.12
Pachycondyla	**0.214**	0.677	0.746	**21.44**	53.17	25.39
P. caffraria	**-0.202**	0.250	0.098	**-20.20**	30.00	90.19
Pheidole	**-0.000**	0.686	0.686	**-0.030**	68.67	31.37
Camponotus	**-0.059**	**0.339**	0.300	**-5.930**	**35.95**	69.98
C. nr. cintellus	**-0.142**	**0.339**	**0.245**	**-14.17**	**38.70**	**75.47**
Camponotus (18S)	**0.069**	0.090	0.915	**6.86**	84.66	8.48
Camponotus (28S)	**-0.238**	0.961	0.952	**-23.81**	119	4.81
Pachycondyla (18S)	-	-	0.462	46.25	-	53.75
Pachycondyla (28S)	0.303	0.439	0.609	30.27	30.60	39.14
Pachycondyla (COI+18S+28S)	**0.391**	0.639	0.780	**39.13**	38.87	22.00

4.3.2 COI and nuclear phylogenies and haplotype networks

Different models of nucleotide substitution were selected for each of the datasets (Table 4.4). The simplest model of nucleotide substitution chosen was the F81 model for the 18S marker for *Camponotus* followed by the HKY model chosen for the COI marker for *L. incisa*. The maximum likelihood and Bayesian trees were congruent for all of the taxa and therefore Bayesian posterior probabilities could be placed on the most likely tree along with bootstrap support from maximum likelihood analysis.

Table 4.4. Best-fit models for COI and nuclear DNA (18S and 28S) selected by the AICc in jModeltest 2.

	Model	Log likelihood (-lnL)	Nucleotide frequencies			
			A	C	G	T
Lepisiota	TIM1+I	1461.08	0.2954	0.1905	0.1136	0.4004
L. incisa	HKY	884.089	0.2894	0.2046	0.1161	0.3900
Pachycondyla	TIM2+I	2788.91	0.3039	0.1705	0.1152	0.4105
P. caffraria	TPM1uf	1046.05	0.3021	0.1783	0.1241	0.3955
Pheidole	TIM+I+G	2937.82	0.3058	0.1883	0.1201	0.3858
Camponotus	TIM2+I+G	3014.99	0.2952	0.1847	0.1174	0.4028
C. nr. cintellus	TrN+I	1038.42	0.2980	0.1659	0.1267	0.4094
Camponotus (18S)	TIM+G	1791.00	0.2468	0.2132	0.2932	0.2468
Camponotus (28S)	F81	597.48	0.2137	0.2914	0.3265	0.1684
Pachycondyla (18S)	TIM3ef+G	1992.32	-	-	-	-
Pachycondyla (28S)	HKY+G	1387.28	0.1891	0.3063	0.3250	0.1795

For *Lepisiota*, there were eight clades recovered on the maximum likelihood and Bayesian trees. Although these clades were quite distinct from each other, there was no variation within each clade. This striking pattern was also recovered in the median-joining haplotype network with each of the eight COI haplotypes separated by a large number of mutational steps. Clade A consisted of individuals from the BIN 264250 and consisted of a single haplotype shared by all individuals from nine localities. The second most well-represented haplotype and clades were F and G which consisted of individuals from four and five localities respectively. Haplotype E and clade E consisted of individuals only from Krantzkloof and Seaton Park nature reserves. Although there is no strong signature of geographically correlated structure, individuals from Vernon Crookes all clustered together with a few individuals from New Germany and UKZN Agric (Figure 4.3).

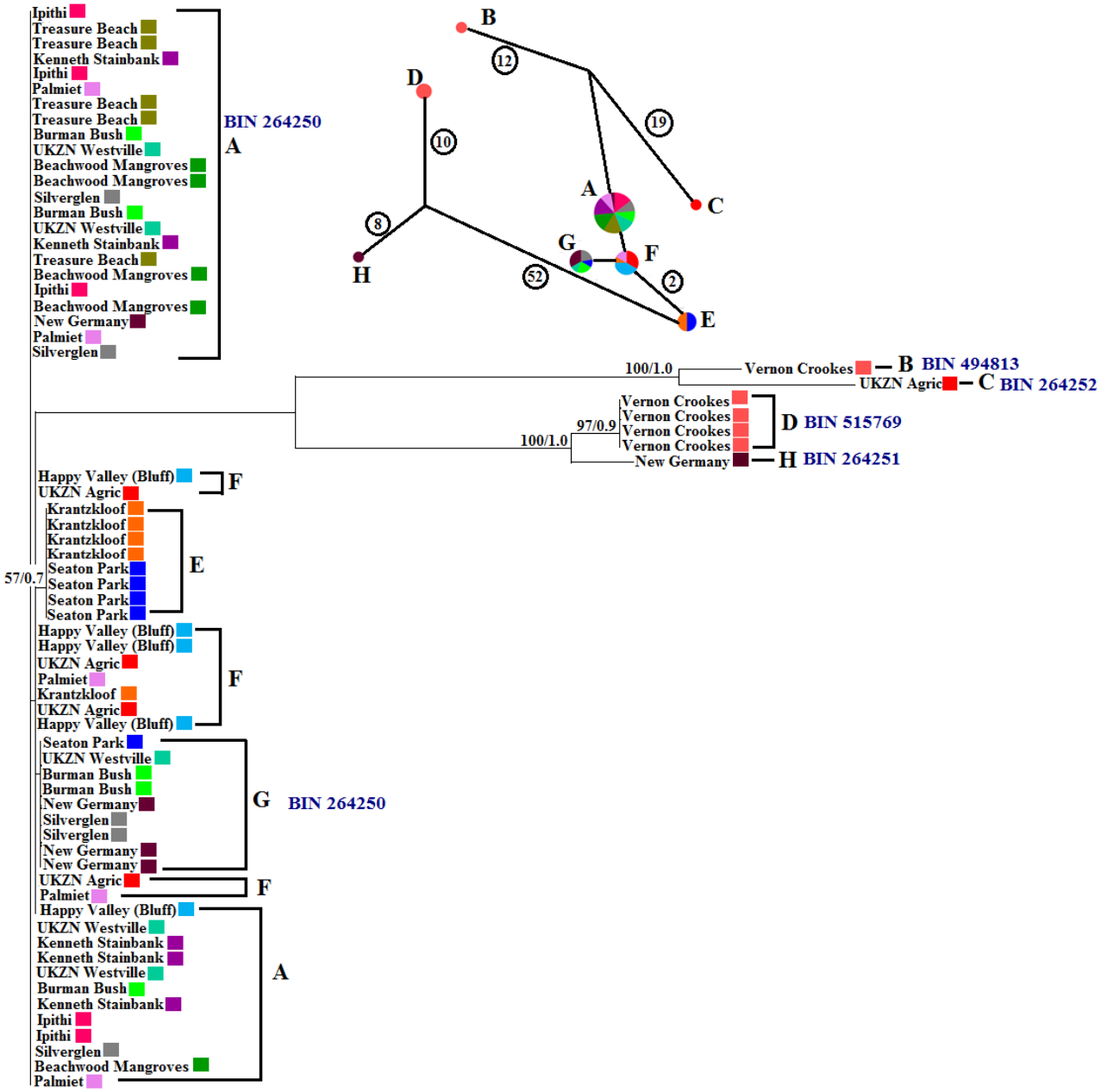


Figure 4.3 Maximum likelihood tree and median-joining haplotype network for *Lepisiota* for COI sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype. The groupings are also indicated.

The maximum likelihood and Bayesian topology for *L. incisa* revealed that there were four clades, again within each of these clades there was no variation and they corresponded to the four haplotypes in the median-joining haplotype network. This species was found to be monophyletic. Just as for *Lepisiota*, haplotype A contained the most individuals, followed by haplotypes B and D. Clade C and haplotype C only consisted of individuals from Krantzkloof and Seaton Park nature reserves (Figure 4.4).

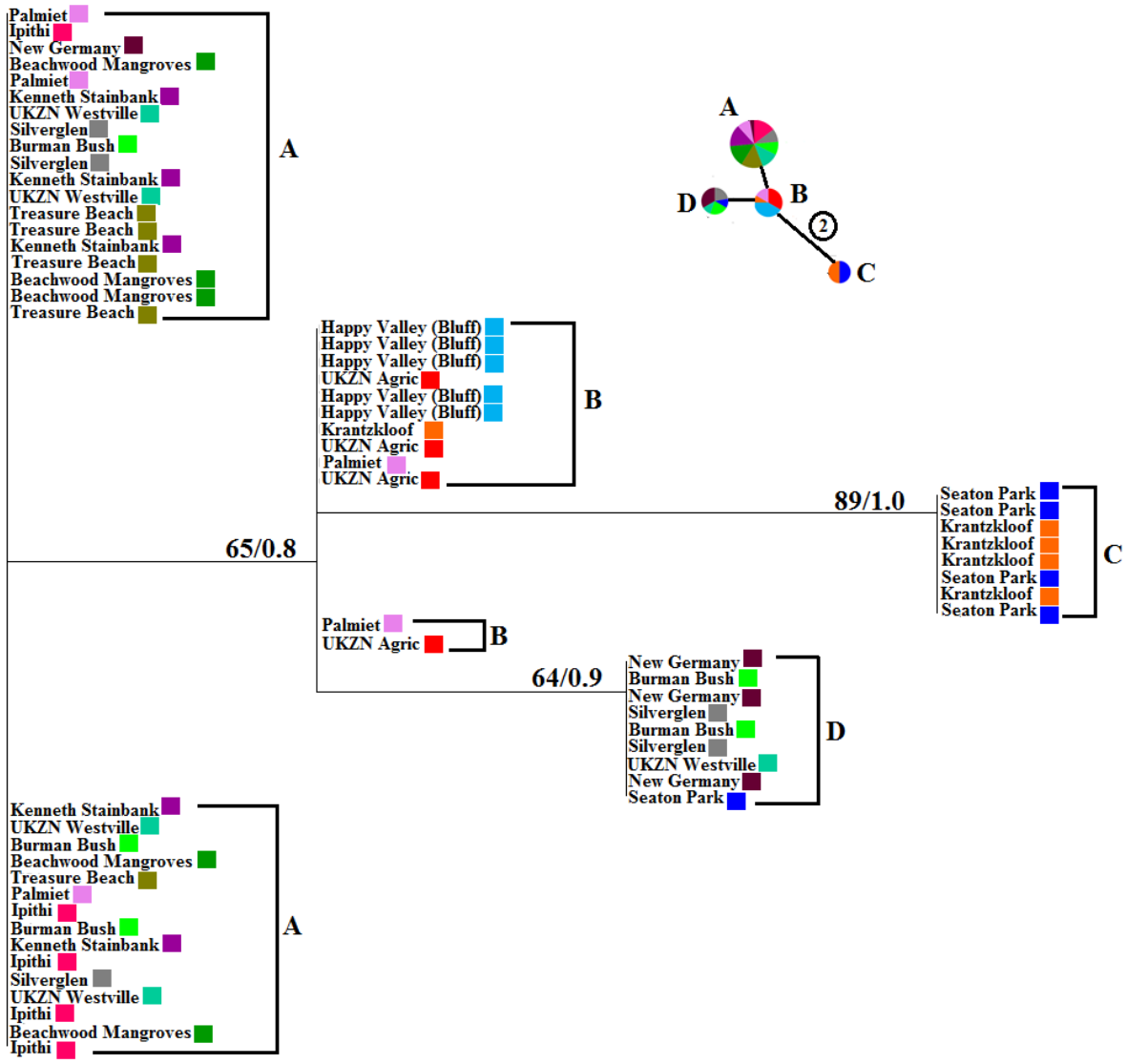


Figure 4.4. Maximum likelihood tree and median-joining haplotype network for *L. incisa* for COI sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype. The groupings are also indicated.

As *L. incisa* is an invasive species, these patterns were expected. There were only four haplotypes and all were shared amongst the reserves. There were no haplotypes unique to a reserve which highlighted the limited genetic variation characteristic of invasive species (Dlugosch and Parker, 2008).

The maximum likelihood and Bayesian topology for *Pheidole* recovered seven distinct clades. There was high bootstrap and posterior probability support for most of the clades and there was some geographic structuring observed. Individuals from Silverglen nature reserve grouped together with high branch support (78, 0.9, Figure 4.5), as did individuals from Beachwood Mangroves (87, 1.0) nature reserve. Individuals from Krantzkloof grouped together but there was lack of branch support (Figure 4.5). There were 23 haplotypes in the median-joining haplotype network. The haplotype groupings are indicated on the maximum likelihood tree. Haplogroup A and clade A consisted of individuals from four localities and all individuals from this clade belonged to BIN 514391. Haplotype D and clade D consisted of individuals from two localities, Ipithi nature reserve and Isipingo Beach. This clade consisted of individuals from BIN 261761. Clade C, with individuals from Krantzkloof also fell under BIN 261761. Haplotype G and clade G consisted of individuals from Springside and New Germany nature reserves. Haplotype H and clade H contained individuals from Burman Bush and Seaton Park. Individuals from clade G and H were from the BIN 261760. There were four haplotypes unique to Darville Resources Park, three unique to Beachwood Mangroves nature reserve, two unique to Springside nature reserve and one each unique to Ipithi, Krantzkloof, New Germany and Silverglen nature reserves (Figure 4.5).

Even though all the individuals from different BINs are grouped under the genus *Pheidole*, morphologically they all resembled *P. megacephala* which is a known invasive species (Fournier et al., 2012). Despite *P. megacephala* being an invasive species, there was a high degree of genetic variation, as well as many private haplotypes. This could be indicative of multiple introductions (Dlugosch and Parker, 2008; Tsutsui and Case, 2001) or cryptic speciation.

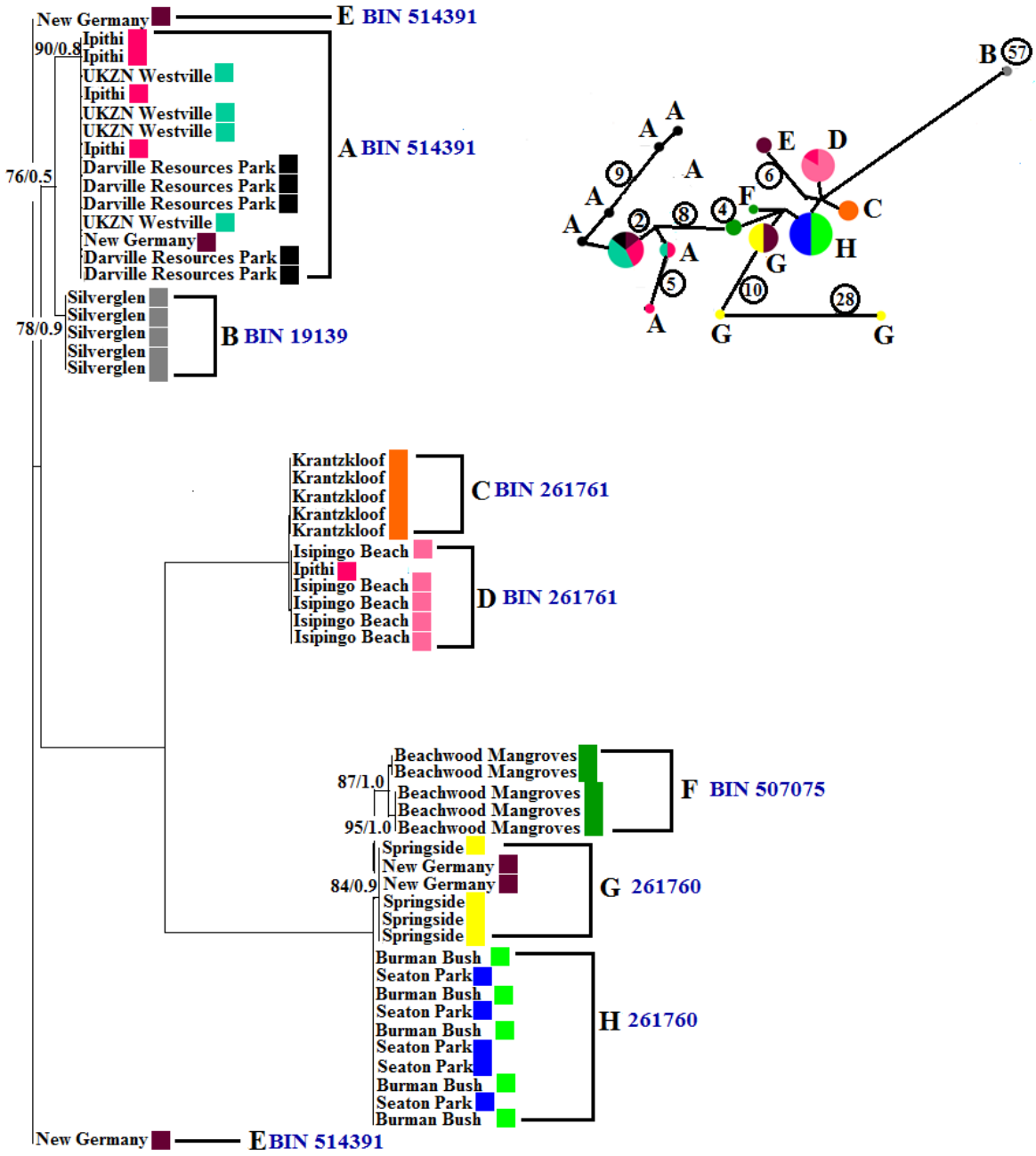


Figure 4.5. Maximum likelihood tree and median-joining haplotype network for *Pheidole* for COI sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype. The groupings are also indicated.

For *Pachycondyla*, the maximum likelihood topology revealed that there was some evidence of geographic structuring. Individuals from Phinda game reserve grouped together in one clade with high bootstrap and posterior probability support (100, 1.0, Figure 4.6). Individuals from Amatikulu nature reserve also grouped together with high branch support (100, 1.0, Figure 4.6). There were a total of 12 haplotypes in the median-joining haplotype network which corresponded to the clades on the maximum likelihood tree. Haplogroup A, and clade A consisted of individuals from four localities and were from BIN 262511. Haplotype J and clade J consisted of individuals from three localities and also found under BIN 262511. Haplotype F and clade F consisted of individuals from two localities and BIN 262510. There were three haplotypes unique to Ipithi nature reserve, two unique to Amatikulu nature reserve, one unique to Phinda game reserve and one unique to Kenneth Stainbank nature reserve (Figure 4.6).

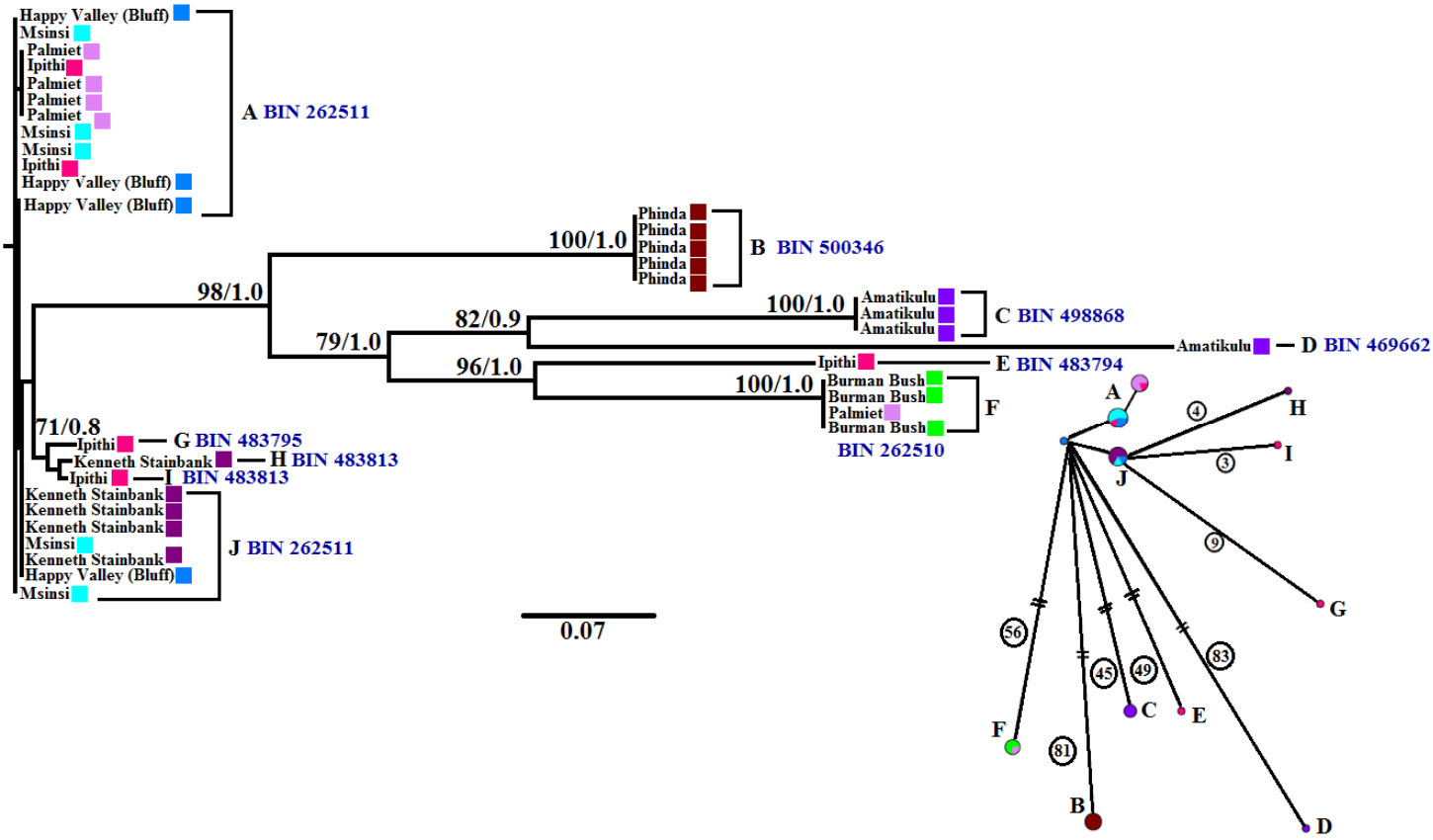


Figure 4.6. Maximum likelihood tree and median-joining haplotype network for *Pachycondyla* for COI sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype. The groupings are also indicated. Dashed lines indicate that lines are not drawn to scale.

The maximum likelihood topology for *P. caffraria* identified five major clades. There was no clear indication of geographic structuring. There were a total of seven haplotypes in the median-joining haplotype network. The haplotypes are indicated on the maximum likelihood tree (Figure 4.7). Haplotype A and clade A consisted of the most number of individuals from three localities. Haplotype G and clade G had individuals from three localities. Haplotype C and clade C had individuals from two localities (Figure 4.7). There were two haplotypes that were unique to Ipithi nature reserve, one unique to Kenneth Stainbank nature reserve and one unique to Happy Valley (Bluff) nature reserve. There were three BINs within this species, 262511, 483795 and 483813.

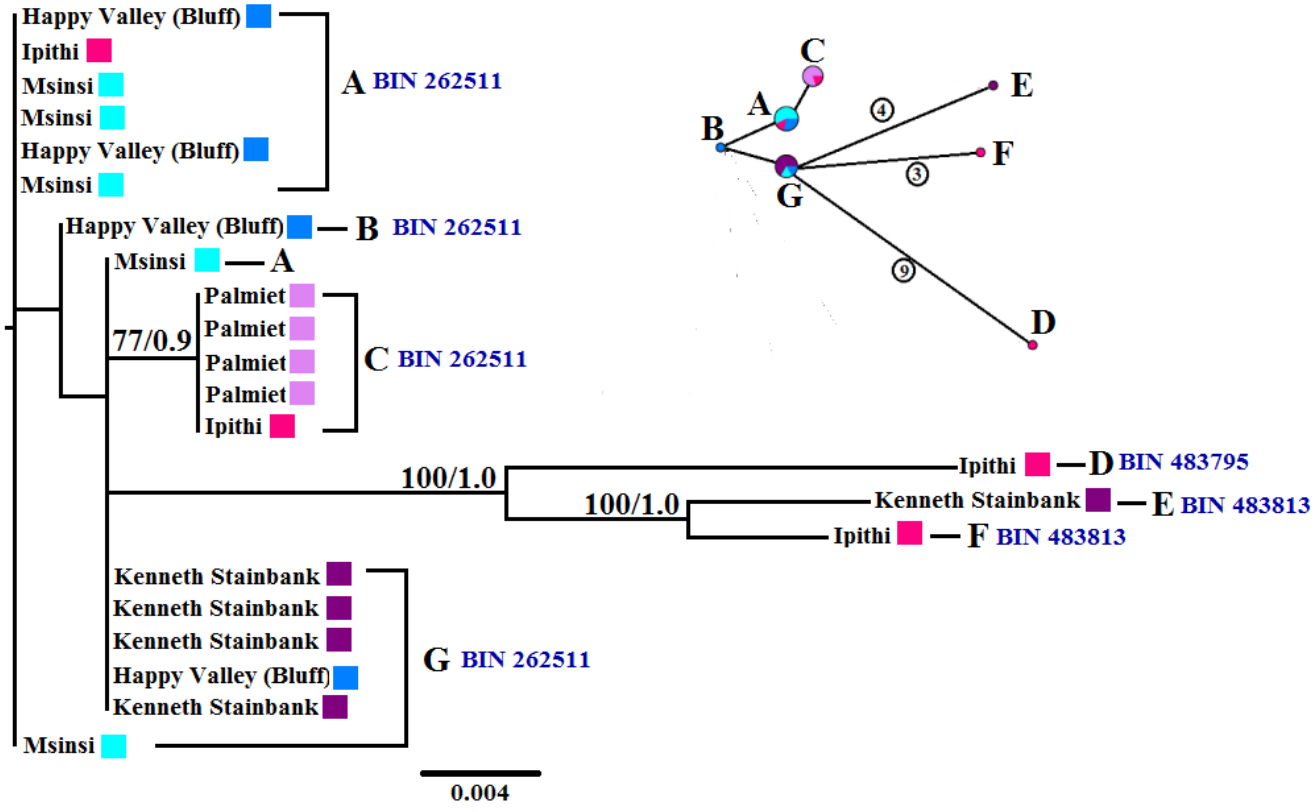


Figure 4.7. Maximum likelihood tree and median-joining haplotype network for *P. caffraria* for COI sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype. The groupings are also indicated.

The maximum likelihood phylogeny for *Pachycondyla sp*. for 18S grouped together the individuals of *P. havilandi* from Springside nature reserve with high bootstrap probability support of 87 and posterior probability support of 0.9. Individuals from Msinsi nature reserve also grouped together with high bootstrap probability support and posterior support of 93 and 0.9 respectively (Figure 4.8). The median-joining haplotype network for *Pachycondyla sp*. for 18S revealed a starburst pattern. All haplotypes were unique and there were no haplotypes shared between localities (Figure 4.8).

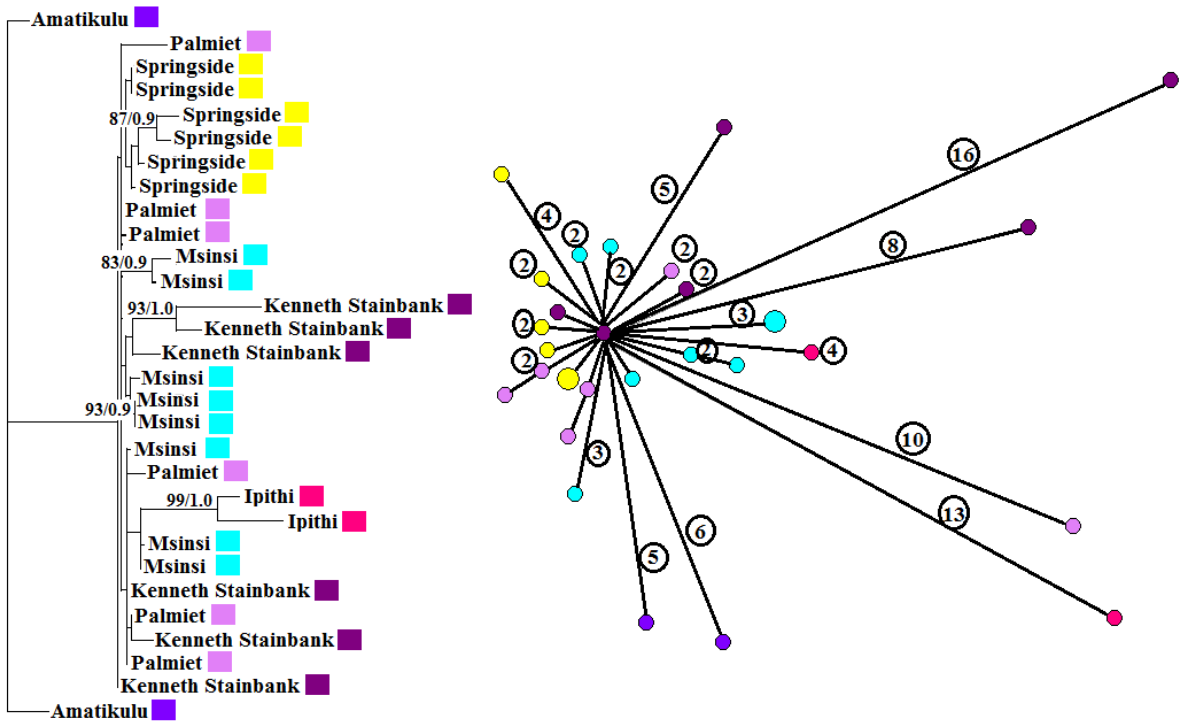


Figure 4.8. Maximum likelihood tree and median-joining haplotype network for *Pachycondyla sp*. for 18S sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype.

The maximum likelihood phylogeny and the median-joining haplotype network for *Pachycondyla sp*. for 28S revealed more structure than 18S. Individuals of *P. havilandi* from Springside nature reserve grouped together with *Pachycondyla sp*. from Amatikulu nature reserve with high bootstrap and posterior probability support of 100 and 1.0 respectively. Individuals from Phinda game reserve grouped together with high branch support (74, 0.9, Figure 4.9). Even though many of the 28S haplotypes were unique to each locality, there were three haplotypes that were shared between localities.

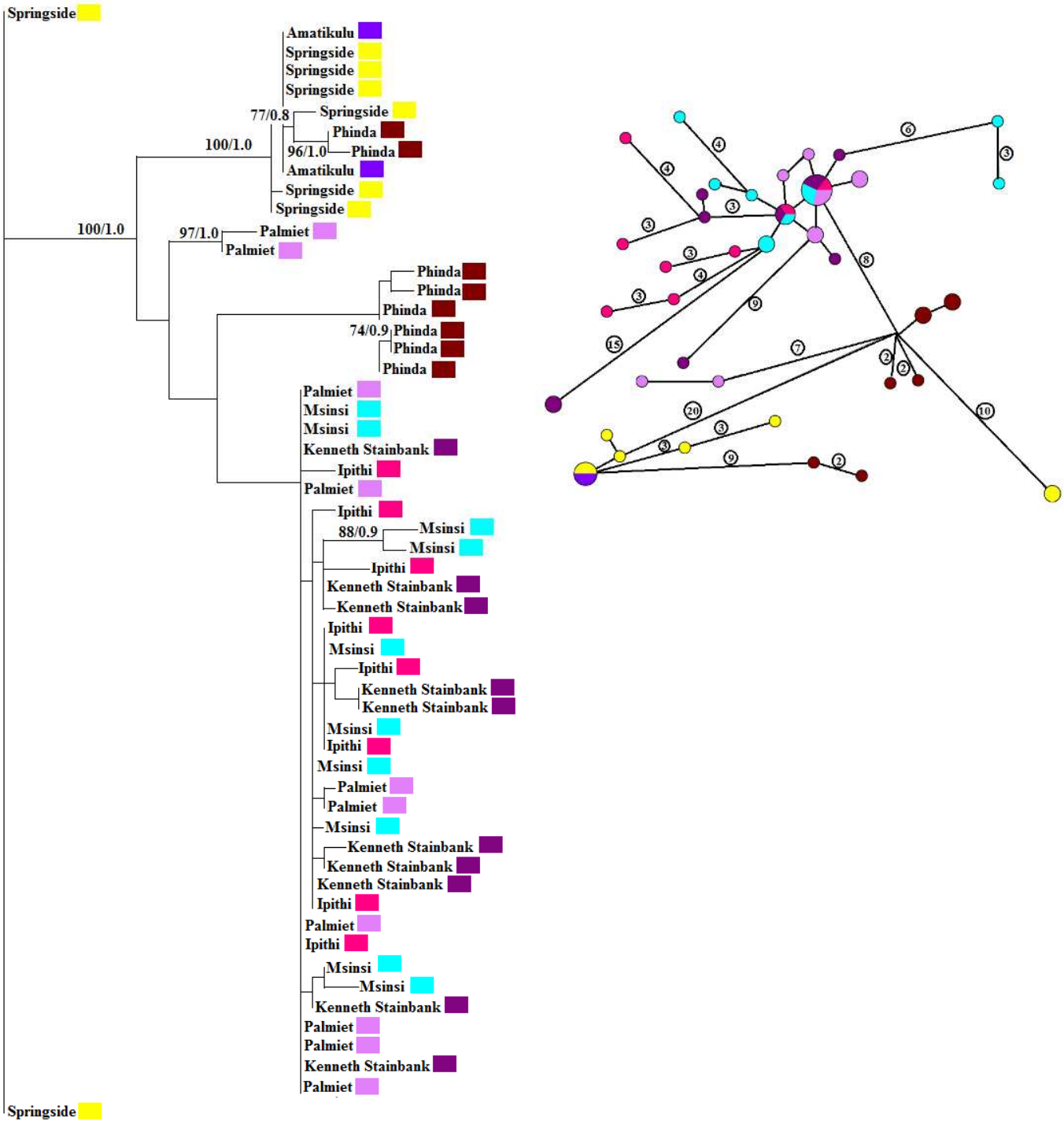


Figure 4.9. Maximum likelihood tree and median-joining haplotype network for *Pachycondyla sp*. for 28S sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype.

One haplotype was shared between Springside and Amatikulu nature reserves, one haplotype was shared between Msinsi, Palmiet, Kenneth Stainbank and Ipithi nature reserves and one was shared between Ipithi, Msinsi and Kenneth Stainbank nature reserves (Figure 4.9). The 28S maximum likelihood phylogeny and haplotype network revealed some similarities to the COI phylogeny. Individuals from Phinda game reserve were also recovered as a single clade for COI and COI haplotypes were also shared between Msinsi, Ipithi and Kenneth Stainbank (Figure 4.6). These findings were further confirmed by the combined maximum likelihood phylogeny for *Pachycondyla sp*. (Figure 4.10).

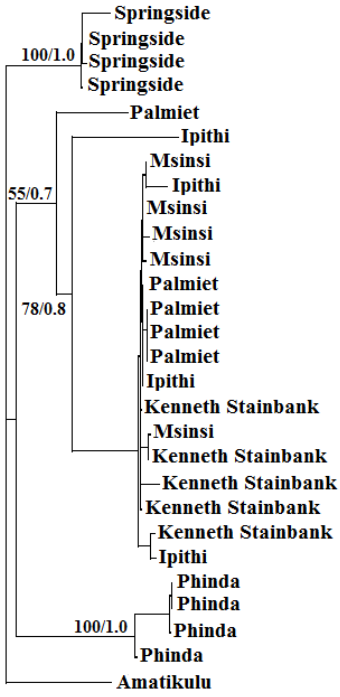


Figure 4.10. Maximum likelihood tree for *Pachycondyla sp*. for the combined dataset (COI+18S+28S). Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown.

The maximum likelihood topology for *Camponotus* revealed some geographic structuring as individuals from New Germany, Springside, Vernon Crookes and Port Shepstone each grouped together. The clades for New Germany, Vernon Crookes and Port Shepstone received high bootstrap and posterior probability support (Figure 4.11). New Germany had a bootstrap probability support of 100 and a posterior probability support of 1.0. Vernon Crookes had a bootstrap probability support and posterior probability support of 85 and 0.8 respectively. The bootstrap probability and posterior probability values for Port Shepstone were 70 and 0.9 respectively. There were 20 haplotypes in the median-joining haplotype network. There were many more unique haplotypes than shared haplotypes. Springside nature reserve had seven unique haplotypes, Palmiet and Vernon Crookes nature reserves had three unique haplotypes each. New Germany, Ipithi, Port Shepstone, Happy Valley (Bluff), Hazelmere Dam, Burman Bush and UKZN Agric had one unique haplotype each (Figure 4.11).

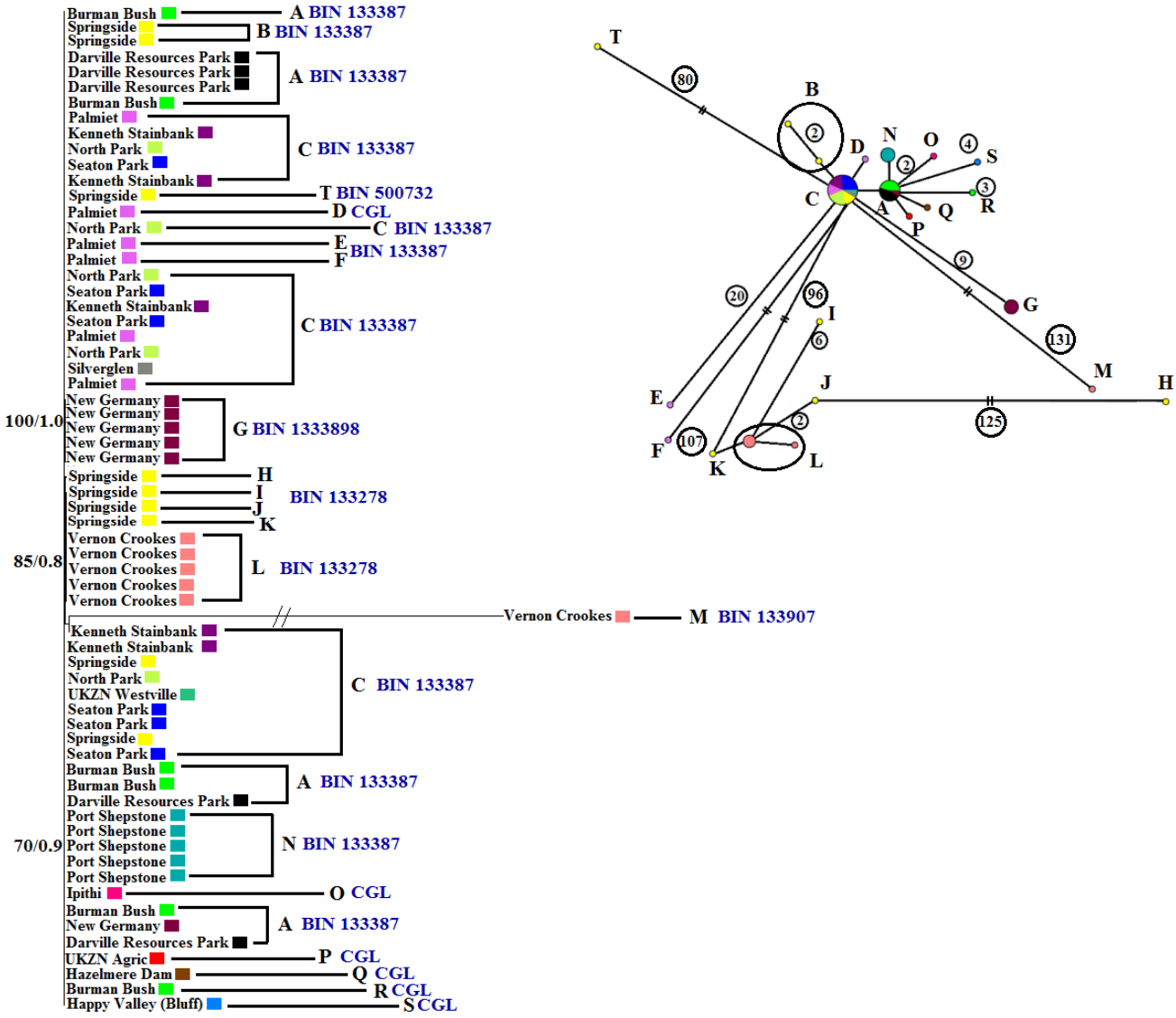


Figure 4.11. Maximum likelihood tree and median-joining haplotype network for *Camponotus* for COI sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype. The groupings are also indicated. Dashed lines indicate that lines/branches are not drawn to scale. CGL indicates that these sequences were generated in the conservation genetics lab at the University of KwaZulu-Natal in Pietermaritzburg.

The maximum likelihood topology for *C. nr. cintellus* consisted of five major clades, however, these did not receive strong support from bootstrap values or posterior probabilities (Figure 4.12). There were 16 haplotypes in the median-joining haplotype network and these are also indicated on the maximum likelihood topology. There were only three shared haplotypes; A, B and C. Haplotype B and clade B was only shared between individuals from Burman Bush and Darville Resources Park. Palmiet nature reserve had three unique haplotypes, Springside nature reserve had two unique haplotypes. Port Shepstone, Kenneth Stainbank, Happy Valley (Bluff), Burman Bush, Hazelmere Dam, Ipithi and UKZN Agric all had one unique haplotype each (Figure 4.12). All individuals were from the BIN 133389 apart from individuals that were sequenced by the CGL.

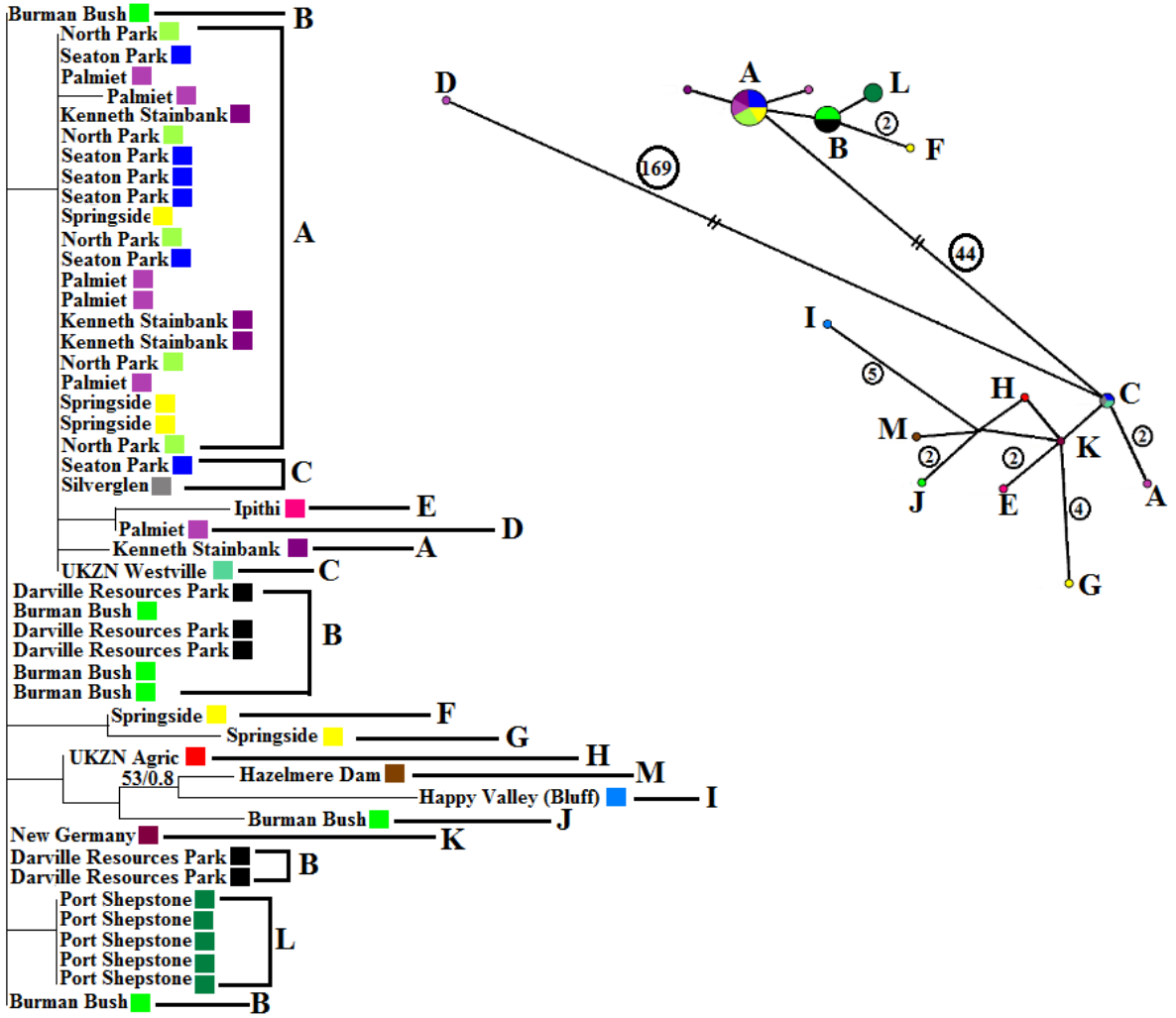


Figure 4.12. Maximum likelihood tree and median-joining haplotype network for *C. nr. cintellus* for COI sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype. The groupings are also indicated. Dashed lines indicate that lines are not drawn to scale.

The maximum likelihood phylogeny for 18S for *C. nr. cintellus* showed high bootstrap and posterior probability support for many of the branches. The median-joining haplotype network indicated that 18S was very variable and each locality had its own unique haplotype. The number of mutational steps ranged from a single mutation to as many as 29 mutational steps (Figure 4.13).

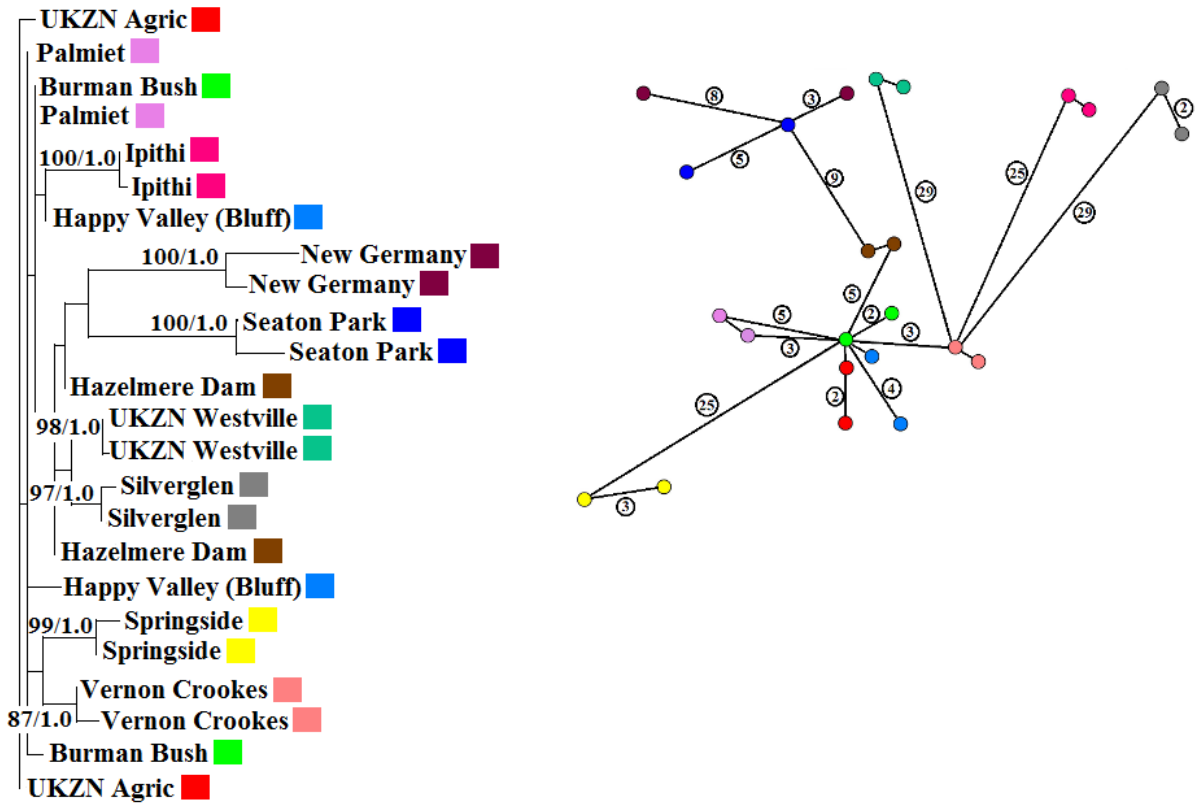


Figure 4.13. Maximum likelihood tree and median-joining haplotype network for *C. nr. cintellus* for 18S sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype.

The maximum likelihood phylogeny for 28S for *C. nr. cintellus* also showed high bootstrap and posterior probabilities for many of the branches. Although there were many unique haplotypes, the median-joining haplotype network showed more structure than the network for 18S (Figure 4.14). The number of mutational steps between haplotypes was less than those for 18S. There was also one haplotype shared between Seaton Park and New Germany. Because only one individual of *C. nr. cintellus* per locality was sequenced for 18S and 28S,

these phylogenies could not accurately be compared to the COI phylogeny. However, localities that contained unique COI haplotypes also contained unique 18S and 28S haplotypes (Figure 4.11 and 4.14).

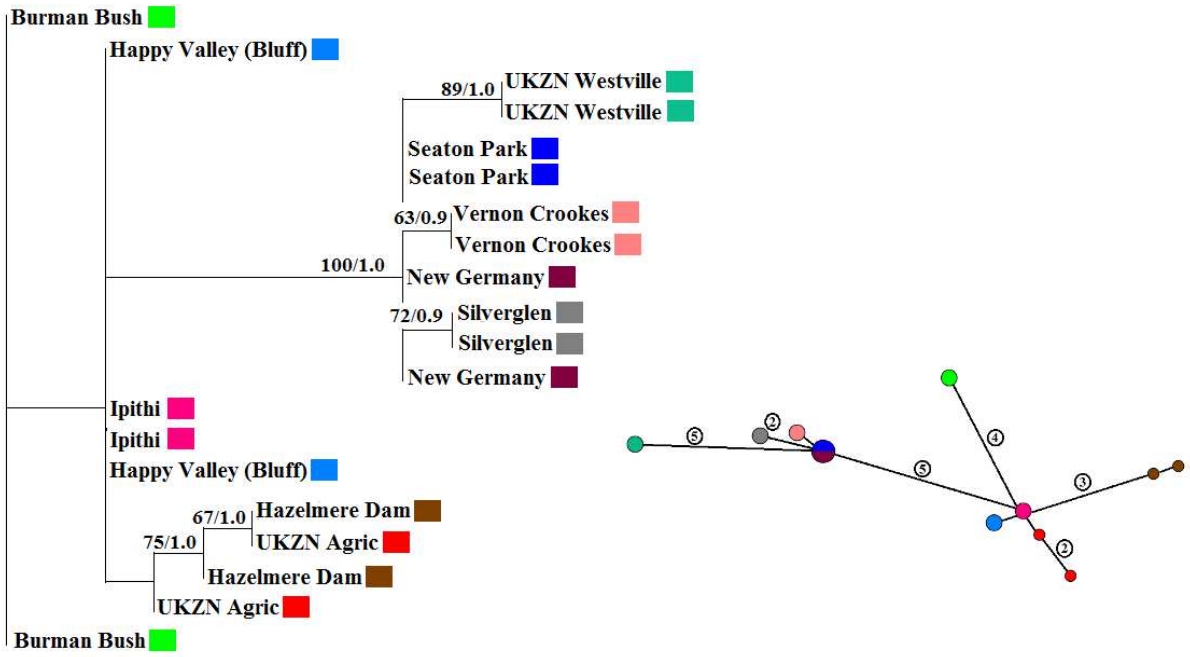


Figure 4.14. Maximum likelihood tree and median-joining haplotype network for *C. nr. cintellus* for 28S sequence data. Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown. The number of mutations on the haplotype network are shown and each circle is proportional to the number of individuals sharing that haplotype.

For the combined dataset (COI+18S+28S) for *C. nr cintellus*, none of the localities grouped together. The topology also received no bootstrap or posterior probability support > 50%/0.5 (Figure 4.15).

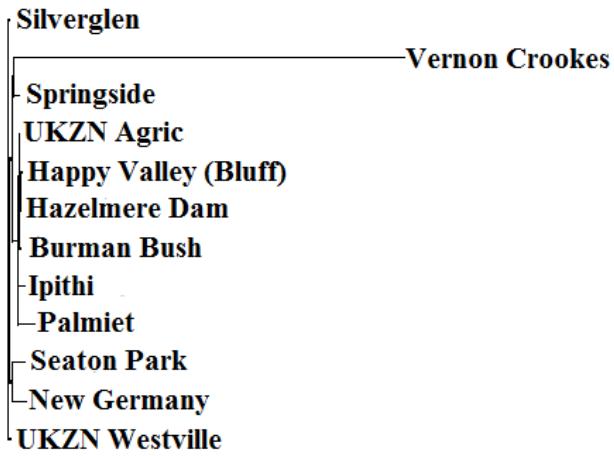


Figure 4.15. Maximum likelihood tree for *C. nr. cintellus* for the combined dataset (COI+18S+28S). Bootstrap values above 50% and Bayesian posterior probabilities over 0.5 are shown.

4.3.3 Comparison of spatial patterns of genetic variation for COI and nuclear data

Mantel tests were used to examine the statistical correlation between genetic and geographic distance. The Mantel test provides a useful indication of whether isolation by distance (IBD) or if landscape features are responsible for genetic differentiation (Balkenhal, 2009). The results of the Mantel tests for COI showed that there was significantly weak IBD for all of the genera and species, except for the genus *Pachycondyla* ($r = 0.53$, $P < 0.05$) which showed strong IBD.

Mantel tests performed for the nuclear DNA datasets showed weak but not significant IBD for *C. nr. cintellus* (18S; $r = 0.23$, $P > 0.05$, 28S; $r = 0.14$, $P > 0.05$). Weak, but significant IBD was detected for *Pachycondyla sp.* using 28S ($r = 0.33$, $P < 0.05$). Moderately significant IBD was detected for *Pachycondyla sp.* using 18S ($r = 0.67$, $P < 0.05$). When the concatenated datasets were considered, *C. nr. cintellus* exhibited weak, non-significant IBD ($r = 0.29$, $P > 0.05$) while *Pachycondyla sp.* exhibited weak but significant IBD ($r = 0.19$, $P < 0.05$).

As an indication of contemporary population structure, a Bayesian Analysis of Population Structure was performed for all the datasets. The highest number of estimated genetically distinct clusters were observed for the genus *Pheidole* (six, P = 0.99, log ln of optimal partition = -5202.67; Table 4.5) and six clusters for *Pachycondyla*, and five for *Camponotus* and three for *Lepisiota*. At the species level, there were four genetically distinct clusters each for *L. incisa* and *C.nr. cintellus* and three clusters for *P. caffraria* (Table 4.5).

BAPS analyses performed on the nuclear datasets for *Camponotus* and *Pachycondyla sp.* showed that the highest number of estimated genetically distinct clusters were found for the *Pachycondyla sp.* 28S dataset (four clusters, P = 1.0, log ln of optimal partition = -1053.56; Table 4.5). A total of three genetically distinct clusters were recovered for both *Camponotus.* (18S) and *Pachycondyla sp.* (18S and combined dataset), while only two genetically distinct clusters were recovered for *Camponotus* with the 18S dataset and the combined dataset (Table 4.5).

Table 4.5. Summary data for the BAPS analysis carried out for COI for all taxa, and for nuclear DNA (18S and 28S) and combined (COI+18S+28S) for *Camponotus* and *Pachycondyla*.

	Estimated no. of clusters	**Probability for no. of clusters**	**Log likelihood**
Lepisiota	3	0.99	-763.21
L. incisa	4	0.85	-29.91
Pachycondyla	6	1.0	-2379.14
P. caffraria	3	0.99	-158.54
Pheidole	5	0.99	-5202.67
Camponotus	5	1.0	-3639.97
C. nr. cintellus	4	0.89	-159.56
Camponotus (18S)	3	1.0	-783.10
Camponotus (28S)	2	0.57	-96.32
Camponotus (Combined)	2	1.0	-2591.21
Pachycondyla (18S)	3	0.51	-985.67
Pachycondyla (28S)	4	1.0	-1053.56
Pachycondyla (Combined)	3	0.97	-7664.06

Genetic landscape shape interpolation plots are useful in highlighting genetic discontinuities across geographic space. On the genetic landscapes, the darkly coloured blue extreme peaks indicate high levels of genetic variability while the lightly coloured green troughs and flat surfaces indicate that the localities have a high degree of genetic connectivity and hence, not much variation.

In Figure 4.16a, three distinct peaks are observed along the north-western edge and southern edge of the genetic landscape for *Lepisiota*. These peaks correspond to UKZN Agric, New Germany and Vernon Crookes nature reserves as these localities have two genetically distinct BAPS clusters each (Figure 4.16b). For *L. incisa*, it is interesting to note that there are four distinct peaks on the genetic landscape, two of which are situated along the north-western edge and two which are situated on the south-eastern edge (Figure 4.16c). In Figure 4.15d, it is seen that there are six localities which have more than two genetically distinct BAPS

clusters (Krantzkloof, New Germany, Seaton Park, Burman Bush, UKZN Westville and Silverglen). This is unexpected as *Lepisiota* showed more variability than *L. incisa* and *L. incisa* sequences are a part of the *Lepisiota* dataset.

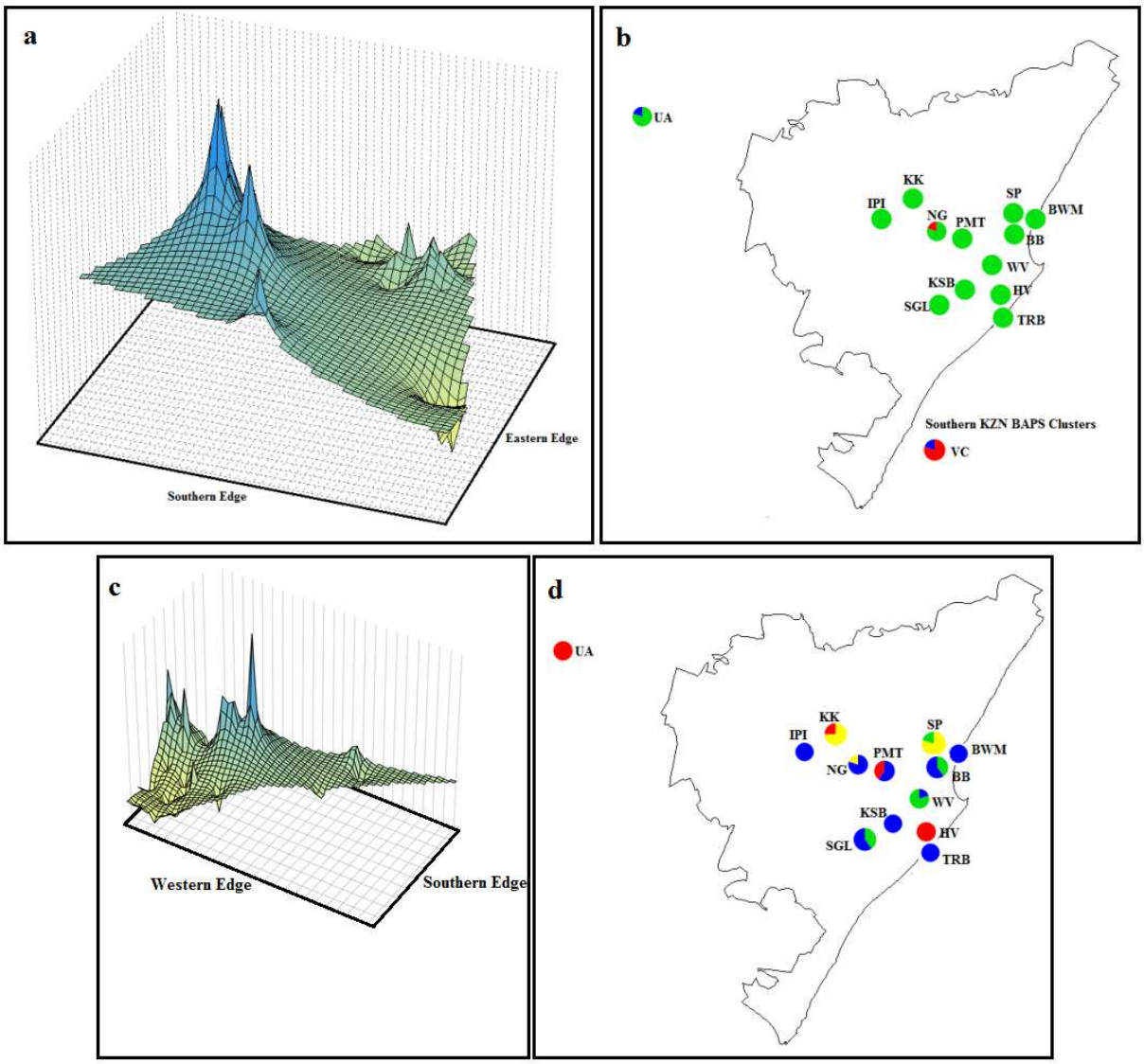


Figure 4.16. (a) Genetic landscape shape interpolation plot for *Lepisiota* using COI sequence data. (b) Map of sampling localities for *Lepisiota* in eThekwini and BAPS clusters. (c) Genetic landscape shape interpolation plot for *L. incisa* using COI sequence data. (d) Map of sampling localities for *L. incisa* in eThekwini and BAPS clusters.

The genetic landscape for *Pheidole* indicated that there were six distinct peaks of high genetic variability which corresponded to the six BAPS clusters (Figure 4.17 a and b). Only four of the localities contained more than one genetically distinct BAPS cluster; New Germany contained five BAPS clusters, while Ipithi and Beachwood Mangroves contained two BAPS clusters each.

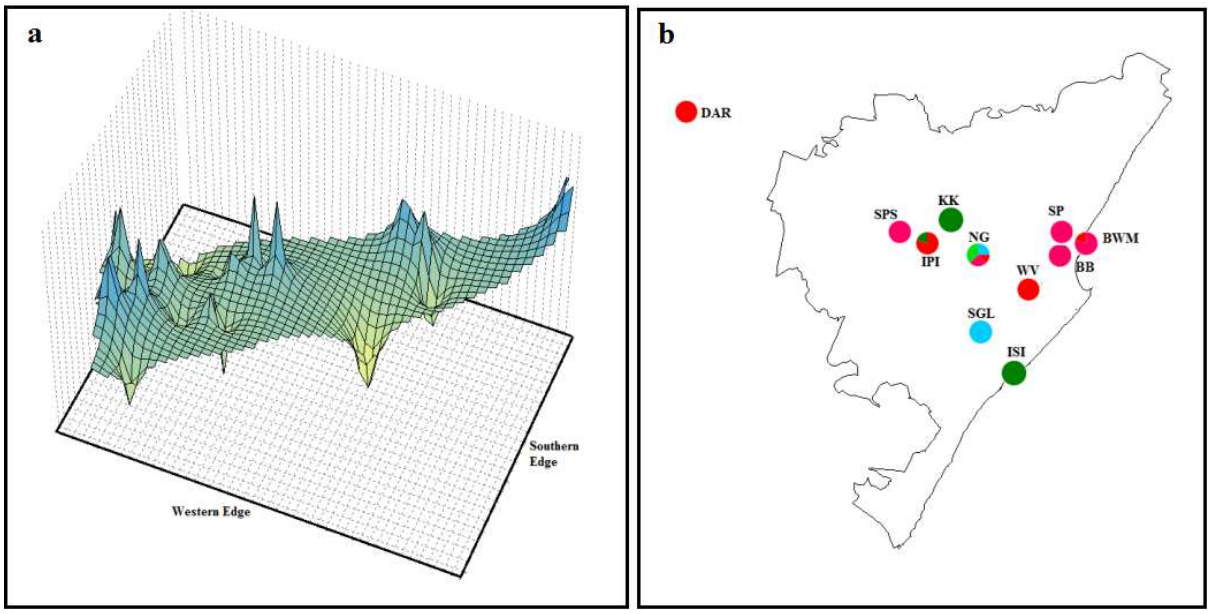


Figure 4.17. (a) Genetic landscape shape interpolation plot for *Pheidole* using COI sequence data. (b) Map of sampling localities for *Pheidole* in eThekwini and BAPS clusters.

Figure 4.18a shows a distinct peak along the northern edge as well as a slight peak in the western edge of the genetic landscape for *Pachycondyla* (Figure 4.18a). These peaks probably correspond to Amatikulu, Ipithi and Palmiet nature reserves in which two genetically distinct BAPS clusters were observed (Figure 4.18b). For *P. caffraria*, there were two subtle peaks observed along the south-eastern edge and western edge which possibly correspond to Ipithi nature reserve which contained three genetically distinct BAPS clusters and Kenneth Stainbank which had two genetically distinct BAPS clusters (Figure 4.18 c and d).

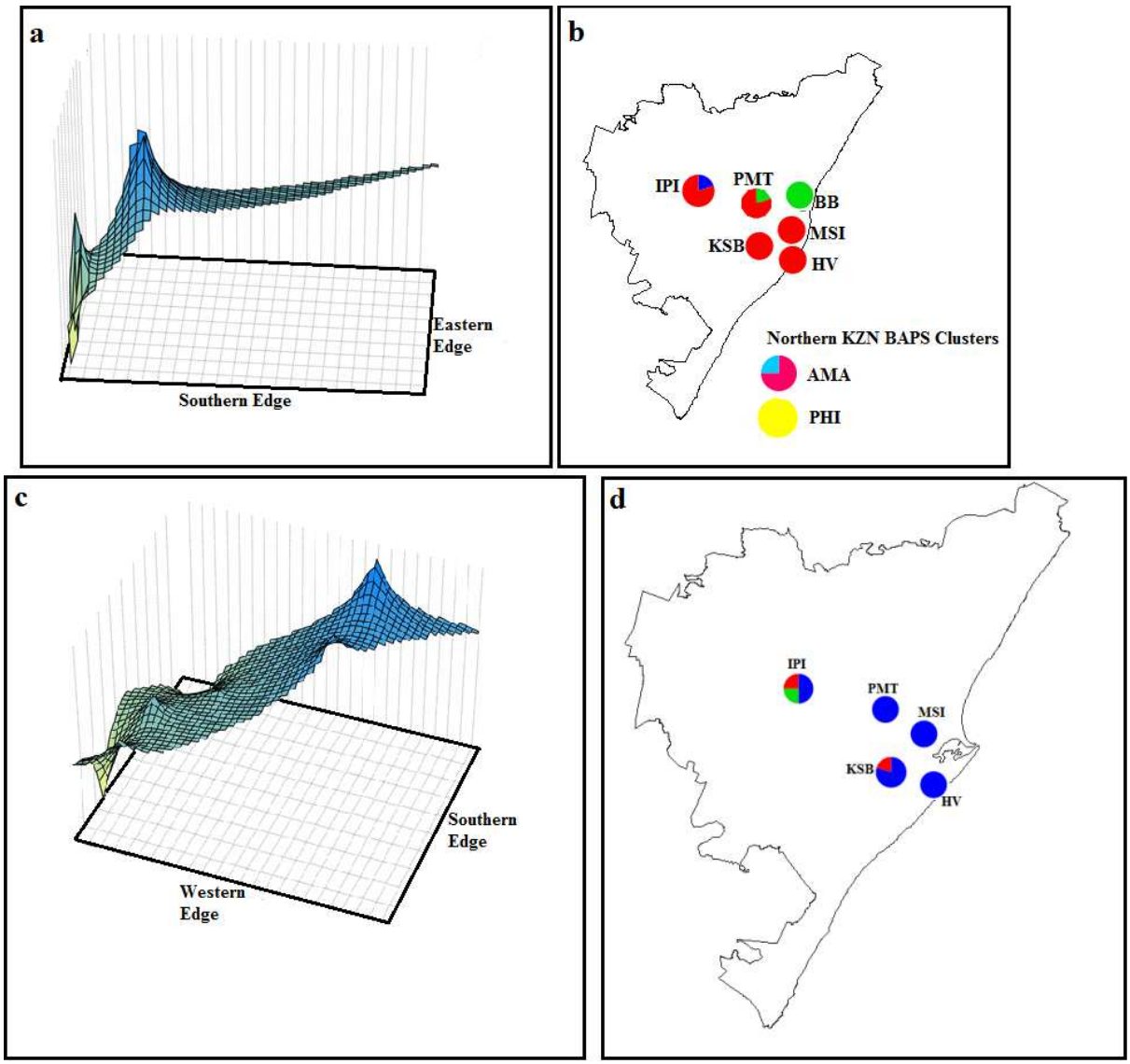
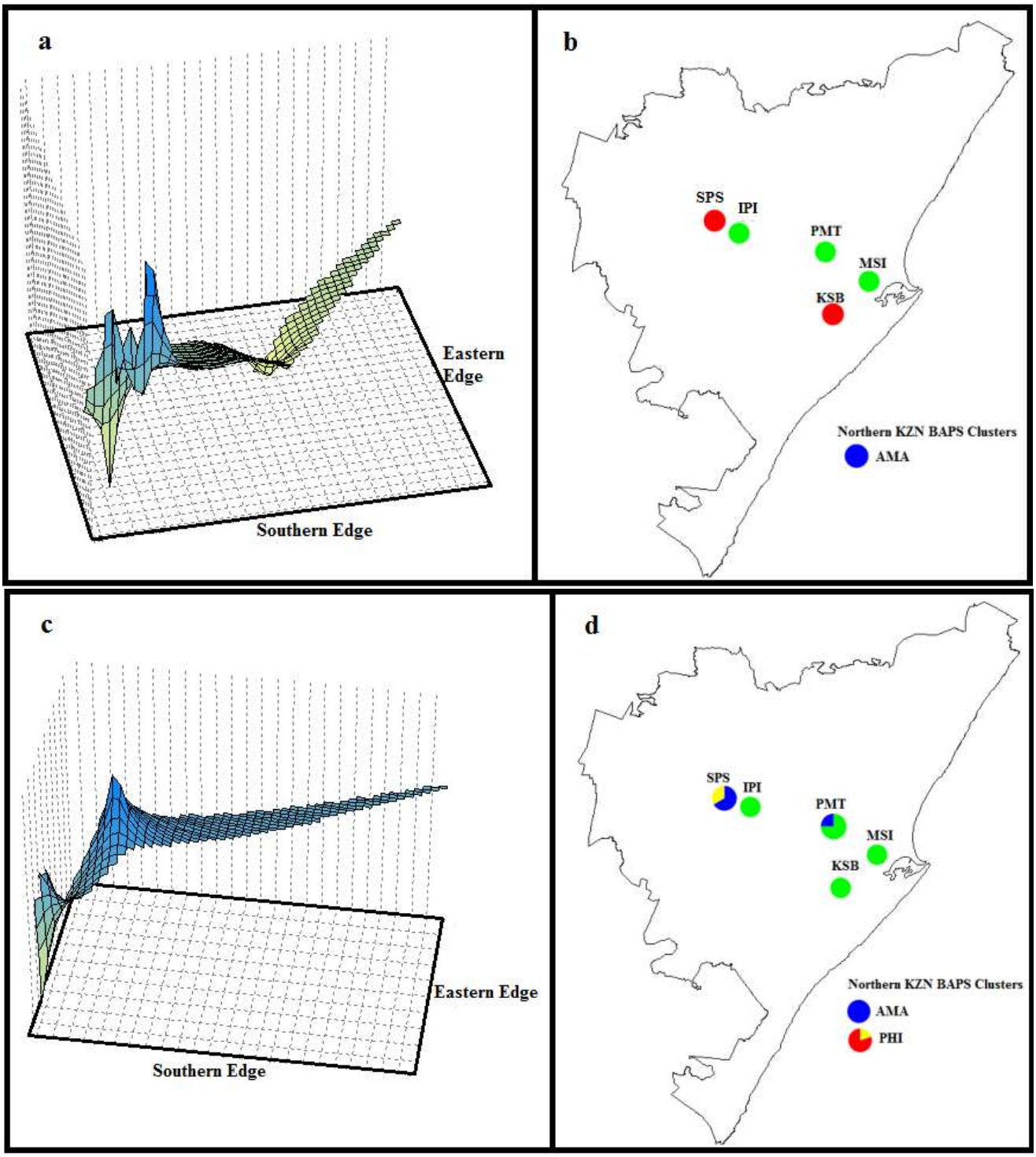


Figure 4.18. (a) Genetic landscape shape interpolation plot for *Pachycondyla* using COI sequence data. (b) Map of sampling localities for *Pachycondyla* in eThekwini and BAPS clusters. (c) Genetic landscape shape interpolation plot for *P. caffraria* using COI sequence data. (d) Map of sampling localities for *P. caffraria* in eThekwini and BAPS clusters.

The genetic landscapes for *Pachycondyla sp.* using 18S and 28S were similar to the genetic landscape for *Pachycondyla sp.* using COI (Figure 4.19 a and c). However, when the combined dataset was used, the peak in the southern edge became a trough (Figure 4.19 e). For *Pachycondyla sp.* at 28S, there were three localities which had two genetically distinct BAPS clusters (Springside, Palmiet and Phinda; Figure 4.19 d). Using the combined dataset, it was observed that there were five localities which had two genetically distinct BAPS

clusters (Springside, Ipithi, Palmiet, Msinsi and Kenneth Stainbank; Figure 4.19 f). These results correlate with those obtained for *Pachycondyla sp.* and *P. caffraria* using the COI dataset (Figure 4.18).



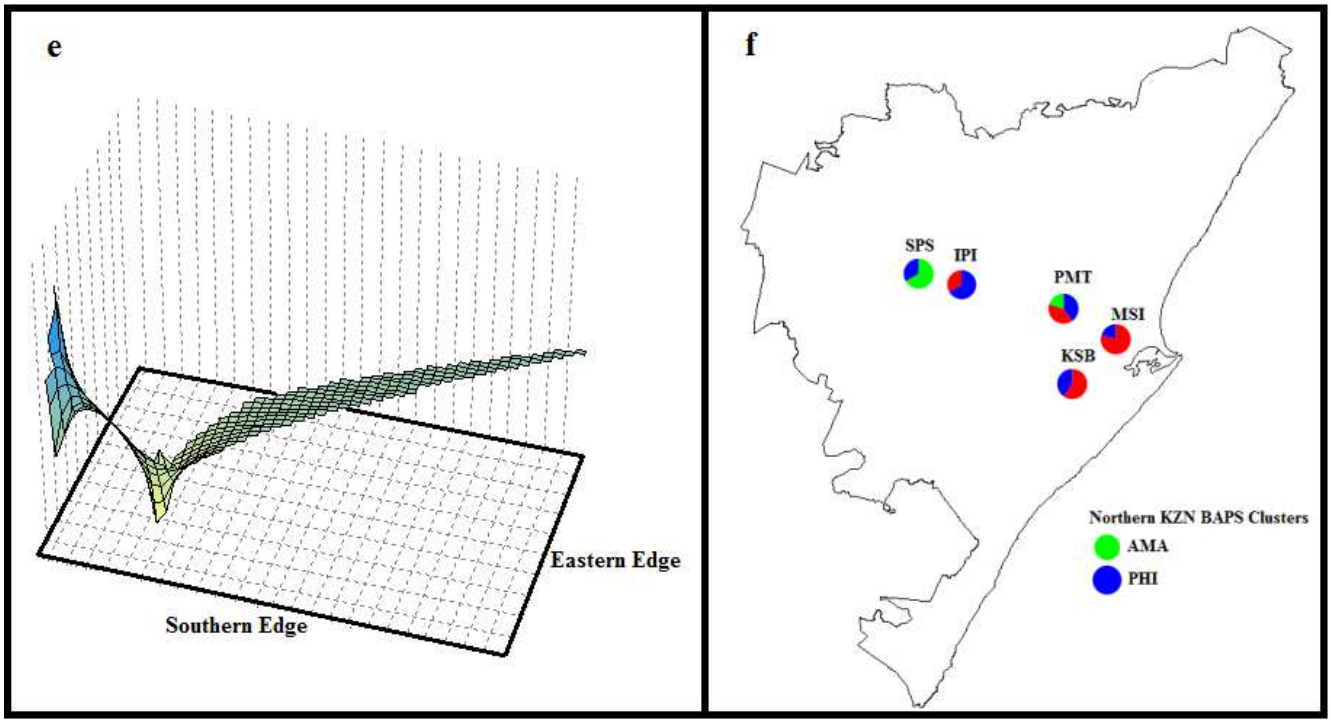


Figure 4.19. (a) Genetic landscape shape interpolation plot for *Pachycondyla sp*. using 18S sequence data. (b) Map of sampling localities for *Pachycondyla sp*. in eThekwini and BAPS clusters. (c) Genetic landscape shape interpolation plot for *Pachycondyla sp*. using COI sequence data. (d) Map of sampling localities for *Pachycondyla sp*. in eThekwini and BAPS clusters. (e) Genetic landscape shape interpolation plot for *Pachycondyla sp*. using the combined dataset. (f) Map of sampling localities for *Pachycondyla sp*. in eThekwini and BAPS clusters.

There were four distinct peaks observed along the genetic landscape for *Camponotus* (Figure 4.20 a). The peak along the western edge of the plot possibly corresponds to Springside nature reserve, in which there were three genetically distinct BAPS clusters (Figure 4.20 b). The peaks on the northern edge could correspond to New Germany and Kenneth Stainbank nature reserves and the peak on the southern edge could correspond to Vernon Crookes nature reserve because each of these localities contained two genetically distinct BAPS clusters (Figure 4.20 b). For *C.nr. cintellus*, there were only two subtle peaks observed along the north-western and south-eastern edges (Figure 4.20 c). It is likely that these peaks correspond to Springside and Burman Bush nature reserves, both of which contained two genetically distinct BAPS clusters (Figure 4.20 d).

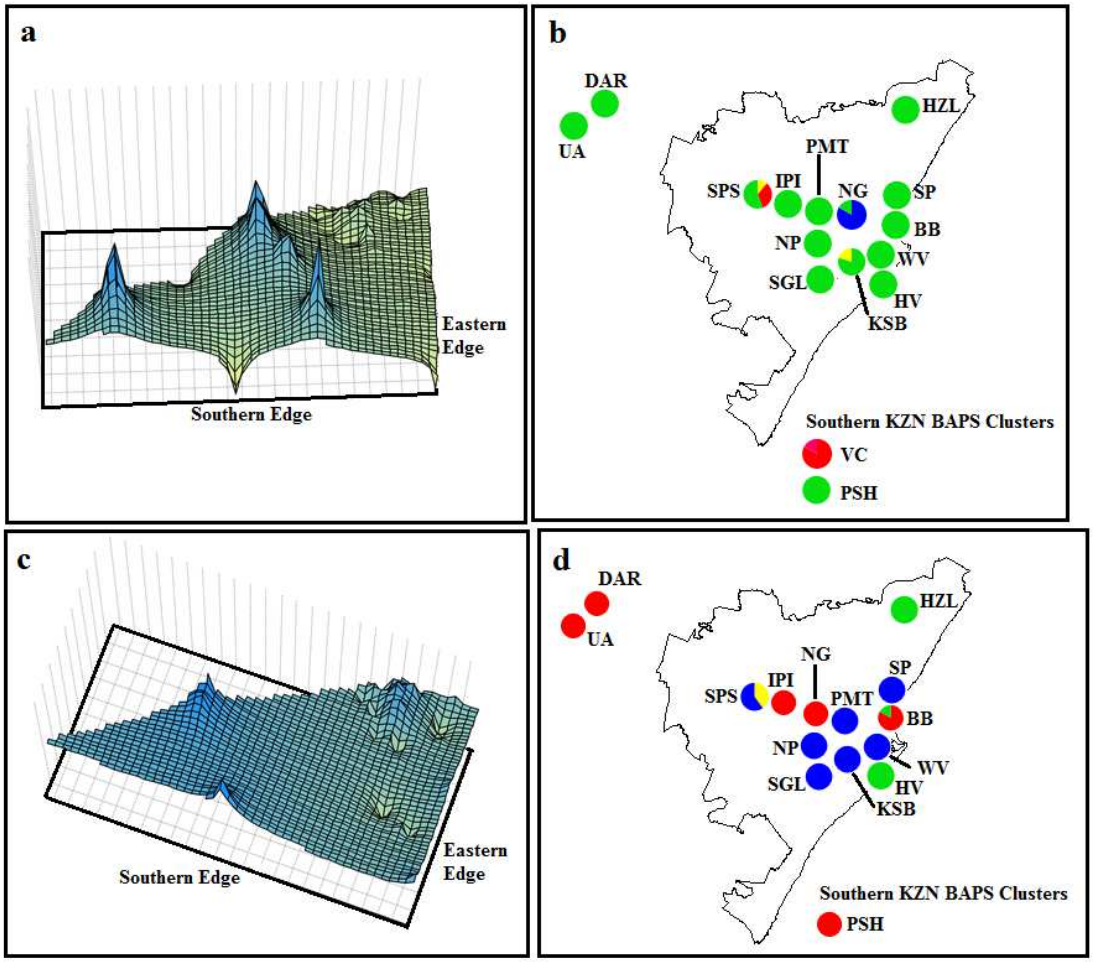
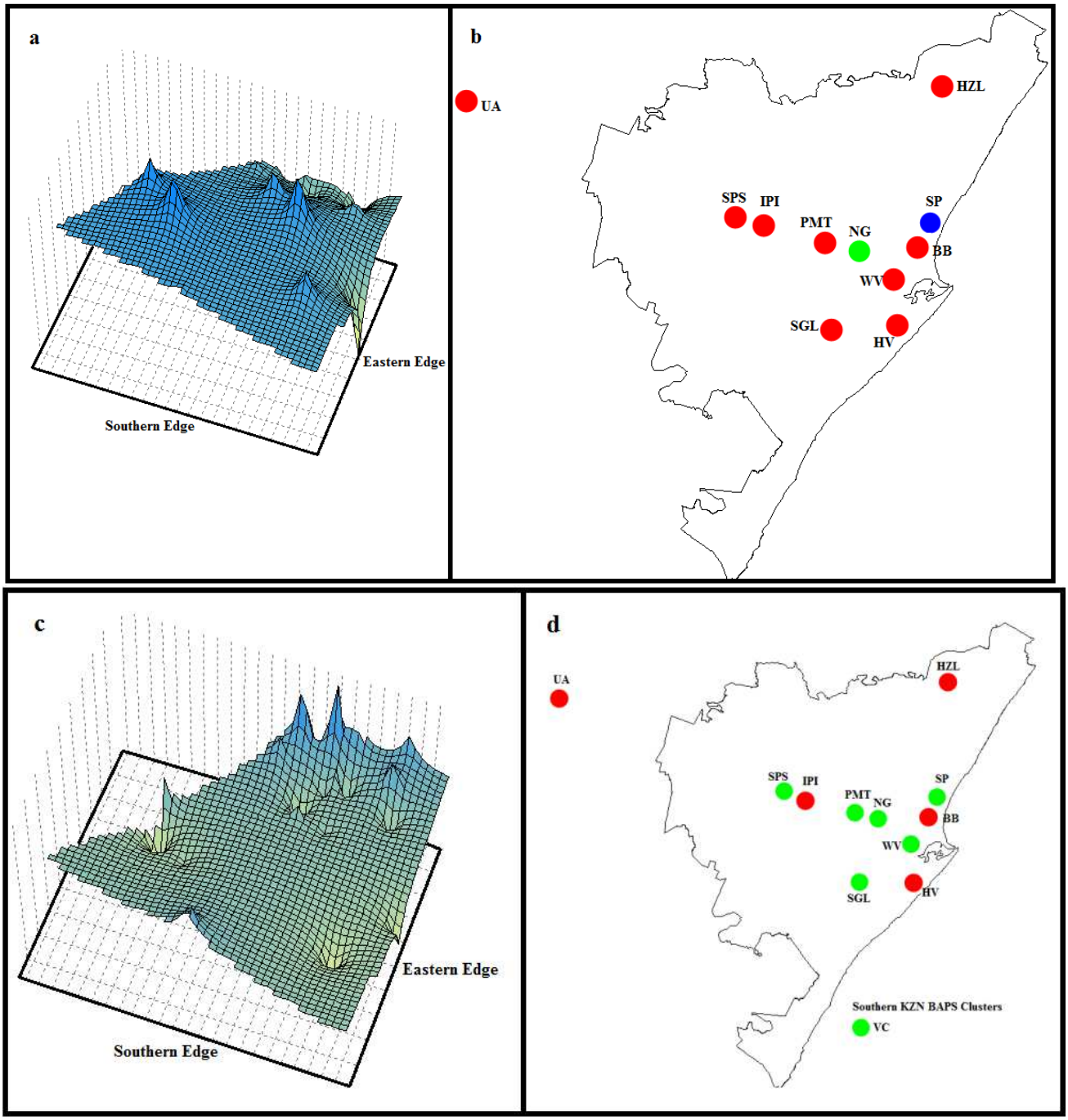


Figure 4.20. (a) Genetic landscape shape interpolation plot for *Camponotus* using COI sequence data. (b) Map of sampling localities for *Camponotus* in eThekwini and BAPS clusters. (c) Genetic landscape shape interpolation plot for *C. nr. cintellus* using COI sequence data. (d) Map of sampling localities for *C. nr. cintellus* in eThekwini and BAPS clusters.

The genetic landscape for *Camponotus* using 18S and the combined dataset were similar to the genetic landscape for *Camponotus* using COI (Figure 4.21 a and e). The genetic landscape for *Camponotus* using 28S recovered only two peaks of genetic variability in the northern edge (Figure 4.21 c). In Figure 4.20 b, New Germany and Seaton Park nature reserves had unique BAPS clusters for 18S. For the combined dataset for *Camponotus*, UKZN Agric and Vernon Crookes had unique BAPS clusters (Figure 4.21 f). None of the localities contained more than one genetically distinct BAPS clusters for 18S, 28S or the combined dataset whereas the dataset with COI alone showed four localities with two genetically distinct BAPS clusters. This suggests that COI shows more population structure than the nuclear DNA in *Camponotus.*


a
Eastern Edge
Southern Edge
b
UA
HZL
SPS
IPI
SP
PMT
NG
BB
WV
SGL
HV
c
Eastern Edge
Southern Edge
d
UA
HZL
SPS
IPI
SP
PMT
NG
BB
WV
SGL
HV
Southern KZN BAPS Clusters
VC

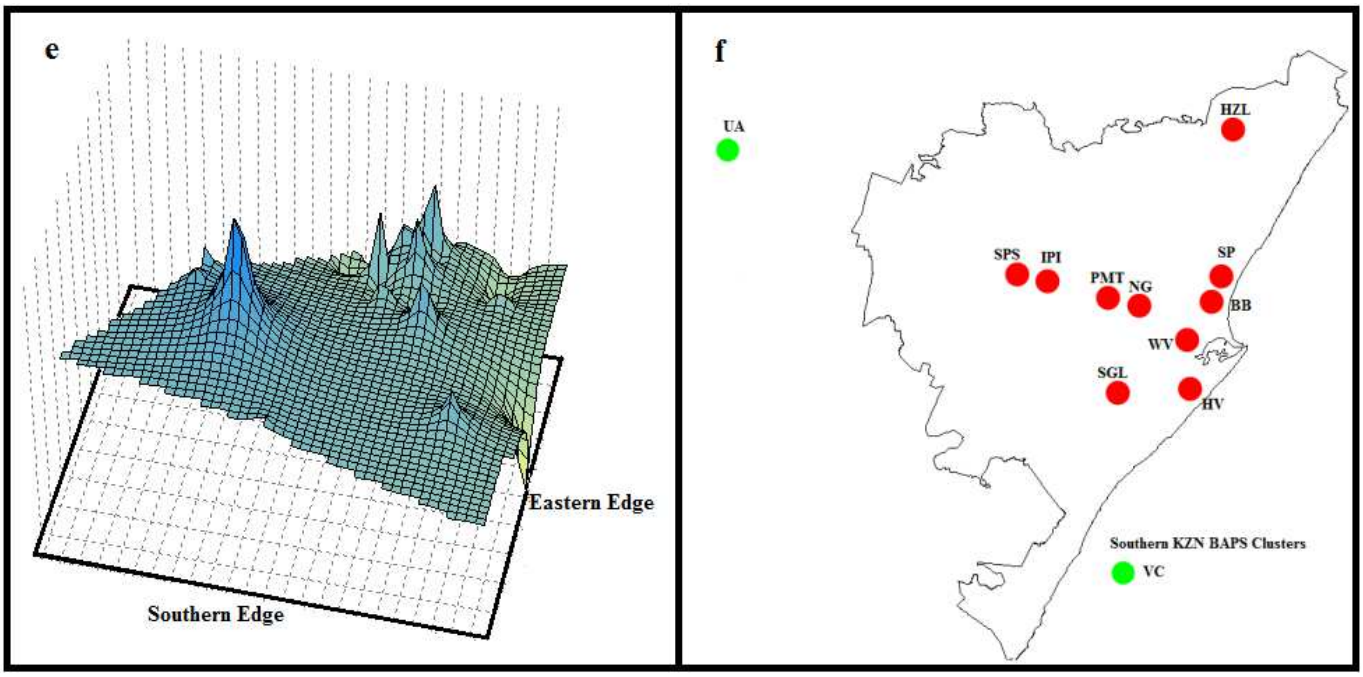


Figure 4.21. (a) Genetic landscape shape interpolation plot for *Camponotus* using 18S sequence data. (b) Map of sampling localities for *Camponotus* in eThekwini and BAPS clusters. (c) Genetic landscape shape interpolation plot for *Camponotus* using COI sequence data. (d) Map of sampling localities for *Camponotus* in eThekwini and BAPS clusters. (e) Genetic landscape shape interpolation plot for *Camponotus* using the combined dataset. (f) Map of sampling localities for *Camponotus* in eThekwini and BAPS clusters.

4.3.4 Evaluation of nuclear markers as a complement to COI

The uncorrected pairwise K2P genetic distances for COI, 18S and 28S were plotted in Figure 4.22. The uncorrected pairwise genetic distance was highest for COI (0.12 ± 0.006) for both *Camponotus* and *Pachycondyla*. *Pachycondyla* had higher uncorrected pairwise genetic distances than *Camponotus* for both 18S and 28S. The genetic distance for 28S (0.024 ± 0.01) was higher than 18S (0.019 ± 0.01) for *Pachycondyla*, whereas the genetic distance was higher for 18S (0.014 ± 0.004) than for 28S (0.013 ± 0.005) for *Camponotus*.

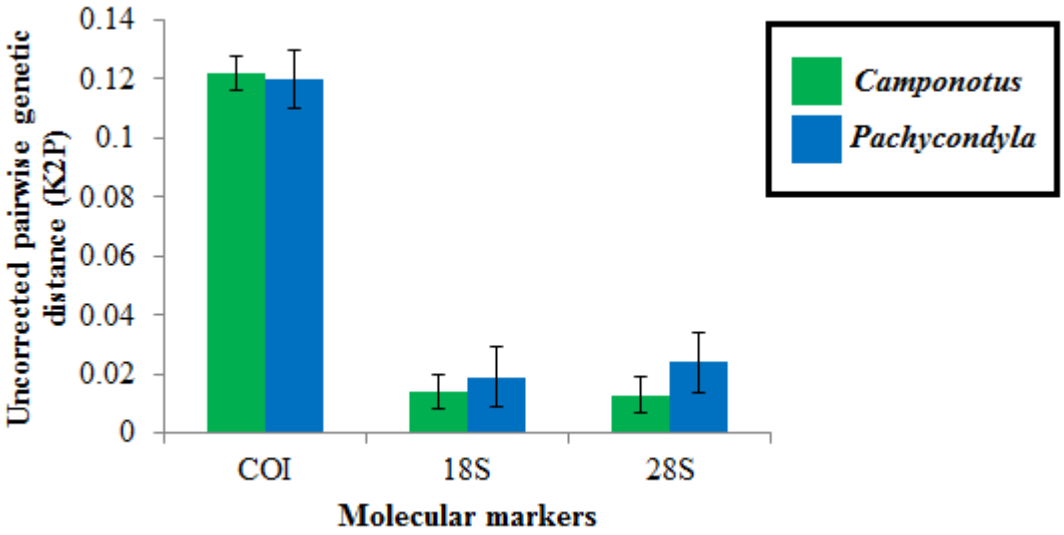


Figure 4.22. Comparison of uncorrected pairwise genetic distances calculated using the K2P distance parameter for mitochondrial COI and nuclear 18S and 28S for *Camponotus* and *Pachycondyla*.

4.4 Discussion

This study highlights the utility of landscape genetics and phylogeographic methods in understanding habitat connectivity and past and present processes and factors that shape the genetic structure of populations in an urban environment. The landscape genetic and phylogeographical patterns among co-distributed populations of ants from the genera *Lepisiota, Pheidole, Pachycondyla* and *Camponotus* at COI were compared and contrasted. These patterns were also compared for species confirmed to be *L. incisa*, *P. caffraria*, *C. nr. cintellus* and for BINs from the genus *Pheidole* that were morphologically thought to be *P. megacephala*.

The landscape genetic and phylogeographic patterns at COI were compared for invasive species *L. incisa* and *P. megacephala*. They were the most widely distributed but there was very little similarity in their genetic structure. Usually, when a single invasion occurs, the species sometimes undergoes founding effects which leads to low genetic diversity when compared to the source population or native species (Dlugosch and Parker, 2008; Tsutsui and Case, 2001). However, separate multiple introductions of an invasive species could lead to new genetic combinations and high genetic variation (Dlugosch and Parker, 2008). Since *P. megacephala* was much more variable than *L. incisa*, this could indicate that there is cryptic speciation within *Pheidole sp*. or multiple introductions whereas *L. incisa* had a single introduction. In addition, *P. megacephala* has several features which promotes its success as an invader; it is unicolonial and has polygynous (multiple queen) colonies (Holldobler and Wilson, 1990; Tsutsui and Suarez, 2003). Cryptic speciation has been encountered in *P. megacephala* from Cameroon. One of the species persisted in rainforests while the other thrived in urban areas (Fournier et al., 2012). In sub-Saharan Africa, *P. megacephala* is further classified into approximately 10 sub-species (Fournier et al., 2012). Understanding what makes invasive species such as *P. megacephala* successful can be useful in the control and prevention of further invasions and can also aid in identifying other potentially invasive species.

Native species *C. nr. cintellus* and *P. caffraria* showed no similarity in their genetic structures. *P. caffraria* was more genetically variable than *C. nr. cintellus.* Unlike the invasive species, species from the genera *Pachycondyla* and *Camponotus* have monogynous (single queen) colonies (Dietemann and Peeters, 2000; Klotz et al., 2008). In this study, nuclear markers 18S and 28S were also sequenced for individuals from *Camponotus* and *Pachycondyla*, to complement the COI data. Using these two classes of molecular markers (mitochondrial and nuclear), allowed the comparison of patterns across species, populations and molecular data types.

The COI phylogeny for *Camponotus* could not be accurately compared to the nuclear phylogenies because only one individual per reserve was sequenced for 18S and 28S. The nuclear phylogenies for *Pachycondyla* were comparable to the COI phylogeny. Landscape genetic analyses such as genetic landscape shape interpolation plots which highlight areas of high genetic variability, and BAPS clustering analysis for *Camponotus* indicated that COI exhibited more population structure than 18S and 28S. For *Pachycondyla*, these analyses revealed that the nuclear data was congruent with the COI data.

Due to the low uncorrected pairwise genetic distances observed for 18S and 28S in the two genera when compared to COI, these markers may not be ideal for use in species delimitation or for studies below the species level. Nevertheless, further testing using other nuclear markers such as introns, which are more variable, should be conducted as nuclear markers have several advantages over mitochondrial markers (Creer et al., 2005; Zhang and Hewitt, 2003). These advantages include being biparentally inherited and containing more informative sites (Creer et al., 2005).

Landscape genetics analyses highlighted that New Germany, Ipithi, Springside and Palmiet nature reserves exhibited exceptionally high levels of genetic variability. These nature reserves also stood out for their diversity in the DNA barcoding analyses (chapter 2) and species assemblage and richness studies (chapter 3). In terms of conservation, these nature reserves should be preserved as they harbour a great deal of genetic diversity, and hence, biodiversity.

There were differences in the genetic structure of the four species of ants in eThekwini. The individual patterns suggest that the differences in evolutionary and life histories are more likely to have shaped population structure instead of effects due to the landscape (Murphy et al., 2010). Nevertheless, the influence of landscape on population genetic structure cannot be ruled out.

4.5 Conclusion

In this study, there was subtle genetic variation at COI for each of the species examined. In order to fully elucidate the population structure patterns which could be expected in eThekwini and surrounding regions, further sampling across more localities is essential. The use of more nuclear markers could also assist in uncovering these unique patterns of genetic variation in an urban setting.

Chapter Five

General Discussion

This MSc study set out to test if the barcoding marker COI could be confirmed as a species diagnostic tool for ants in the eThekwini region. The utility of this marker was also evaluated for its potential to uncover genetic diversity below the species level in four species of ants.

In chapter one, the status and progress of DNA barcoding in Africa was reviewed. The current challenges, successes and the utility of DNA barcoding were also reviewed. Using data available on BOLD from the biggest ant barcoding project in Africa to date, a total of 118 individuals of A. madagascarensis sampled from 39 localities across Madagascar and the neigbouring island of Mayotte, were analysed as a case study to show how the barcoding marker could reveal interesting population genetic, phylogeographic and landscape genetic patterns (Smith et al., 2005).

In order to demonstrate how the barcoding marker COI and the techniques used in chapter one could be useful for revealing diversity below the species level in South African ants, a DNA barcode library was assembled for the ants of eThekwini. The barcode library currently contains 619 individuals from 80 putative species, sampled from 23 geographic localities within eThekwini and surrounding regions. The ability of COI to accurately delineate species was evaluated using statistical methods implemented in the R package Spider (Brown et al., 2012).

The results of DNA barcoding to delineate species were compared to traditional morphological identification. Although DNA barcoding uncovered 80 species while the traditional taxonomic approach identified 51 morphospecies, the two methods should be used in unison to achieve a more sound and meaningful result. Thus, the cases of paraphyly and cryptic speciation highlighted by the COI marker should be subject to further review by taxonomists. More importantly, it also emphasized the need to sample multiple individuals of

the same species from different geographical localities in order to increase the chances of finding a correct match for a species and eliminating false positives and false negatives (Bergsten et al., 2012; Fisher, 1999; Packer et al., 2009).

In chapter three, data from the DNA barcode library for the ants of eThekwini was used to investigate the diversity and species richness of ants as well as highlight the natural and open spaces within eThekwini that contribute to maintaining biodiversity within the city. The use of haplotype accumulation curves and extrapolation measures of diversity demonstrated that many more species of ants could exist in eThekwini. The investigation of ant species assemblages in the different vegetation types of each of the green areas sampled revealed that forest and grassland habitats each supported unique species assemblages. One of the aims of this study was to determine if open spaces within eThekwini could assist in maintaining native ant diversity in an urban environment. The presence of only two invasive ant species, L. incisa and P. megacephala, suggest that native species of ants are not being outcompeted by the invasive species.

In chapter four, the phylogeographical and landscape genetic patterns were compared between these two invasive species and the native species, Pachycondyla caffraria and Camponotus nr. cintellus. Different patterns of genetic structure were obtained for each of the species which suggested that differences in evolutionary and life history shaped population genetic structure instead of landscape effects (Murphy et al., 2010).

DNA barcode data coupled with further research on life history, biology and behaviour could yield valuable insight on how these factors and the influence of environmental factors, shape the population genetics of ants not only in eThekwini, but the rest of South Africa. Future research should incorporate the use of highly variable nuclear DNA markers such as introns. Increasing the scale of spatial sampling is also crucial in uncovering all the species that could be encountered. The techniques and methods of analysis used in this study could be applied to building DNA barcode libraries for other organisms in other regions and thus contribute to the barcode of life initiative to catalogue all of Earth’s biodiversity.

6. References

Anonymous (2010). State of Biodiversity Report, e. Municipality, ed. (Durban: eThekwini Municipality).

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., and Saunders, N.C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics *18*, 489 - 522.

Azuma, N., Ogata, K., Kikuchi, T., and Higashi, S. (2006). Phylogeography of Asian weaver ants, *Oecophylla smaragdina*. Ecological Resources *21*, 126 - 136.

Bergsten, J., Bilton, D.T., Fujisawa, T., Elliot, M., Monaghan, M.T., Balke, M., Hendrich, L., Geijer, J., Herrmann, J., Foster, G.N., *et al.* (2012). The effect of geographical scale of sampling on DNA barcoding. Systematic Biology *61*, 851 - 869.

Bermingham, E., and Moritz, C. (1998). Comparative phylogeography: concepts and applications. Molecular Ecology *7*, 367-369.

Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K., and Das, I. (2006). Cryptic species as a window on diversity and conservation. TRENDS in Ecology and Evolution *22*, 148 - 155.

Blackburn, T.M., and Gaston, K.J. (1996). A sideways look at patterns in species richness, or why there are so few species outside the tropics. Biodiversity Letters *3*, 44 - 53.

Blaustein, R. (2013). Urban biodiversity gains new converts. BioScience *63*, 72 - 77.

Bolger, D.T., Suarez, A.V., Crooks, K.R., Morrison, S.A., and Case, T.J. (2000). Arthropods in Urban Habitat Fragments in Southern California: Area, Age and Edge Effects. Ecological Applications *10*, 1230 - 1248.

Bolton, B. (1994). Identification Guide to the Ant Genera of the World (United States of America: President and Fellows of Harvard College).

Bray, J.R., and Curtis, J.T. (1957). An ordination of the upland forest communities of Southern Wisconsin. Ecological Monographs *27*, 325 - 349.

Bromilow, S.M., and Sperling, F.A.H. (2011). Phylogeographic signal variation in mitochondrial DNA among geographically isolated grassland butterflies. Journal of Biogeography *38*, 299 - 310.

Brown, S.D.J., Collins, R.A., Boyer, S., Lefort, M.-C., Malumbres-Olarte, J., Vink, C.J., and Cruickshank, R.H. (2012). SPIDER: An R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Molecular Ecology Resources *12*, 562 - 565.

Buczkowski, G., and Richmond, D.S. (2012). The effect of urbanization on ant abundance and diversity: a temporal examination of factors affecting biodiversity. PLoS ONE *7*, e41729.

Chao, A., Chazdon, R.L., Colwell, R.K., and Shen, T.-J. (2005). A new statistical approach for assessing similarity of species composition with incidence and abundance data. Ecology Letters *8*, 148 - 159.

Chao, A., Chazdon, R.L., Colwell, R.K., and Shen, T.-J. (2006). Abundance-based similarity indices and their estimation when there are unseen species in samples. Biometrics *62*, 361 - 371.

Chatzimanolis, S., and Caterino, M.S. (2008). Phylogepgraphy and conservation genetics of Californian coastal terrestrial communities: a comparative study using three beetles. Insect Conservation and Diversity *1*, 222 - 232.

Cincotta, R.P., Wisnewski, J., and Engelman, R. (2000). Human population in the biodiversity hotspots. Nature *404*, 990 - 992.

Clarke, K.M., Fisher, B.L., and LeBuhn, G. (2008). The influence of urban park characteristics on ant (Hymenoptera, Formicidae) communities. Urban Ecosystems *11*, 317 - 334.

Collins, R.A., and Cruickshank, R.H. (2012). The seven deadly sins of DNA barcoding. Molecular Ecology Resources, 1 -7.

Collwell, R.K., and Coddington, J.A. (1994). Estimating terrestrial biodiversity through extrapolation. Philosophical Transactions of the Royal Society Biological Sciences *39*, 1 - 22.

Corander, J., and Marttinen, P. (2006). Bayesian identification of admixture events using multilocus molecular markers. Molecular Ecology *15*, 2833 - 2843.

Corander, J., Waldmann, P., and Sillanpaa, M.J. (2003). Bayesian analysis of genetic differentiation between populations. Genetics *163*, 367 - 374.

Cox, C.B., and Moore, P.D. (2010). Biogeography: An Ecological and Evolutionary Approach, 8 edn (United States of America: John Wiley & Sons).

Creer, S., Malhotra, A., Thorpe, R.S., and Pook, C.E. (2005). Targeting optimal introns for phylogenetic analysis in non-model taxa: experimental results in Asian pitvipers. Cladistics *21*, 390 - 395.

Croucamp, A. (2009). Our Biodiverse City, e.M.E.P.a.C.P. Department, ed. (Durban: eThekwini Municipality), pp. 6,8-9.

Crozier, R.H., and Crozier, Y.C. (1993). The Mitochondrial Genome of the Honeybee *Apis-Mellifera* - Complete Sequence and Genome Organization. Genetics *133*, 97 - 117.

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2012). jModeltest 2: more models, new heuristics and parallel computing. Nature Methods *9*, 772.

Dasgupta, B. (2005). DNA-BAR: distinguisher selection for DNA barcoding. Bioinformatics *21*, 3424 - 3426.

Dasmahapatra, K.K., Elias, M., Hill, R.I., Hoffman, J.I., and Mallet, J. (2010). Mitochondrial DNA barcoding detects some species that are real, and some that are not. Molecular Ecology Resources *10*, 264 - 273.

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

Davis, G.A., Havill, N.P., Adelman, Z.N., Caccone, A., Kok, L.T., and Salam, S.M. (2011). DNA barcodes and molecular diagnostics to distinguish an introduced and native *Laricobius* (Coleoptera: Derodontidae) species in eastern North America. Biological Control *58*, 53 - 59.

deWaard, J.R., Hebert, P.D.N., and Humble, L.M. (2011). A Comprehensive DNA Barcode Library for the Looper Moths (Lepidoptera: Geometridae) of British Columbia, Canada. PLoS ONE *6*, e18290.

Diederichs, N., Nichols, G., van Niekerk, M., and Boon, R. (2010). Green Landscaping Guideline, e. Municipality, ed. (Durban: eThekwini Municipality).

Dietemann, V., and Peeters, C. (2000). Queen influence on the shift from trophic to reproductive eggs laid by workers of *Pachycondyla apicalis*. Insectes Sociaux *47*, 223 - 228.

Dinca, V., Zakharov, E.V., Hebert, P.D.N., and Vila, R. (2010). Complete DNA barcode reference library for a country's butterfly fauna reveals high performance for temperate Europe. Proceedings of the Royal Society Biological Sciences *278*, 347 - 355.

Dlugosch, K.M., and Parker, I.M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Molecular Ecology *17*, 431 - 449.

Eddy, S.R. (1998). Profile hidden Markov models. Bioinformatics Review *14*, 755 - 763.

Ekrem, T., Willassen, E., and Stur, E. (2007). A comprehensive DNA sequence library is essential for identification with DNA barcodes. Molecular Phylogenetics and Evolution *43*, 530 - 542.

Excoffier, L., Laval, G., and Schneider, S. (2005). ARLEQUIN version 3.0: an integrated software package for population genetics data analysis. Evolutionary Bioinformatica Online *1*, 47 - 50.

Fairbanks, D.H.K., Reyers, B., and van Jaarsveld, A.S. (2001). Species and environment representation: Selecting reserves for the retention of avian diversity in KwaZulu-Natal, South Africa. Biological Conservation *98*, 365-379.

Felsenstein, J. (2005). PHYLIP (Phylogeny Inference Package) (Distributed by the author).

Fisher, B.L. (1999). Improving Inventory Efficiency: A Case Study of Leaf-Litter Ant Diversity in Madagascar. Ecological Applications *9*, 714 - 731.

Fisher, R.A., Corbet, A.S., and Williams, C.B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology *12*, 42 - 58.

Fitzhugh, K. (2006). DNA barcoding: An instance of technology-driven science? BioScience *56*, 462 - 463.

Flot, J.-F. (2010). SeqPhase: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. Molecular Ecology Resources *10*, 162 - 166.

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit 1 from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology *3*, 294-299.

Fournier, D., Tindo, M., Kenne, M., Masse, P.S.M., Van Bossche, V., De Coninck, E., and Aron, S. (2012). Genetic Structure, Nestmate Recognition and Behaviour of Two Cryptic Species of the Invasive Big-Headed Ant *Pheidole megacephala*. PLoS ONE *7*, e31480.

Frankham, R., Ballou, J.D., and Briscoe, D.A. (2010). Introduction to Conservation Genetics (Cambridge: Cambridge University Press).

Frezal, L., and Leblois, R. (2008). Four years of DNA barcoding: Current advances and prospects. Infection, Genetics and Evolution *8*, 727 - 736.

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

Gariepy, T.D., Lindsay, R., Ogden, N., and Gregory, T.R. (2012). Identifying the last supper: utility of the DNA barcode library for bloodmeal identification in ticks. Molecular Ecology Resources *12*, 646 - 652.

Gaston, K.J. (2000). Global patterns in biodiversity. Nature *405*, 220 - 227.

Graham, J.H., Krzysik, A.J., Kovacic, D.A., Duda, J.J., Freeman, D.C., Emlen, J.M., Zak, J.C., Long, W.R., Wallace, M.P., Chamberlain-Graham, C., *et al.* (2009). Species richness, equitability, and abundance of ants in disturbed landscapes. Ecological Indicators *9*, 866 - 877.

Greenstone, M.H., Rowley, D.L., Heimbach, U., Lundgren, G., Pfannenstiel, R.S., and Rehner, S.A. (2005). Barcoding generalist predators by polymerase chain reaction: carabids and spiders. Molecular Ecology *14*, 3247 - 3266.

Hajibabaei, M. (In Press). The golden age of DNA metasystematics. Trends in Genetics *983*, 1 - 3.

Hajibabaei, M., DeWaard, J.R., Ivanova, N.V., Ratnasingham, S., Dooh, R.T., Kirk, S.L., Mackie, P.M., and Hebert, P.D.N. (2005). Critical factors for assembling a high volume of DNA barcodes. Philosophical Transactions of the Royal Society Biological Sciences.

Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W., and Hebert, P.D.N. (2006). DNA barcodes distinguish species of tropical Lepidoptera. PNAS *103*, 968-971.

Hajibabaei, M., Singer, G.A.C., Hebert, P.D.N., and Hickey, D.A. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends in Genetics *23*, 167 - 172.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series *41*, 95 - 98.

Hamaide, B., ReVelle, C.S., and Malcom, S.A. (2006). Biological reserves, rare species and the trade-off between species abundance and species diversity. Ecological Economics *56*, 570 - 583.

Hebert, P.D.N., Cywinska, A., Ball, S.L., and DeWaard, J.R. (2003a). Biological identifications through DNA barcodes. Proceedings of the Royal Society London: Biological Sciences *270*, 313 - 321.

Hebert, P.D.N., Penton, E., Burns, J., and al., e. (2004a). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences *101*, 14812 - 14817.

Hebert, P.D.N., Ratnasingham, S., and DeWaard, J.R. (2003b). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society Biological Sciences *270*, S96 - S99.

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., and Francis, C.M. (2004). Identification of birds through DNA barcodes. PLoS Biology *2*, 312.

Heino, J., Parviainen, J., Paavola, R., Jehle, M., Louhi, P., and Muotka, T. (2005). Characterizing macroinvertebrate assemblage structure in relation to stream size and and tributary position. Hydrobiologia *539*, 121 - 130.

Heterick, B.E., Lythe, M., and Smithyman, C. (2013). Urbanisation factors impacting on ant (Hymenoptera: Formicidae) biodiversity in the Perth metropolitan area, Western Australia: Two case studies. Urban Ecosystems *16*, 145 - 173.

Heywood, V.H. (1995). The Global Biodiversity Assessment (Cambridge: Cambridge University Press).

Hickerson, M.J., Meyer, C.P., and Moritz, C. (2006). DNA barcoding will often fail to discover new animal species over broad parameter space. Systematic Biology *55*, 729 - 739.

Hijmans, R.J., Guarino, L., and Mathur, P. (2012). DIVA-GIS. Version 7.5.

Holderegger, R., Kamin, U., and Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: implications for landscape genetics. Landscape Ecology *21*, 797-807.

Holderegger, R., and Wagner, H.H. (2006). A brief guide to landscape genetics. Landscape Ecology *21*, 793-796.

Holderegger, R., and Wagner, H.H. (2008). Landscape Genetics. BioScience *3*, 199-207.

Holldobler, B., and Wilson, E.O. (1990). The Ants (Cambridge: Harvard University Press).

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

Hulme, P.E. (2009). Trade, transport and trouble: managing invasive species pathways in an era of globalization. Journal of Applied Ecology *46*, 10 - 18.

Imai, H.T., and Kubota, M. (1972). Karyological studies of Japanese ants (Hymenoptera, Formicidae). Chromosoma *37*, 193 - 200.

Ives, C.D., Taylor, M.P., Nipperess, D.A., and Hose, G.C. (2013). Effect of catchment urbanization on ant diversity in remnant riparian corridors. Landscape and Urban Planning *110*, 155 - 163.

James, P.M.A., Coltman, D.W., Murray, B.W., Hamelin, R.C., and Sperling, F.A.H. (2011). Spatial genetic structure of a symbiotic beetle-fungal system: toward multi-taxa integrated landscape genetics. PLoS ONE *6*, e25359.

Janzen, D.H., Hajibabaei, M., Burns, J.M., Hallwachs, W., Remigio, E., and Hebert, P.D.N. (2005). Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. Philosophical Transactions of the Royal Society Biological Sciences *360*, 1835 - 1845.

Jinbo, U., Kato, T., and Ito, M. (2011). Current progress in DNA barcoding and future implications for entomology. Entomological Science *14*, 107 - 124.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution *16*, 111 - 120.

Klotz, J., Hansen, L., Pospischil, R., and Rust, M. (2008). Urban ants of North America and Europe, identification, biology and management (New York: Cornell University Press).

Kremen, C., Colwell, R.K., Erwin, T.L., Murphy, D.D., Noss, R.F., and Sanjayan, M.A. (1993). Terrestrial arthropod assemblages: Their use in conservation planning. Society for Conservation Biology *7*, 796 -808.

Kubisz, D., Kajtoch, L., Mazur, M.A., and Rizun, V. (2012). Molecular barcoding for central-eastern European *Crioceris* leaf-beetles (Coleoptera: Chrysomelidae). Central European Journal of Biology *7*, 69 - 76.

Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, J.G., and Savolainen, V. (2008). DNA barcoding the floras of biodiversity hotspots. PNAS *105*, 2923-2928.

Leppanen, J., Vepsalainen, K., Anthoni, H., and Savolainen, R. (2013). Comparative phylogeography of the ants *Myrmica ruginodis* and *Myrmica rubra*. Journal of Biogeography *40*, 479 - 491.

Lessard, J.-P., and Buddle, C.M. (2005). The effects of urbanization on ant assemblages (Hymenoptera: Formicidae) associated with the Molson Nature Reserve, Quebec. Canadian Entomology *137*, 215 - 225.

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

Linehan, J., Gross, M., and Finn, J. (1995). Greenway planning: developing a landscape ecological network approach. Landscape and Urban Planning *33*, 179 - 193.

Lunt, D.H., Zhang, D.X., Szymura, J.M., and Hewitt, G.M. (1996). The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. Insect Molecular Biology *5*, 153-165.

Magurran, A.E. (2004). Measuring biological diversity (USA: Blackwell Publishing).

Manel, S., Schwartz, M.K., Luikart, G., and Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. TRENDS in Ecology and Evolution *18*, 189 - 197.

Margalef, R. (1958). Information theory in ecology. General Systems *3*, 36 - 71.

McGinley, M., and Hogan, C.M. (2011). Biodiversity in Africa (United Nations Environment Programme).

McKinney, M.L. (2006). Urbanization as a major cause of biotic homogenization. Biological Conservation *127*, 247 - 260.

Meier, R., Shiyang, K., Vaidya, G., and Ng, P.K.L. (2006). DNA barcoding and taxonomy in Diptera: A tale of high intraspecific variability and low identification success. Systematic Biology *55*, 715 - 728.

Menke, S.B., Booth, W., Dunn, R.R., Schal, C., Vargo, E.L., and Silverman, J. (2010a). Is It Easy to Be Urban? Convergent Success in Urban Habitats among Lineages of a Widespread Native Ant. PLoS ONE *5*, e9194.

Menke, S.B., Guenard, B., Sexton, J.O., Weiser, M.D., Dunn, R.R., and Silverman, J. (2010b). Urban areas may serve as habitat and corridors for dry-adapted, heat tolerant species; an example from ants. Urban Ecosystems *14*, 135 - 163.

Meyer, C.P., and Paulay, G. (2005). DNA Barcoding: Error Rates Based on Comprehensive Sampling. PLoS Biology *3*, e422.

Miller, M.P. (2005). Alleles In Space: computer software for the joint-analysis of inter-individual spatial and genetic information. Journal of Heredity *96*, 722 - 724.

Mitrovich, M.J., Matsuda, T., Pease, K.H., and Fisher, R.N. (2010). Ants as a measure of effectiveness of habitat conservation planning in southern California. Conservation Biology *24*, 1239 - 1248.

Moritz, C., and Cicero, C. (2004). DNA barcoding: promise and pitfalls. PLoS Biology *2*, e279.

Murphy, N.P., Guzik, M.T., and Wilmer, J.W. (2010). The influence of landscape on population structure of four invertebrates in groundwater springs. Freshwater Biology *55*, 2499 - 2509.

Myers, N. (1988). Threatened biotas: "Hot spots in tropical forests". The Environmentalist *8*, 187 - 208.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. Nature *403*, 853 - 858.

Nielson, R., and Matz, M. (2006). Statistical Approaches for DNA Barcoding. Systematic Biology *55*, 162-169.

Pacheco, R., and Vasconcelos, H.L. (2007). Invertebrate conservation in urban areas: Ants in the Brazilian Cerrado. Landscape and Urban Planning *81*, 193 - 199.

Packer, L., Gibbs, J., Sheffield, C., and Hanner, R. (2009). DNA barcoding and the mediocrity of morphology. Molecular Ecology Resources *9*, 42 - 50.

Pecarevic, M., Danoff-Burg, J., and Dunn, R.R. (2010). Biodiversity on broadway - enigmatic diversity of the societies of ants (Formicidae) on the streets on New York City. PLoS ONE *5*, e13222.

Perez-Espona, S., Mcleod, J.E., and Franks, N.R. (2012). Landscape genetics of a top neotropical predator. Molecular Ecology *21*, 5969 - 5985.

Potter, S., Eldridge, M.D.B., Cooper, S.J.B., Paplinska, J.Z., and Taggart, D.A. (2012). Habitat connectivity, more than species' biology, influences genetic differentiation in a habitat specialist, the short-eared rock-wallaby (*Petrogale brachyotis*). Conservation Genetics *13*, 937 - 952.

Ratnasingham, S., and Hebert, P.D.N. (2007). BOLD: the barcode of life data system (http://www.barcodinglife.org). Molecular Ecology Notes *7*, 355 - 364.

Reid, W.V. (1998). Biodiversity hotspots. TREE *13*, 275 - 280.

Rubinoff, D. (2006). Utility of Mitochondrial DNA Barcodes in Species Conservation. Conservation Biology *20*, 1026 - 1033.

Saux, C., Fisher, B.L., and Spicer, G.S. (2004). Dracula ant phylogeny as inferred by nuclear 28S rDNA sequences and implications for ant systematics (Hymenoptera: Formicidae: Amblyoponinae). Molecular Phylogenetics and Evolution *33*, 457 - 468.

Savolainen, V., Cowan, R.S., Vogler, A.P., Roderick, G.K., and Lane, R. (2005). Towards writing the encyclopedia of life: an introduction to DNA barcoding. Philosophical Transactions of the Royal Society Biological Sciences *360*, 1805-1811.

Schilthuizen, M., Scholte, C., van Wijk, R.E.J., Dommershuijzen, J., van der Horst, D., zu Schlochtern, M.M., Lievers, R., and Groenenberg, D.S.J. (2011). Using DNA-barcoding to make the necrobiont beetle family Cholevidae accessible for forensic entomology. Forensic Science International *210*, 91 -95.

Schoeman, C.S., and Foord, S.H. (2012). A checklist of epigaeic ants (Hymenoptera: Formicidae) from the Marakele National Park, Limpopo, South Africa. Koedoe *54*, 7 pages.

Seal, J.N., Kellner, K., Trindl, A., and Heinze, J. (2011). Phylogeography of the parthenogenic ant *Platythyrea punctata:* highly successful colonization of the West Indies by a poor disperser. Journal of Biogeography *38*, 868 - 882.

Segelbacher, G., Cushman, S.A., Epperson, B.K., Fortin, M.J., Francois, O., Hardy, O.J., Holderegger, R., Taberlet, P., Waits, L.P., and Manel, S. (2010). Applications of landscape genetics in conservation biology: concepts and challenges. Conservation Genetics *11*, 375-385.

Seppa, P. (2008). Do ants (Hymenoptera: Formicidae) need conservation and does ant conservation need genetics? Myrmecological News *11*, 161 - 172.

Shannon, C.E., and Weaver, W. (1963). The Mathematical Theory of Communication (Illinois, USA: The University of Illinois Press).

Sheffield, C.S., Hebert, P.D.N., Kevan, P.G., and Packer, L. (2009). DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies. Molecular Ecology Resources *9*, 196 - 207.

Simonsen, K.L., Churchill, G.A., and Aquardo, C.F. (1995). Properties of statistical tests of neutrality for DNA polymorphism data. Genetics *141*, 413 - 429.

Simpson, E.H. (1949). Measurement of diversity. Nature *163*, 688.

Sithole, H., Smit, I.P.J., and Parr, C.L. (2009). Preliminary investigations into a potential ant invader in Kruger National Park, South Africa. African Journal of Ecology *48*, 736 - 743.

Smith, M.A., and Fisher, B.L. (2009). Invasions, DNA barcodes, and rapid biodiversity assessment using ants of Mauritius. Frontiers in Zoology *6*, doi: 10.1186/1742-9994-1186-1131.

Smith, M.A., Fisher, B.L., and Hebert, P.D.N. (2005). DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philosophical Transactions of the Royal Society Biological Sciences *360*, 1825 - 1834.

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

Stoeckle, M.Y., and Hebert, P.D.N. (2008). Barcode of Life: DNA tags help classify animals. Scientific American *298*, 39 - 43.

Storfer, A., Murphy, M.A., Evans, J.S., Goldberg, C.S., Robinson, S., Vierling, L., and Waits, L.P. (2007). Putting the 'landscape' in landscape genetics. Heredity *98*, 128-142.

Suarez, A.V., Bolger, D.T., and Case, T.J. (1998). Effects of Fragmentation and Invasion on Native Ant Communities in Coastal Southern California. Ecology *79*, 2041 - 2056.

Swartz, E.R., Mwale, M., and Hanner, R. (2008). A role for barcoding in the study of African fish diversity and conservation. South African Journal of Science *104*, 293 - 298.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphisms. Genetics *123*, 585 - 595.

Tamura, K., Peterson, D., N., P., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA 5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Molecular Biology and Evolution *28*, 2731 - 2739.

Thompson, B., and McLachlan, S. (2007). The effects of urbanization on ant communities and myrmecochory in Manitoba, Canada. Urban Ecosystems *10*, 43 - 52.

Tollenaere, C., Brouat, C., Duplantier, J.-M., Rahalison, L., Rahelinirina, S., Pascal, M., Mone, H., Mouahid, G., Leirs, H., and Cosson, J.-F. (2010). Phylogeography of the introduced species *Rattus rattus* in the western Indian Ocean, with special emphasis on the colonization history of Madagascar. Journal of Biogeography *37*, 398 - 410.

Tsutsui, N.D., and Case, T.J. (2001). Population genetics and colony structure of the Argentine ant (*Linepithema humile*) in its native and introduced ranges. Evolution *55*, 976 - 985.

Tsutsui, N.D., and Suarez, A.V. (2003). The colony structure and population biology of invasive ants. Conservation Biology *17*, 48 - 58.

Valentini, A., Pompanon, F., and Taberlet, P. (2008a). DNA barcoding for ecologists. TRENDS in Ecology and Evolution *24*, 110 - 117.

Valentini, A., Pompanon, F., and Taberlet, P. (2008b). DNA barcoding for ecologists. Trends in Ecology and Evolution *24*, 110-117.

van Velzen, R., Weitschek, E., Felici, G., and Bakker, F.T. (2012). DNA Barcoding of Recently Diverged Species: Relative Performance of Matching Methods. PLoS ONE *7*, e30490.

Vergnes, A., Le Viol, I., and Clergeau, P. (2012). Green corridors in urban landscapes affect the arthropod communities of domestic gardens. Biological Conservation *145*, 171 - 178.

Vernooy, R., Haribabu, E., Muller, M.R., Vogel, J.H., Hebert, P.D.N., Schindel, D.E., Shimura, J., and Singer, G.A.C. (2010). Barcoding Life to Conserve Biological Diversity: Beyond the Taxonomic Imperative. PLoS Biology *8*, e1000417.

Walton, C.H., J.M., Tun-Lin, W., Collins, F.H., Harbach, R.E., Baimal, V., and Butlin, R.K. (2000). Population structure and population history of *Anopheles dirus* mosquitoes in Southeast Asia. Molecular Biology and Evolution *17*, 962 - 974.

Wang, I.J. (2010). Recognizing the temporal distinctions between landscape genetics and phylogeography. Molecular Ecology *19*, 2605 - 2608.

Waugh, J. (2007). DNA barcoding in animal species: progress, potential and pitfalls. BioEssays *29*, 188 - 197.

Webb, J.M., Jacobus, L.M., Funk, D.H., Zhou, X., Kondratieff, B., Geraci, C.J., DeWalt, E., Baird, D.J., Richard, B., Phillips, I., *et al.* (2012). A DNA Barcode Library for North American Ephemeroptera: Progress and Prospects. PLoS ONE *7*, e38063.

Weitschek, E., Van Velzen, R., Felici, G., and Bertolazzi, P. (2013). BLOG 2.0: a software system for character-based species classification with DNA Barcode sequences. What it does, how to use it. Molecular Ecology Resources.

Wiemers, M., and Fiedler, K. (2007). Does the DNA barcoding gap? - a case study in blue butterflies (Lepidoptera: Lycaenidae). Frontiers in Zoology *4*, 10.1186/1742-9994-1184-1188.

Willig, M.R., Kaufman, D.M., and Stevens, R.D. (2003). Latitudinal gradients of biodiversity: pattern, process, scale and synthesis. Annual Review of Ecology, Evolution and Systematics *34*, 273 - 309.

Wilson, J.J., Rougerie, R., Schonfeld, J., Janzen, D.H., Hallwachs, W., Hajibabaei, M., Kitching, I.J., Haxaire, J., and Hebert, P.D. (2011). When species matches are unavailable are DNA barcodes correctly assigned to higher taxa? An assessment using sphingid moths. BMC Ecology *11*, http://www.biomedcentral.com/1472-6785/1411/1418.

Yamaguchi, T. (2005). Influence of urbanization on ant distribution in parks of Tokyo and Chiba City, Japan II. Analysis of species. Entomological Science *8*, 17 - 25.

Yasuda, M., and Koike, F. (2009). The contribution of the bark of isolated trees as habitat for ants in an urban landscape. Landscape and Urban Planning *92*, 276 - 281.

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

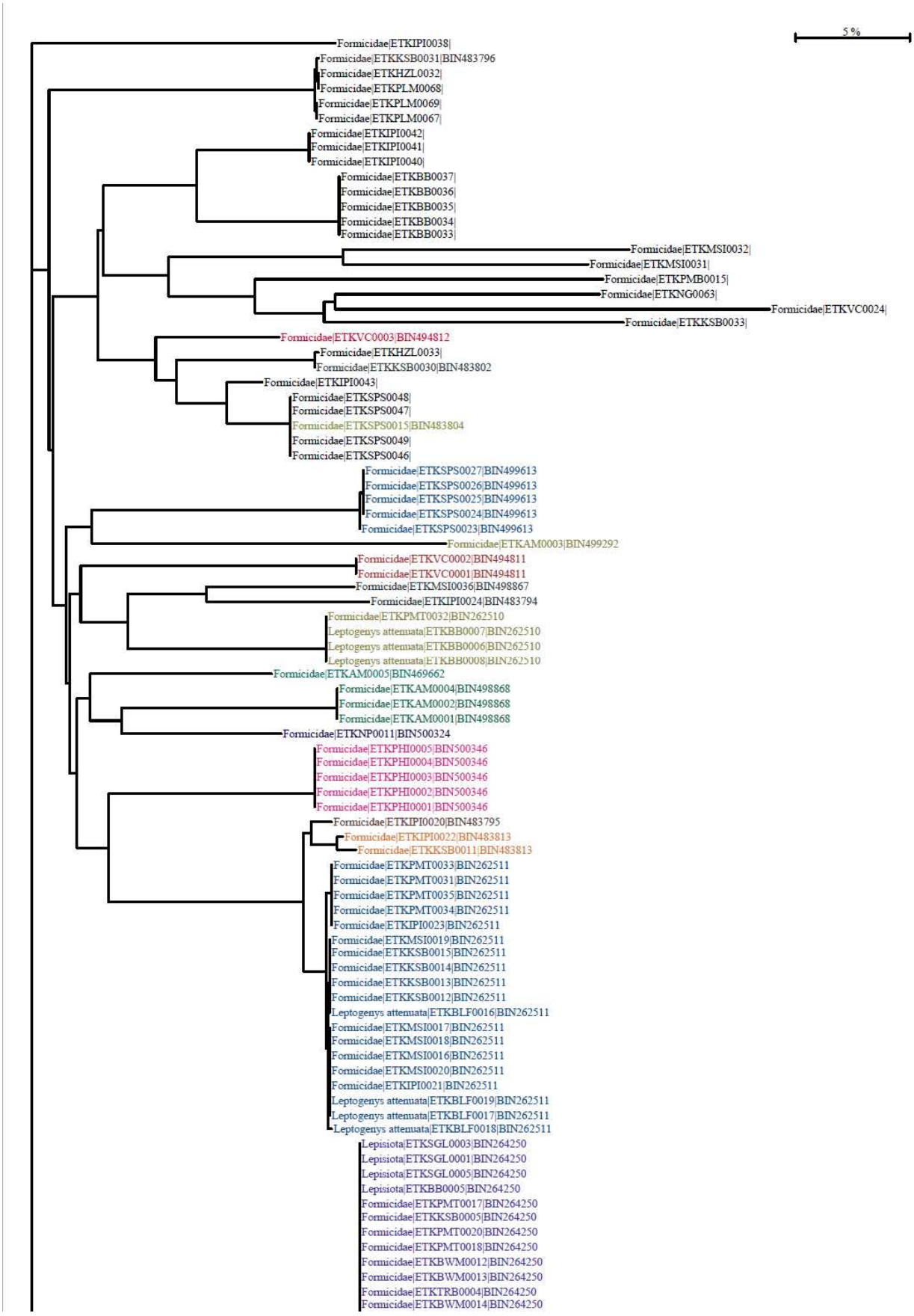
Zhang, Y.-Z., Si, S.-L., Zheng, J.-T., Li, H.-L., Fang, Y., Zhu, C.-D., and Vogler, A.P. (2011). DNA barcoding of endoparasitoid wasps in the genus *Anicetus* reveals high levels of host specificity (Hymenoptera: Encyrtidae). Biological Control *58*, 182 - 191.

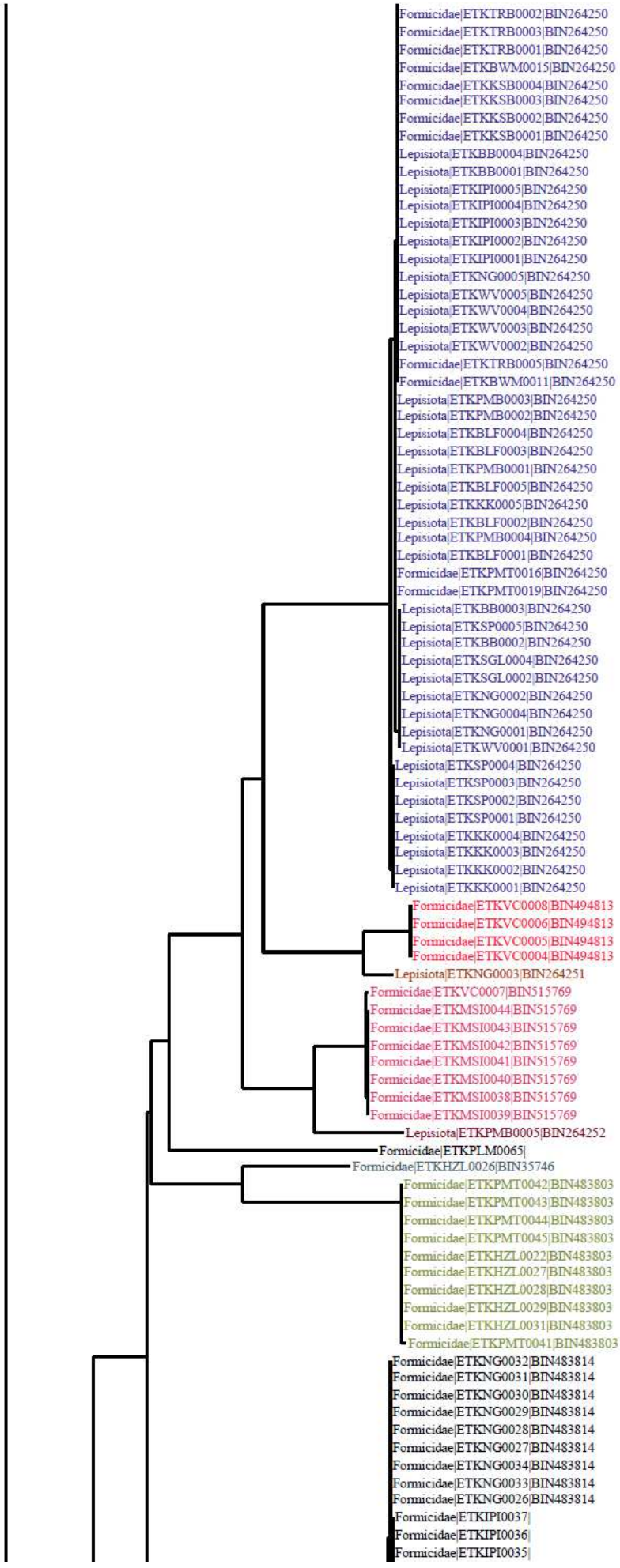
Zhou, X., Adamowicz, S.J., Jacobus, L.M., DeWalt, E., and Hebert, P.D.N. (2009). Towards a comprehensive barcode library for arctic life - Ephemeroptera, Plecoptera, and Trichoptera of Churchill, Manitoba, Canada. Frontiers in Zoology *6*.

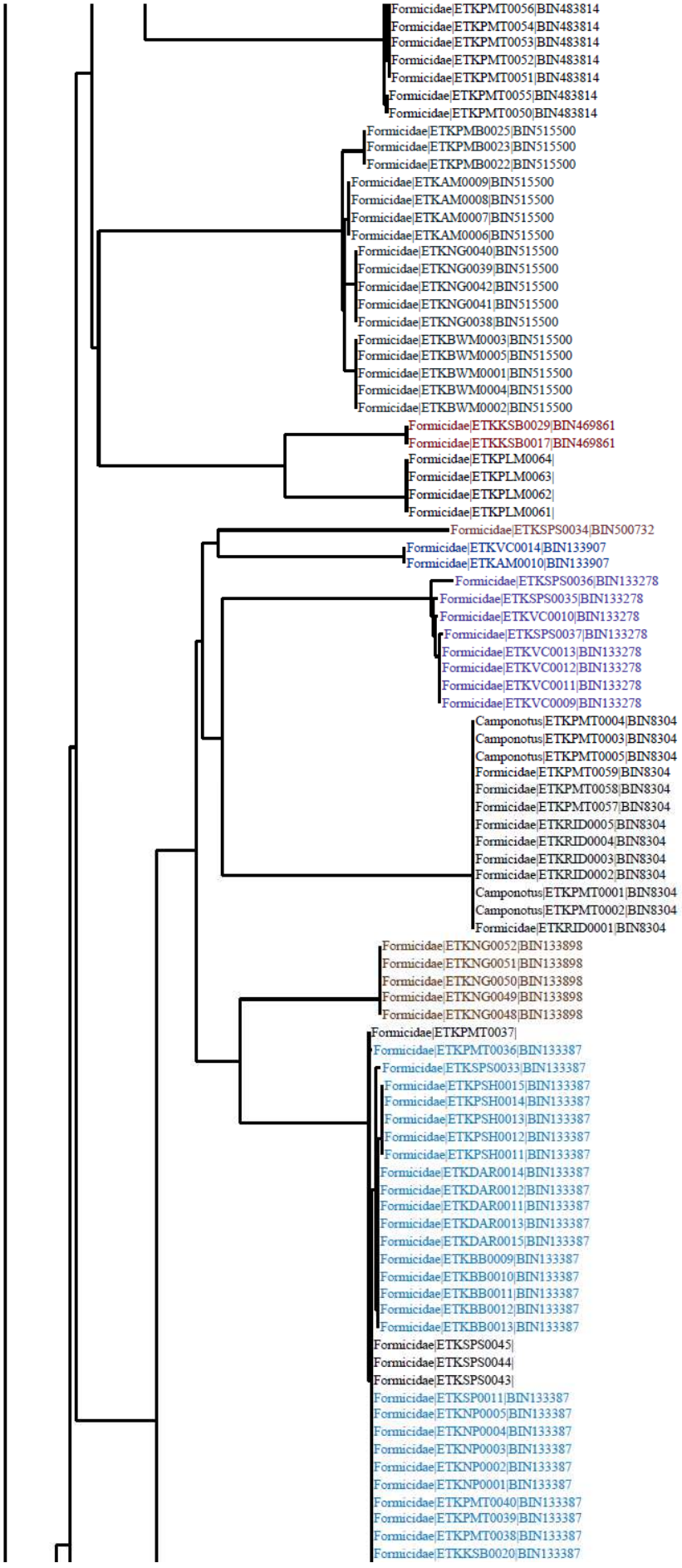
Zhou, X., Robinson, J.L., Geraci, C.J., Parker, C.R., Flint, O.S., Etnier, D.A., Ruiter, D., DeWalt, E., Jacobus, L.M., and Hebert, P.D.N. (2011). Accelerated construction of a regional DNA-barcode reference library: caddisflies (Trichoptera) in the Great Smoky Mountains National Park. Journal of the North American Benthological Society *30*, 131 - 162.

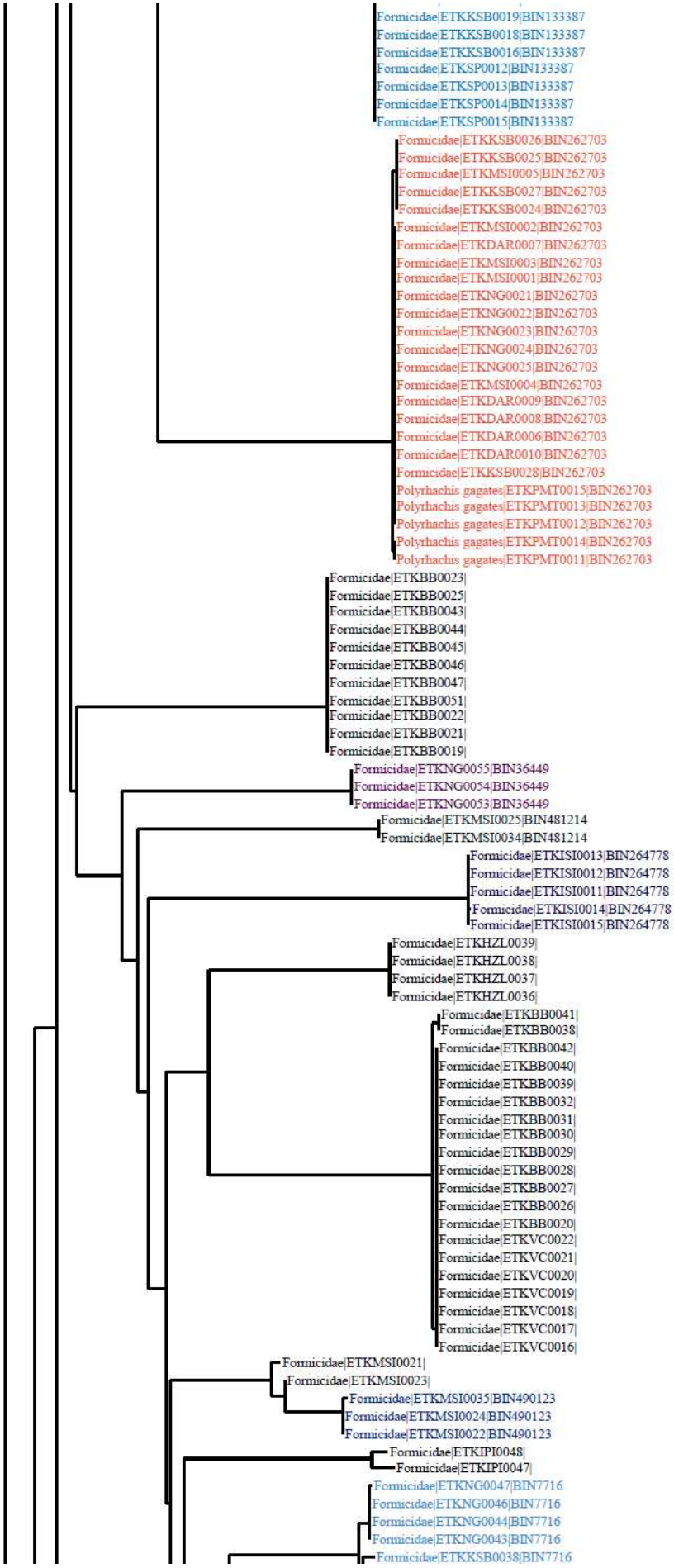
Zwickl, D.J. (2006). Genetic algorithm application for the phylogenetic analyses of large biological sequence datasets under maximum likelihood (University of Texas).

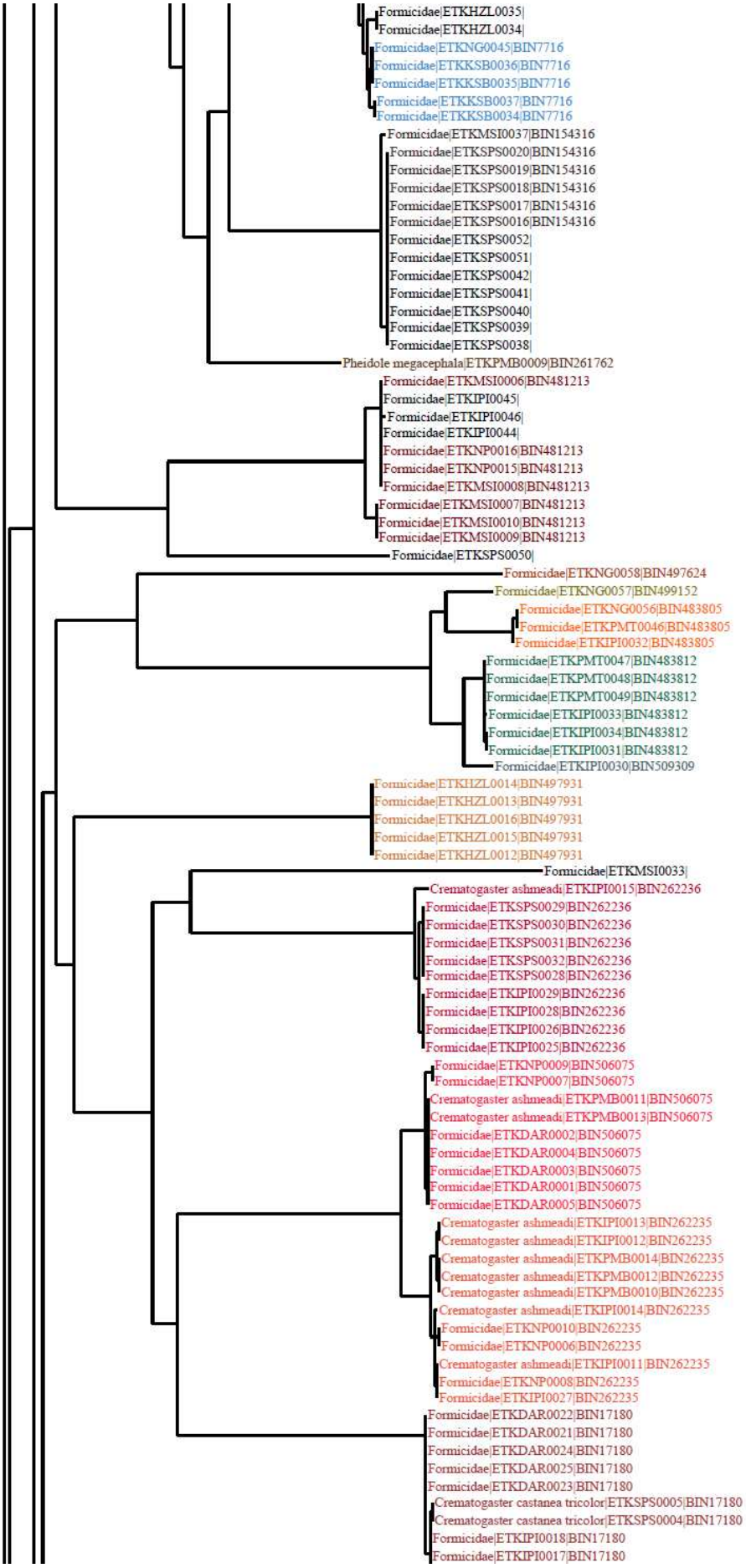
7. Appendix 1 – Neighbour-joining tree (K2P parameter) of all 624 specimens of ants represented in the DNA barcode library for eThekwini.

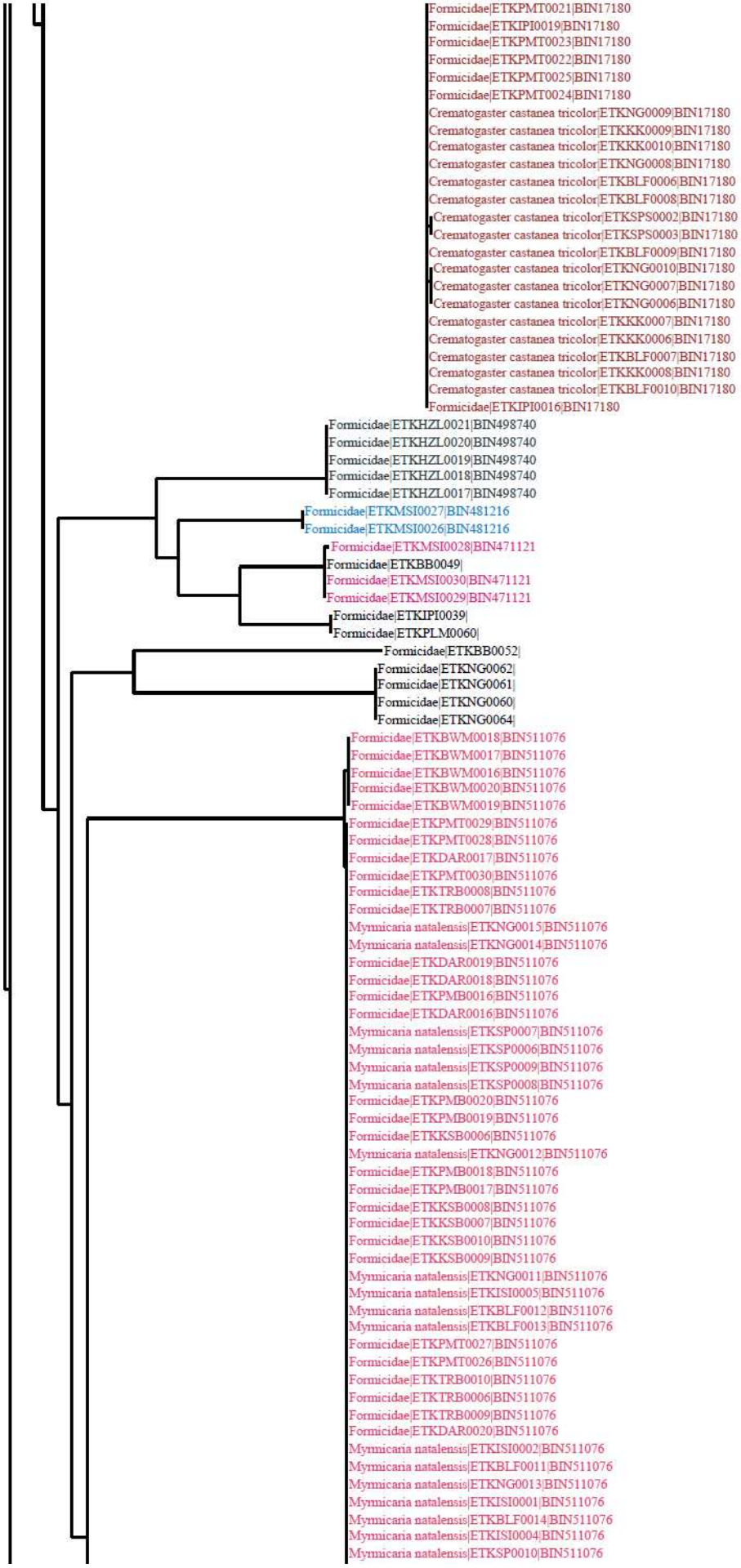


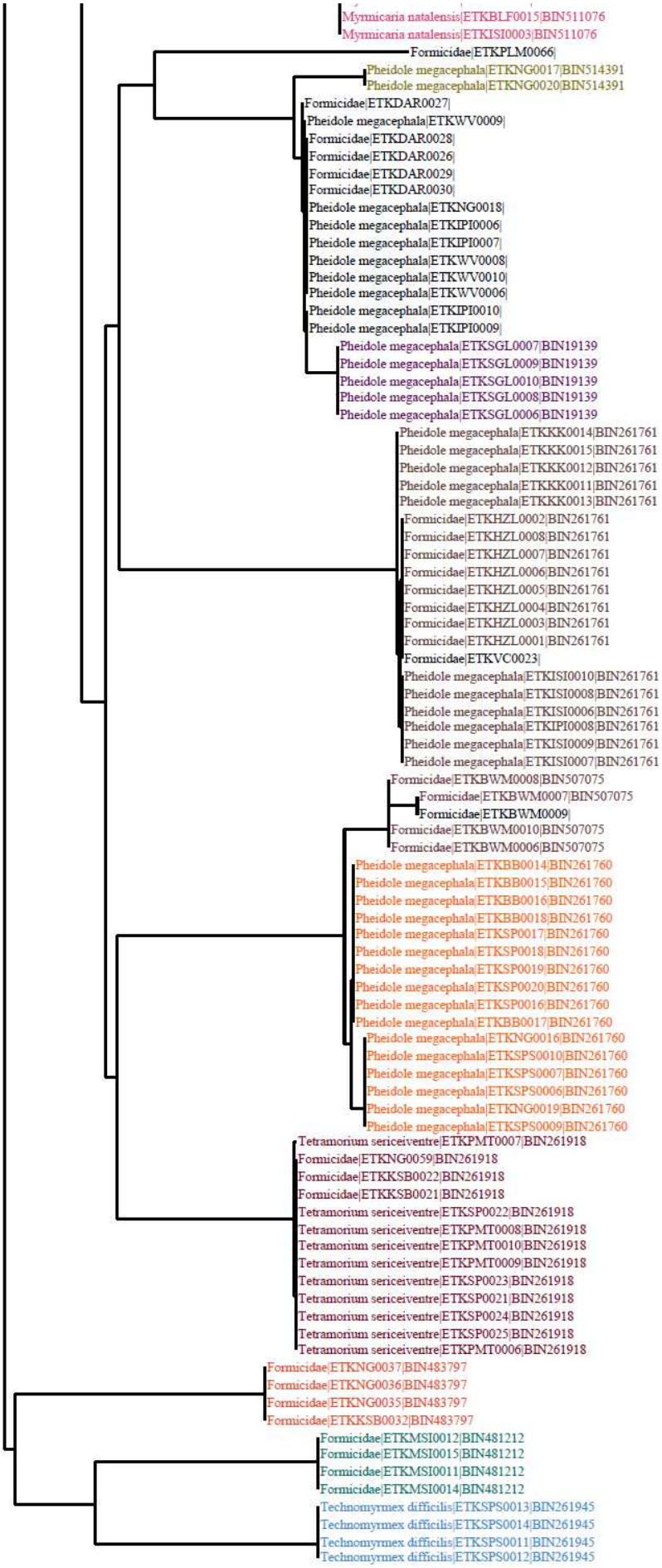

Formicidae|ETKTRB0002|BIN264250
Formicidae|ETKTRB0003|BIN264250
Formicidae|ETKTRB0001|BIN264250
Formicidae|ETKBWM0015|BIN264250
Formicidae|ETKKSB0004|BIN264250
Formicidae|ETKKSB0003|BIN264250
Formicidae|ETKKSB0002|BIN264250
Formicidae|ETKKSB0001|BIN264250
Lepisiota|ETKBB0004|BIN264250
Lepisiota|ETKBB0001|BIN264250
Lepisiota|ETKIPI0005|BIN264250
Lepisiota|ETKIPI0004|BIN264250
Lepisiota|ETKIPI0003|BIN264250
Lepisiota|ETKIPI0002|BIN264250
Lepisiota|ETKIPI0001|BIN264250
Lepisiota|ETKNG0005|BIN264250
Lepisiota|ETKWV0005|BIN264250
Lepisiota|ETKWV0004|BIN264250
Lepisiota|ETKWV0003|BIN264250
Lepisiota|ETKWV0002|BIN264250
Formicidae|ETKTRB0005|BIN264250
Formicidae|ETKBWM0011|BIN264250
Lepisiota|ETKPMB0003|BIN264250
Lepisiota|ETKPMB0002|BIN264250
Lepisiota|ETKBLF0004|BIN264250
Lepisiota|ETKBLF0003|BIN264250
Lepisiota|ETKPMB0001|BIN264250
Lepisiota|ETKBLF0005|BIN264250
Lepisiota|ETKKK0005|BIN264250
Lepisiota|ETKBLF0002|BIN264250
Lepisiota|ETKPMB0004|BIN264250
Lepisiota|ETKBLF0001|BIN264250
Formicidae|ETKPMT0016|BIN264250
Formicidae|ETKPMT0019|BIN264250
Lepisiota|ETKBB0003|BIN264250
Lepisiota|ETKSP0005|BIN264250
Lepisiota|ETKBB0002|BIN264250
Lepisiota|ETKSGL0004|BIN264250
Lepisiota|ETKSGL0002|BIN264250
Lepisiota|ETKNG0002|BIN264250
Lepisiota|ETKNG0004|BIN264250
Lepisiota|ETKNG0001|BIN264250
Lepisiota|ETKWV0001|BIN264250
Lepisiota|ETKSP0004|BIN264250
Lepisiota|ETKSP0003|BIN264250
Lepisiota|ETKSP0002|BIN264250
Lepisiota|ETKSP0001|BIN264250
Lepisiota|ETKKK0004|BIN264250
Lepisiota|ETKKK0003|BIN264250
Lepisiota|ETKKK0002|BIN264250
Lepisiota|ETKKK0001|BIN264250
Formicidae|ETKVC0008|BIN494813
Formicidae|ETKVC0006|BIN494813
Formicidae|ETKVC0005|BIN494813
Formicidae|ETKVC0004|BIN494813
Lepisiota|ETKNG0003|BIN264251
Formicidae|ETKVC0007|BIN515769
Formicidae|ETKMSI0044|BIN515769
Formicidae|ETKMSI0043|BIN515769
Formicidae|ETKMSI0042|BIN515769
Formicidae|ETKMSI0041|BIN515769
Formicidae|ETKMSI0040|BIN515769
Formicidae|ETKMSI0038|BIN515769
Formicidae|ETKMSI0039|BIN515769
Lepisiota|ETKPMB0005|BIN264252
Formicidae|ETKPLM0065|
Formicidae|ETKHZL0026|BIN35746
Formicidae|ETKPMT0042|BIN483803
Formicidae|ETKPMT0043|BIN483803
Formicidae|ETKPMT0044|BIN483803
Formicidae|ETKPMT0045|BIN483803
Formicidae|ETKHZL0022|BIN483803
Formicidae|ETKHZL0027|BIN483803
Formicidae|ETKHZL0028|BIN483803
Formicidae|ETKHZL0029|BIN483803
Formicidae|ETKHZL0031|BIN483803
Formicidae|ETKPMT0041|BIN483803
Formicidae|ETKNG0032|BIN483814
Formicidae|ETKNG0031|BIN483814
Formicidae|ETKNG0030|BIN483814
Formicidae|ETKNG0029|BIN483814
Formicidae|ETKNG0028|BIN483814
Formicidae|ETKNG0027|BIN483814
Formicidae|ETKNG0034|BIN483814
Formicidae|ETKNG0033|BIN483814
Formicidae|ETKNG0026|BIN483814
Formicidae|ETKIPI0037|
Formicidae|ETKIPI0036|
Formicidae|ETKIPI0035|


Formicidae|ETKPMT0056|BIN483814
Formicidae|ETKPMT0054|BIN483814
Formicidae|ETKPMT0053|BIN483814
Formicidae|ETKPMT0052|BIN483814
Formicidae|ETKPMT0051|BIN483814
Formicidae|ETKPMT0055|BIN483814
Formicidae|ETKPMT0050|BIN483814
Formicidae|ETKPMB0025|BIN515500
Formicidae|ETKPMB0023|BIN515500
Formicidae|ETKPMB0022|BIN515500
Formicidae|ETKAM0009|BIN515500
Formicidae|ETKAM0008|BIN515500
Formicidae|ETKAM0007|BIN515500
Formicidae|ETKAM0006|BIN515500
Formicidae|ETKNG0040|BIN515500
Formicidae|ETKNG0039|BIN515500
Formicidae|ETKNG0042|BIN515500
Formicidae|ETKNG0041|BIN515500
Formicidae|ETKNG0038|BIN515500
Formicidae|ETKBWM0003|BIN515500
Formicidae|ETKBWM0005|BIN515500
Formicidae|ETKBWM0001|BIN515500
Formicidae|ETKBWM0004|BIN515500
Formicidae|ETKBWM0002|BIN515500
Formicidae|ETKKSB0029|BIN469861
Formicidae|ETKKSB0017|BIN469861
Formicidae|ETKPLM0064|
Formicidae|ETKPLM0063|
Formicidae|ETKPLM0062|
Formicidae|ETKPLM0061|
Formicidae|ETKSPS0034|BIN500732
Formicidae|ETKVC0014|BIN133907
Formicidae|ETKAM0010|BIN133907
Formicidae|ETKSPS0036|BIN133278
Formicidae|ETKSPS0035|BIN133278
Formicidae|ETKVC0010|BIN133278
Formicidae|ETKSPS0037|BIN133278
Formicidae|ETKVC0013|BIN133278
Formicidae|ETKVC0012|BIN133278
Formicidae|ETKVC0011|BIN133278
Formicidae|ETKVC0009|BIN133278
Camponotus|ETKPMT0004|BIN8304
Camponotus|ETKPMT0003|BIN8304
Camponotus|ETKPMT0005|BIN8304
Formicidae|ETKPMT0059|BIN8304
Formicidae|ETKPMT0058|BIN8304
Formicidae|ETKPMT0057|BIN8304
Formicidae|ETKRID0005|BIN8304
Formicidae|ETKRID0004|BIN8304
Formicidae|ETKRID0003|BIN8304
Formicidae|ETKRID0002|BIN8304
Camponotus|ETKPMT0001|BIN8304
Camponotus|ETKPMT0002|BIN8304
Formicidae|ETKRID0001|BIN8304
Formicidae|ETKNG0052|BIN133898
Formicidae|ETKNG0051|BIN133898
Formicidae|ETKNG0050|BIN133898
Formicidae|ETKNG0049|BIN133898
Formicidae|ETKNG0048|BIN133898
Formicidae|ETKPMT0037|
Formicidae|ETKPMT0036|BIN133387
Formicidae|ETKSPS0033|BIN133387
Formicidae|ETKPSH0015|BIN133387
Formicidae|ETKPSH0014|BIN133387
Formicidae|ETKPSH0013|BIN133387
Formicidae|ETKPSH0012|BIN133387
Formicidae|ETKPSH0011|BIN133387
Formicidae|ETKDAR0014|BIN133387
Formicidae|ETKDAR0012|BIN133387
Formicidae|ETKDAR0011|BIN133387
Formicidae|ETKDAR0013|BIN133387
Formicidae|ETKDAR0015|BIN133387
Formicidae|ETKBB0009|BIN133387
Formicidae|ETKBB0010|BIN133387
Formicidae|ETKBB0011|BIN133387
Formicidae|ETKBB0012|BIN133387
Formicidae|ETKBB0013|BIN133387
Formicidae|ETKSPS0045|
Formicidae|ETKSPS0044|
Formicidae|ETKSPS0043|
Formicidae|ETKSP0011|BIN133387
Formicidae|ETKNP0005|BIN133387
Formicidae|ETKNP0004|BIN133387
Formicidae|ETKNP0003|BIN133387
Formicidae|ETKNP0002|BIN133387
Formicidae|ETKNP0001|BIN133387
Formicidae|ETKPMT0040|BIN133387
Formicidae|ETKPMT0039|BIN133387
Formicidae|ETKPMT0038|BIN133387
Formicidae|ETKKSB0020|BIN133387


Formicidae|ETKKSB0019|BIN133387
Formicidae|ETKKSB0018|BIN133387
Formicidae|ETKKSB0016|BIN133387
Formicidae|ETKSP0012|BIN133387
Formicidae|ETKSP0013|BIN133387
Formicidae|ETKSP0014|BIN133387
Formicidae|ETKSP0015|BIN133387
Formicidae|ETKKSB0026|BIN262703
Formicidae|ETKKSB0025|BIN262703
Formicidae|ETKMSI0005|BIN262703
Formicidae|ETKKSB0027|BIN262703
Formicidae|ETKKSB0024|BIN262703
Formicidae|ETKMSI0002|BIN262703
Formicidae|ETKDAR0007|BIN262703
Formicidae|ETKMSI0003|BIN262703
Formicidae|ETKMSI0001|BIN262703
Formicidae|ETKNG0021|BIN262703
Formicidae|ETKNG0022|BIN262703
Formicidae|ETKNG0023|BIN262703
Formicidae|ETKNG0024|BIN262703
Formicidae|ETKNG0025|BIN262703
Formicidae|ETKMSI0004|BIN262703
Formicidae|ETKDAR0009|BIN262703
Formicidae|ETKDAR0008|BIN262703
Formicidae|ETKDAR0006|BIN262703
Formicidae|ETKDAR0010|BIN262703
Formicidae|ETKKSB0028|BIN262703
Polyrhachis gagates|ETKPMT0015|BIN262703
Polyrhachis gagates|ETKPMT0013|BIN262703
Polyrhachis gagates|ETKPMT0012|BIN262703
Polyrhachis gagates|ETKPMT0014|BIN262703
Polyrhachis gagates|ETKPMT0011|BIN262703
Formicidae|ETKBB0023|
Formicidae|ETKBB0025|
Formicidae|ETKBB0043|
Formicidae|ETKBB0044|
Formicidae|ETKBB0045|
Formicidae|ETKBB0046|
Formicidae|ETKBB0047|
Formicidae|ETKBB0051|
Formicidae|ETKBB0022|
Formicidae|ETKBB0021|
Formicidae|ETKBB0019|
Formicidae|ETKNG0055|BIN36449
Formicidae|ETKNG0054|BIN36449
Formicidae|ETKNG0053|BIN36449
Formicidae|ETKMSI0025|BIN481214
Formicidae|ETKMSI0034|BIN481214
Formicidae|ETKISI0013|BIN264778
Formicidae|ETKISI0012|BIN264778
Formicidae|ETKISI0011|BIN264778
Formicidae|ETKISI0014|BIN264778
Formicidae|ETKISI0015|BIN264778
Formicidae|ETKHZL0039|
Formicidae|ETKHZL0038|
Formicidae|ETKHZL0037|
Formicidae|ETKHZL0036|
Formicidae|ETKBB0041|
Formicidae|ETKBB0038|
Formicidae|ETKBB0042|
Formicidae|ETKBB0040|
Formicidae|ETKBB0039|
Formicidae|ETKBB0032|
Formicidae|ETKBB0031|
Formicidae|ETKBB0030|
Formicidae|ETKBB0029|
Formicidae|ETKBB0028|
Formicidae|ETKBB0027|
Formicidae|ETKBB0026|
Formicidae|ETKBB0020|
Formicidae|ETKVC0022|
Formicidae|ETKVC0021|
Formicidae|ETKVC0020|
Formicidae|ETKVC0019|
Formicidae|ETKVC0018|
Formicidae|ETKVC0017|
Formicidae|ETKVC0016|
Formicidae|ETKMSI0021|
Formicidae|ETKMSI0023|
Formicidae|ETKMSI0035|BIN490123
Formicidae|ETKMSI0024|BIN490123
Formicidae|ETKMSI0022|BIN490123
Formicidae|ETKIPI0048|
Formicidae|ETKIPI0047|
Formicidae|ETKNG0047|BIN7716
Formicidae|ETKNG0046|BIN7716
Formicidae|ETKNG0044|BIN7716
Formicidae|ETKNG0043|BIN7716
Formicidae|ETKKSB0038|BIN7716


Formicidae|ETKHZL0035|
Formicidae|ETKHZL0034|
Formicidae|ETKNG0045|BIN7716
Formicidae|ETKKSB0036|BIN7716
Formicidae|ETKKSB0035|BIN7716
Formicidae|ETKKSB0037|BIN7716
Formicidae|ETKKSB0034|BIN7716
Formicidae|ETKMSI0037|BIN154316
Formicidae|ETKSPS0020|BIN154316
Formicidae|ETKSPS0019|BIN154316
Formicidae|ETKSPS0018|BIN154316
Formicidae|ETKSPS0017|BIN154316
Formicidae|ETKSPS0016|BIN154316
Formicidae|ETKSPS0052|
Formicidae|ETKSPS0051|
Formicidae|ETKSPS0042|
Formicidae|ETKSPS0041|
Formicidae|ETKSPS0040|
Formicidae|ETKSPS0039|
Formicidae|ETKSPS0038|
Pheidole megacephala|ETKPMB0009|BIN261762
Formicidae|ETKMSI0006|BIN481213
Formicidae|ETKIPI0045|
Formicidae|ETKIPI0046|
Formicidae|ETKIPI0044|
Formicidae|ETKNP0016|BIN481213
Formicidae|ETKNP0015|BIN481213
Formicidae|ETKMSI0008|BIN481213
Formicidae|ETKMSI0007|BIN481213
Formicidae|ETKMSI0010|BIN481213
Formicidae|ETKMSI0009|BIN481213
Formicidae|ETKSPS0050|
Formicidae|ETKNG0058|BIN497624
Formicidae|ETKNG0057|BIN499152
Formicidae|ETKNG0056|BIN483805
Formicidae|ETKPMT0046|BIN483805
Formicidae|ETKIPI0032|BIN483805
Formicidae|ETKPMT0047|BIN483812
Formicidae|ETKPMT0048|BIN483812
Formicidae|ETKPMT0049|BIN483812
Formicidae|ETKIPI0033|BIN483812
Formicidae|ETKIPI0034|BIN483812
Formicidae|ETKIPI0031|BIN483812
Formicidae|ETKIPI0030|BIN509309
Formicidae|ETKHZL0014|BIN497931
Formicidae|ETKHZL0013|BIN497931
Formicidae|ETKHZL0016|BIN497931
Formicidae|ETKHZL0015|BIN497931
Formicidae|ETKHZL0012|BIN497931
Formicidae|ETKMSI0033|
Crematogaster ashmeadi|ETKIPI0015|BIN262236
Formicidae|ETKSPS0029|BIN262236
Formicidae|ETKSPS0030|BIN262236
Formicidae|ETKSPS0031|BIN262236
Formicidae|ETKSPS0032|BIN262236
Formicidae|ETKSPS0028|BIN262236
Formicidae|ETKIPI0029|BIN262236
Formicidae|ETKIPI0028|BIN262236
Formicidae|ETKIPI0026|BIN262236
Formicidae|ETKIPI0025|BIN262236
Formicidae|ETKNP0009|BIN506075
Formicidae|ETKNP0007|BIN506075
Crematogaster ashmeadi|ETKPMB0011|BIN506075
Crematogaster ashmeadi|ETKPMB0013|BIN506075
Formicidae|ETKDAR0002|BIN506075
Formicidae|ETKDAR0004|BIN506075
Formicidae|ETKDAR0003|BIN506075
Formicidae|ETKDAR0001|BIN506075
Formicidae|ETKDAR0005|BIN506075
Crematogaster ashmeadi|ETKIPI0013|BIN262235
Crematogaster ashmeadi|ETKIPI0012|BIN262235
Crematogaster ashmeadi|ETKPMB0014|BIN262235
Crematogaster ashmeadi|ETKPMB0012|BIN262235
Crematogaster ashmeadi|ETKPMB0010|BIN262235
Crematogaster ashmeadi|ETKIPI0014|BIN262235
Formicidae|ETKNP0010|BIN262235
Formicidae|ETKNP0006|BIN262235
Crematogaster ashmeadi|ETKIPI0011|BIN262235
Formicidae|ETKNP0008|BIN262235
Formicidae|ETKIPI0027|BIN262235
Formicidae|ETKDAR0022|BIN17180
Formicidae|ETKDAR0021|BIN17180
Formicidae|ETKDAR0024|BIN17180
Formicidae|ETKDAR0025|BIN17180
Formicidae|ETKDAR0023|BIN17180
Crematogaster castanea tricolor|ETKSPS0005|BIN17180
Crematogaster castanea tricolor|ETKSPS0004|BIN17180
Formicidae|ETKIPI0018|BIN17180
Formicidae|ETKIPI0017|BIN17180


Formicidae|ETKPMT0021|BIN17180
Formicidae|ETKIPI0019|BIN17180
Formicidae|ETKPMT0023|BIN17180
Formicidae|ETKPMT0022|BIN17180
Formicidae|ETKPMT0025|BIN17180
Formicidae|ETKPMT0024|BIN17180
Crematogaster castanea tricolor|ETKNG0009|BIN17180
Crematogaster castanea tricolor|ETKKK0009|BIN17180
Crematogaster castanea tricolor|ETKKK0010|BIN17180
Crematogaster castanea tricolor|ETKNG0008|BIN17180
Crematogaster castanea tricolor|ETKBLF0006|BIN17180
Crematogaster castanea tricolor|ETKBLF0008|BIN17180
Crematogaster castanea tricolor|ETKSPS0002|BIN17180
Crematogaster castanea tricolor|ETKSPS0003|BIN17180
Crematogaster castanea tricolor|ETKBLF0009|BIN17180
Crematogaster castanea tricolor|ETKNG0010|BIN17180
Crematogaster castanea tricolor|ETKNG0007|BIN17180
Crematogaster castanea tricolor|ETKNG0006|BIN17180
Crematogaster castanea tricolor|ETKKK0007|BIN17180
Crematogaster castanea tricolor|ETKKK0006|BIN17180
Crematogaster castanea tricolor|ETKBLF0007|BIN17180
Crematogaster castanea tricolor|ETKKK0008|BIN17180
Crematogaster castanea tricolor|ETKBLF0010|BIN17180
Formicidae|ETKIPI0016|BIN17180
Formicidae|ETKHZL0021|BIN498740
Formicidae|ETKHZL0020|BIN498740
Formicidae|ETKHZL0019|BIN498740
Formicidae|ETKHZL0018|BIN498740
Formicidae|ETKHZL0017|BIN498740
Formicidae|ETKMSI0027|BIN481216
Formicidae|ETKMSI0026|BIN481216
Formicidae|ETKMSI0028|BIN471121
Formicidae|ETKBB0049|
Formicidae|ETKMSI0030|BIN471121
Formicidae|ETKMSI0029|BIN471121
Formicidae|ETKIPI0039|
Formicidae|ETKPLM0060|
Formicidae|ETKBB0052|
Formicidae|ETKNG0062|
Formicidae|ETKNG0061|
Formicidae|ETKNG0060|
Formicidae|ETKNG0064|
Formicidae|ETKBWM0018|BIN511076
Formicidae|ETKBWM0017|BIN511076
Formicidae|ETKBWM0016|BIN511076
Formicidae|ETKBWM0020|BIN511076
Formicidae|ETKBWM0019|BIN511076
Formicidae|ETKPMT0029|BIN511076
Formicidae|ETKPMT0028|BIN511076
Formicidae|ETKDAR0017|BIN511076
Formicidae|ETKPMT0030|BIN511076
Formicidae|ETKTRB0008|BIN511076
Formicidae|ETKTRB0007|BIN511076
Myrmicaria natalensis|ETKNG0015|BIN511076
Myrmicaria natalensis|ETKNG0014|BIN511076
Formicidae|ETKDAR0019|BIN511076
Formicidae|ETKDAR0018|BIN511076
Formicidae|ETKPMB0016|BIN511076
Formicidae|ETKDAR0016|BIN511076
Myrmicaria natalensis|ETKSP0007|BIN511076
Myrmicaria natalensis|ETKSP0006|BIN511076
Myrmicaria natalensis|ETKSP0009|BIN511076
Myrmicaria natalensis|ETKSP0008|BIN511076
Formicidae|ETKPMB0020|BIN511076
Formicidae|ETKPMB0019|BIN511076
Formicidae|ETKKSB0006|BIN511076
Myrmicaria natalensis|ETKNG0012|BIN511076
Formicidae|ETKPMB0018|BIN511076
Formicidae|ETKPMB0017|BIN511076
Formicidae|ETKKSB0008|BIN511076
Formicidae|ETKKSB0007|BIN511076
Formicidae|ETKKSB0010|BIN511076
Formicidae|ETKKSB0009|BIN511076
Myrmicaria natalensis|ETKNG0011|BIN511076
Myrmicaria natalensis|ETKISI0005|BIN511076
Myrmicaria natalensis|ETKBLF0012|BIN511076
Myrmicaria natalensis|ETKBLF0013|BIN511076
Formicidae|ETKPMT0027|BIN511076
Formicidae|ETKPMT0026|BIN511076
Formicidae|ETKTRB0010|BIN511076
Formicidae|ETKTRB0006|BIN511076
Formicidae|ETKTRB0009|BIN511076
Formicidae|ETKDAR0020|BIN511076
Myrmicaria natalensis|ETKISI0002|BIN511076
Myrmicaria natalensis|ETKBLF0011|BIN511076
Myrmicaria natalensis|ETKNG0013|BIN511076
Myrmicaria natalensis|ETKISI0001|BIN511076
Myrmicaria natalensis|ETKBLF0014|BIN511076
Myrmicaria natalensis|ETKISI0004|BIN511076
Myrmicaria natalensis|ETKSP0010|BIN511076


Myrmicaria natalensis|ETKBLF0015|BIN511076
Myrmicaria natalensis|ETKISI0003|BIN511076
Formicidae|ETKPLM0066|
Pheidole megacephala|ETKNG0017|BIN514391
Pheidole megacephala|ETKNG0020|BIN514391
Formicidae|ETKDAR0027|
Pheidole megacephala|ETKWV0009|
Formicidae|ETKDAR0028|
Formicidae|ETKDAR0026|
Formicidae|ETKDAR0029|
Formicidae|ETKDAR0030|
Pheidole megacephala|ETKNG0018|
Pheidole megacephala|ETKIPI0006|
Pheidole megacephala|ETKIPI0007|
Pheidole megacephala|ETKWV0008|
Pheidole megacephala|ETKWV0010|
Pheidole megacephala|ETKWV0006|
Pheidole megacephala|ETKIPI0010|
Pheidole megacephala|ETKIPI0009|
Pheidole megacephala|ETKSGL0007|BIN19139
Pheidole megacephala|ETKSGL0009|BIN19139
Pheidole megacephala|ETKSGL0010|BIN19139
Pheidole megacephala|ETKSGL0008|BIN19139
Pheidole megacephala|ETKSGL0006|BIN19139
Pheidole megacephala|ETKKK0014|BIN261761
Pheidole megacephala|ETKKK0015|BIN261761
Pheidole megacephala|ETKKK0012|BIN261761
Pheidole megacephala|ETKKK0011|BIN261761
Pheidole megacephala|ETKKK0013|BIN261761
Formicidae|ETKHZL0002|BIN261761
Formicidae|ETKHZL0008|BIN261761
Formicidae|ETKHZL0007|BIN261761
Formicidae|ETKHZL0006|BIN261761
Formicidae|ETKHZL0005|BIN261761
Formicidae|ETKHZL0004|BIN261761
Formicidae|ETKHZL0003|BIN261761
Formicidae|ETKHZL0001|BIN261761
Formicidae|ETKVC0023|
Pheidole megacephala|ETKISI0010|BIN261761
Pheidole megacephala|ETKISI0008|BIN261761
Pheidole megacephala|ETKISI0006|BIN261761
Pheidole megacephala|ETKIPI0008|BIN261761
Pheidole megacephala|ETKISI0009|BIN261761
Pheidole megacephala|ETKISI0007|BIN261761
Formicidae|ETKBWM0008|BIN507075
Formicidae|ETKBWM0007|BIN507075
Formicidae|ETKBWM0009|
Formicidae|ETKBWM0010|BIN507075
Formicidae|ETKBWM0006|BIN507075
Pheidole megacephala|ETKBB0014|BIN261760
Pheidole megacephala|ETKBB0015|BIN261760
Pheidole megacephala|ETKBB0016|BIN261760
Pheidole megacephala|ETKBB0018|BIN261760
Pheidole megacephala|ETKSP0017|BIN261760
Pheidole megacephala|ETKSP0018|BIN261760
Pheidole megacephala|ETKSP0019|BIN261760
Pheidole megacephala|ETKSP0020|BIN261760
Pheidole megacephala|ETKSP0016|BIN261760
Pheidole megacephala|ETKBB0017|BIN261760
Pheidole megacephala|ETKNG0016|BIN261760
Pheidole megacephala|ETKSPS0010|BIN261760
Pheidole megacephala|ETKSPS0007|BIN261760
Pheidole megacephala|ETKSPS0006|BIN261760
Pheidole megacephala|ETKNG0019|BIN261760
Pheidole megacephala|ETKSPS0009|BIN261760
Tetramorium sericeiventre|ETKPMT0007|BIN261918
Formicidae|ETKNG0059|BIN261918
Formicidae|ETKKSB0022|BIN261918
Formicidae|ETKKSB0021|BIN261918
Tetramorium sericeiventre|ETKSP0022|BIN261918
Tetramorium sericeiventre|ETKPMT0008|BIN261918
Tetramorium sericeiventre|ETKPMT0010|BIN261918
Tetramorium sericeiventre|ETKPMT0009|BIN261918
Tetramorium sericeiventre|ETKSP0023|BIN261918
Tetramorium sericeiventre|ETKSP0021|BIN261918
Tetramorium sericeiventre|ETKSP0024|BIN261918
Tetramorium sericeiventre|ETKSP0025|BIN261918
Tetramorium sericeiventre|ETKPMT0006|BIN261918
Formicidae|ETKNG0037|BIN483797
Formicidae|ETKNG0036|BIN483797
Formicidae|ETKNG0035|BIN483797
Formicidae|ETKKSB0032|BIN483797
Formicidae|ETKMSI0012|BIN481212
Formicidae|ETKMSI0015|BIN481212
Formicidae|ETKMSI0011|BIN481212
Formicidae|ETKMSI0014|BIN481212
Technomyrmex difficilis|ETKSPS0013|BIN261945
Technomyrmex difficilis|ETKSPS0014|BIN261945
Technomyrmex difficilis|ETKSPS0011|BIN261945
Technomyrmex difficilis|ETKSPS0012|BIN261945