Genetic diversity of populations of a Southern African millipede, *Bicoxidens flavicollis* (Diplopoda, Spirostreptida, Spirostreptidae)

by

Yevette Gounden
212502571

Submitted in fulfillment of the academic requirements for the degree of
Master of Science (Genetics)
School of Life Sciences,
University of KwaZulu-Natal
Westville campus

November 2018

As the candidate's supervisor I have/have not approved this thesis/dissertation for submission.

Signed: _____ Name: _____ Date: _____

ABSTRACT

The African millipede genus *Bicoxidens* is endemic to Southern Africa, inhabiting a variety of regions ranging from woodlands to forests. Nine species are known within the genus but *Bicoxidens flavicollis* is the most dominant and wide spread species found across Zimbabwe. *Bicoxidens flavicollis* individuals have been found to express phenotypic variation in several morphological traits. The most commonly observed body colours are brown and black. In the Eastern Highlands of Zimbabwe body colour ranges from orange- yellow to black, individuals from North East of Harare have a green-black appearance and a range in size (75–110 mm). There is disparity in body size which has been noted with individuals ranging from medium to large and displaying variation in the number of body rings. Although much morphological variation has been observed within this species, characterization based on gonopod morphology alone cannot distinguish or define variation between phenotypically distinct individuals. Morphological classification has been found to be too inclusive and hiding significant genetic variation. Taxa must be re-assessed with the implementation of DNA molecular methods to identify the variation between individuals. This study aimed to detect genetic divergence of *B. flavicollis* due to isolation by distance of populations across Zimbabwe. The mitochondrial DNA 16S and 12S rRNA genes were used to detect levels of genetic variation as mitochondrial markers express high variability making them suitable for phylogenetic studies. Sequence analysis of the 16S rRNA gene resulted in the generation of 22 haplotypes, derived from 42 sequences with strong haplotype diversity (H_d > 0.9). Analysis of Molecular Variance (AMOVA) analysis determined that variation among the populations was significantly greater (> 80 %) than the variation occurring within populations (< 12 %). A high fixation index (F_{ST} = 0.88229) indicated a high level of population genetic differentiation. With analysis of the 16S rRNA gene, *B. flavicollis* individuals demonstrated both distinctive phylogeographic diversity and genetic similarity for specific regions within Zimbabwe. Phylogenetic analyses using the 12S rRNA gene provided evidence of a more distinct genetic structure between localities. Nineteen haplotypes were derived from 19 sequences, which indicated a genetically distinct population structure (H_d = 1.000). The AMOVA analysis demonstrated that variation among the populations was greater (> 60 %) than the variation occurring within populations (< 40 %), although both were quite high. A low fixation index (F_{ST} = 0.37466) suggests a predominantly homozygous population structure. Both genes indicated distinctly structured populations, whilst the 16S rRNA also suggested the existence of closely clustered populations based on PCoA analyses, which is further supported by the presence of admixed haplotypes. The results are significant for *B. flavicollis* as a genetically diverse species. The findings of this study can be considered for future comparative research within the genus *Bicoxidens* or against other geographically distant genera. Additional markers such as those of nuclear origin, can be used along with mitochondrial markers to investigate and identify more diplopods which exhibit this level of

genetic divergence although belonging to a single species. Acquired knowledge and understanding from phylogeographic studies will provide researchers with greater taxonomic awareness.

Key words: *Bicoxidens flavicollis*; Fixation index; Haplotypes; Mitochondrial DNA; Morphological classification; Phylogenetic; Phylogeographic; Variation

PREFACE

The experimental work described in this thesis was completed by Yevette Gounden in the Discipline of Genetics, School of Life Sciences, College of Agriculture, Engineering and Science at the University of KwaZulu–Natal, Westville Campus, under the supervision of Dr Oliver Tendayi Zishiri and Prof. Tarombera Mwabvu.

These studies represent original work by the author and have not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of the work of others it is duly acknowledged in the text.

I certify that the above information is correct

_____

Dr Oliver Tendayi Zishiri (Supervisor)

Date:.....

PLAGIARISM DECLARATION

I, Yevette Gounden 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: ……

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 2

ACKNOWLEDGMENTS

I would firstly like to thank and extend my appreciation to my supervisors, Dr Oliver Tendayi Zishiri and Prof. Tarombera Mwabvu. I am fortunate enough to have worked with patient and supportive individuals who assisted and guided me at every step when help was needed. Dr Zishiri's laugher and words of encouragement during stressful times made the duration of this MSc degree enjoyable and possible and of course without Prof. Mwabvu I would not have had a research project which introduced me to a new area of interest and allowed me to gain skills and knowledge which will be useful in future endeavours.

I acknowledge and thank the National Research Foundation for financial assistance during my two years of study. The scholarship has been most helpful and lifted a weight of my shoulders during a stressful time.

I am thankful to my peers at the Department of Genetics for their kind words and assistance but most importantly, their friendship. Shirwin Pillay, Leah Maharaj, Kerisha Govender and Natalie Naidoo, kept me sane during this time and always lent a helping hand when needed. The friendships formed and grown over the years will remain for a lifetime.

I would like to thank my parents, Marlene and Dan Gounden for their continuous support and sacrifice to get me to where I am today. I will never have enough words to express my appreciation for you guys.

CHAPTER 1
GENERAL INTRODUCTION

1.1 Background

Biodiversity is often used as a measurement to monitor and assess the health of ecosystems and is taken into great consideration when planning conservation efforts (Morrison *et al*., 2012). One or more taxa can be used as an indicator for the determination of the diversity of larger taxonomic groups. This has been acknowledged to be useful especially when studying invertebrate fauna, which can be found in abundance in various environments (Morrison *et al*., 2012). Cardosa *et al*. (2011) reported that invertebrates account for 80 % of all described species across the globe and are dominant in species richness and biomass. However, although their presence is prominent in all domains, there is still a lack of literature to accompany it (Cardosa *et al*., 2011). Such is especially true when investigating the arthropod class, Diplopoda, commonly referred to as millipedes, which is understudied (Brewer *et al*., 2012).

Millipedes are terrestrial invertebrates which form an essential component of the soil environment (Stašiov *et al*., 2012). The term millipede refers to the large number of legs these arthropods possess (Fontanetti *et al*., 2002). As saprophages, millipedes partake in the fragmentation and transformation of dead organic matter, making use of their bodily movements to move the decomposing matter along with mineral soil components deeper into the ground (Sierwald and Bond, 2007, Stašiov *et al*., 2012). In Southern Africa, 39 % of litter standing crop is consumed by millipedes and they play a key role in accelerating nutrient availability for plants (Dangerfield and Milner, 1996; Vohland and Hamer, 2013) Millipedes are commonly found in both tropical and temperate areas (Sierwald and Bond, 2007). More than 12 000 millipede species have been described (Sierwald and Bond, 2007). The existing high level of classification is directed towards aiding with the identification rather than analysing natural groupings. This is due to a paucity of well supported phylogenetic studies using optimal criterion (Brewer *et al*., 2012, Brewer and Bond, 2013).

Millipedes have rather narrow distributions with individual species seldom occupying a wide–ranging area (Redman and Hamer, 2003). Low dispersal ranges and habitat transformation occurring at exponential rates, conservation mitigation measures are required to ensure the survival of the species (McGeoch *et al*., 2011). In Africa, the biogeography of millipede fauna is vaguely understood, however, that of South Africa is considered to be well understood scientifically with approximately two hundred or more species found exclusively in the province of KwaZulu– Natal (Redman and Hamer, 2003, McGeoch *et al*., 2011). There are a vast number of Southern African millipedes which are not yet known to science due to limited assessment of habitats, and minimal analysis and study of preserved specimens in museums (Mwabvu *et al*., 2009). It is because of the previously mentioned reasons that Mwabvu *et al*. (2009) stated that our knowledge of millipede diversity and distribution in

the southern African region is far from complete. Spirostreptidans have been recognized as the most frequently encountered diplopod group in southern Africa, with the genus *Bicoxidens* being endemic to the area (Hamer, 1999, Tinago *et al*., 2017). *Bicoxidens* species are dominant in the Eastern, Southern and Central regions of Zimbabwe, however, there is a lack of genetic based studies on the genus (Mwabvu *et al*., 2009, Tinago *et al*., 2017).

1.2 Justification

This study was undertaken to broaden our knowledge on *Bicoxidens*, focusing on *Bicoxidens flavicollis* from various localities of Zimbabwe. There is a dearth in literature pertaining to the species, with very limited DNA molecular studies having been undertaken and reported. The present study will contribute to existing knowledge by utilising mitochondrial DNA markers to distinguish the phylogenetic background of the species based on one site geographic location. It is essential to unravel the cryptic nature of the species as it will be beneficial in future identification and conservation efforts.

1.3 Hypothesis

It was hypothesised that DNA molecular based methods of identification would categorise individuals belonging to the genus *Bicoxidens* (Spirostreptidae) more efficiently than morphology based methods and *Bicoxidens flavicollis* will exhibit phylogenetic diversity whilst being a complex species.

1.4 Aims

1.4.1 To establish the existence of genetic divergence between populations of *B. flavicollis*

1.4.2 To use statistical and phylogenetic analysis to determine the genetic diversity present between populations.

1.5 Objectives

1.5.1 To elute DNA from specimens using Quick-DNA™ Miniprep kit

1.5.2 To amplify the 12S and 16S rRNA genes within the selected sample pool using PCR

1.5.3 To conduct sequencing on all samples that positively amplified the genes

1.5.4 To use statistical analyses to determine the presence of polymorphisms with *B. flavicollis* populations and

1.5.5 To generate haplotype networks, principle coordinates analysis and distance matrices to examine the relationship between *B. flavicollis* individuals/ populations.

1.6 References

Brewer, M. and Bond, J. (2013). Ordinal– Level Phylogenomics of the Arthropod Class Diplopoda (Millipedes) Based on an Analysis of 221 Nuclear Protein– Coding Loci Generated Using Next– Generation Sequence Analyses. *PLoS ONE*, 8(11): p.e79935.

Brewer, M., Sierwald, P. and Bond, J. (2012). Millipede Taxonomy after 250 Years: Classification and Taxonomic Practices in a Mega– Diverse yet Understudied Arthropod Group. *PLoS ONE*, 7(5): p.e37240.

Cardoso, P., Erwin, T., Borges, P. and New, T. (2011). The seven impediments in invertebrate conservation and how to overcome them. *Biological Conservation*, 144(11): 2647– 2655.

Dangerfield, J. and Milner, A. (1996). Millipede Fecal Pellet Production in Selected Natural and Managed Habitats of Southern Africa: Implications for Litter Dynamics. *Biotropica*, 28(1): 113– 120.

Fontanetti, C., Campos, K., Prado, R. and da Silva Souza, T. (2002). Cytogenetic Studies in Diplopoda. *Cytologia*, 67(3): 253– 260.

Hamer, M. (1999). An illustrated key to the spirostreptidan (Diplopoda: Spirostreptida) genera of Southern Africa. *Annals of the Natal Museum*, 40: 1– 22.

McGeoch, M., Sithole, H., Samways, M., Simaika, J., Pryke, J., Picker, M., Uys, C., Armstrong, A., Dippenaar– Schoeman, A., Engelbrecht, I., Braschler, B. and Hamer, M. (2011). Conservation and monitoring of invertebrates in terrestrial protected areas. *Koedoe – African Protected Area Conservation and Science*, 53(2): 1–13.

Morrison, W., Waller, J., Brayshaw, A., Hyman, D., Johnson, M. and Fraser, A. (2012). Evaluating Multiple Arthropod Taxa as Indicators of Invertebrate Diversity in Old Fields. *The Great Lakes Entomologist*, 45(1): 56– 57.

Mwabvu, T., Hamer, M. and Slotow, R. (2009). A new species of *Bicoxidens* Attems, 1928 (Diplopoda, Spirostreptida, Spirostreptidae) from northern Zimbabwe. *ZooKeys*, 7: 75– 81.

Redman, G. and Hamer, M. (2003). The distribution of Southern African Harpagophoridae Attems, 1909 (Diplopoda: Spirostreptida). *African Invertebrates*, 44(1): 213–226.

Sierwald, P. and Bond, J. (2007). Current Status of the Myriapod Class Diplopoda (Millipedes): Taxonomic Diversity and Phylogeny. *Annual Review of Entomology*, 52(1): 401– 420.

Stašiov, S., Stašiová, A., Svitok, M., Michalková, E., Slobodník, B. and Lukáčik, I. (2012). Millipede (Diplopoda) communities in an arboretum: Influence of tree species and soil properties. *Biologia*, 67(5): 945 – 946.

Tinago, T., Mwabvu, T. and MacDonald, A. (2017). Evidence of multiple divergent mitochondrial lineages within the southern African diplopod genus *Bicoxidens* Attems, 1928 (Spirostreptida). *African Zoology*, 52(4): 229– 235.

Vohland, K. and Hamer, M. (2013). A Review of the Millipedes (Diplopoda) of Namibia, with Identification Keys and Descriptions of Two New Genera and Five New Species. *African Invertebrates*, 54(1): 251– 304.

CHAPTER 2
LITERATURE REVIEW

This chapter focuses on the literature highlighting various aspects of Diplopods, focusing on their global distribution and the environments which they inhabit. Methods currently employed for identification purposes, both morphological and molecular based are reviewed. Primary focus was given to the family Spirostreptidae regarding their presence in and around Southern Africa. Emphasis was directed towards the species, *Bicoxidens flavicollis*, highlighting the need for further genetic and taxonomic studies of this species.

2.1 Taxonomy of millipedes

The first comprehensive review of the Myriapoda subphylum was undertaken in 1997, and focused on fossil records to draw conclusions regarding the evolution of the subphylum (Shear and Edgecombe, 2010). Myriapoda is a terrestrial subphylum of Arthropoda which consists of four classes, namely; Diplopoda (millipedes), Chilopoda (centipedes), Pauropoda (pauropods) and Symphyla (symphylids) (Shear and Edgecombe, 2010; Miyazawa *et al*., 2014). There are four recognized modes of development (epimorphosis, euanamorphosis, hemianamorphosis and teloanamorphosis), during which molting and the addition of body segments occur (Miyazawa *et al*., 2014). Myriapods within the same order may have the same post-embryonic development, however, it is not conserved at class level (Miyazawa *et al*., 2014). In epimorphic (epimorphosis) development, there is no observed addition of new segments as the full body complement is produced during embryogenesis, hence individuals molt but remain with the same number of segments (Fusco, 2005; Minelli and Fusco, 2013; Miyazawa *et al*., 2014). The remaining three modes of development are known as anamorphic development, as individuals acquire their full set of body segments is reached at a later stage (Fusco, 2005). During euanamorphosis, new segments are added after each molt. The number of new segments added each time can be greater than or equal to one (Fusco, 2005; Miyazawa *et al*., 2014). Individuals will continue to add new segments until the final molt, with an unknown terminal number (Fusco, 2005). Hemianamorphosis is characterized by the addition of new segments during initial molts, thereafter once the final maximum number of segments is reached, growth continues with further molts however there is no addition of new body segments (Fusco, 2005; Miyazawa *et al*., 2014). During teloanamorphosis, once the maximum number of segments is attained, molting and the addition of further segments cease and will no longer occur thereafter (Fusco, 2005; Miyazawa *et al*., 2014).

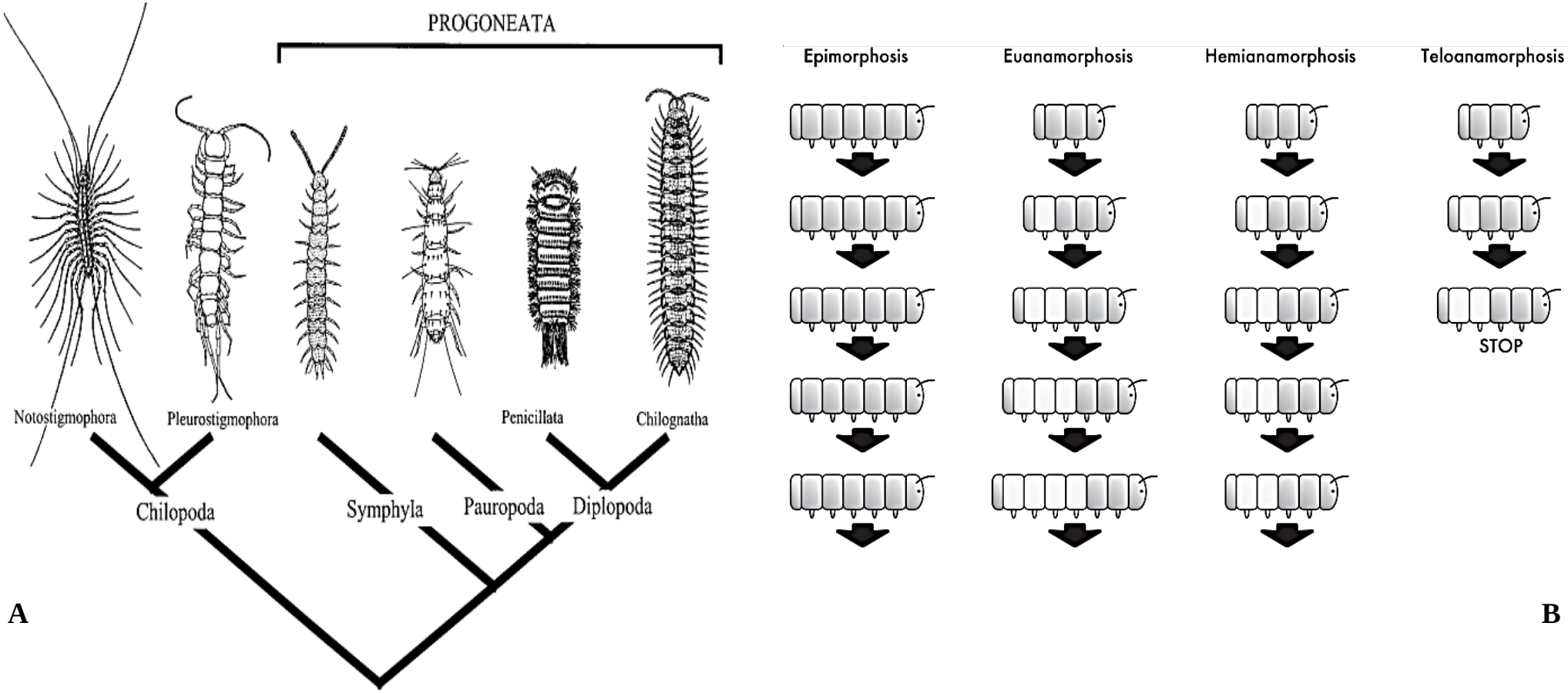

Figure 2.1: A: Diagrammatic representation of the relationships within the subphylum Myriapoda based on morphological features. Chilopoda and Diplopoda are further divided into sub-classes. Progoneata is indicative of the genital apertures being located near the anterior end of the body (Edgecombe and Giribet, 2002). **B:** Diagrammatic representation of the four modes of post-embryonic development which occur in Myriapods (Miyazawa *et al*., 2014).

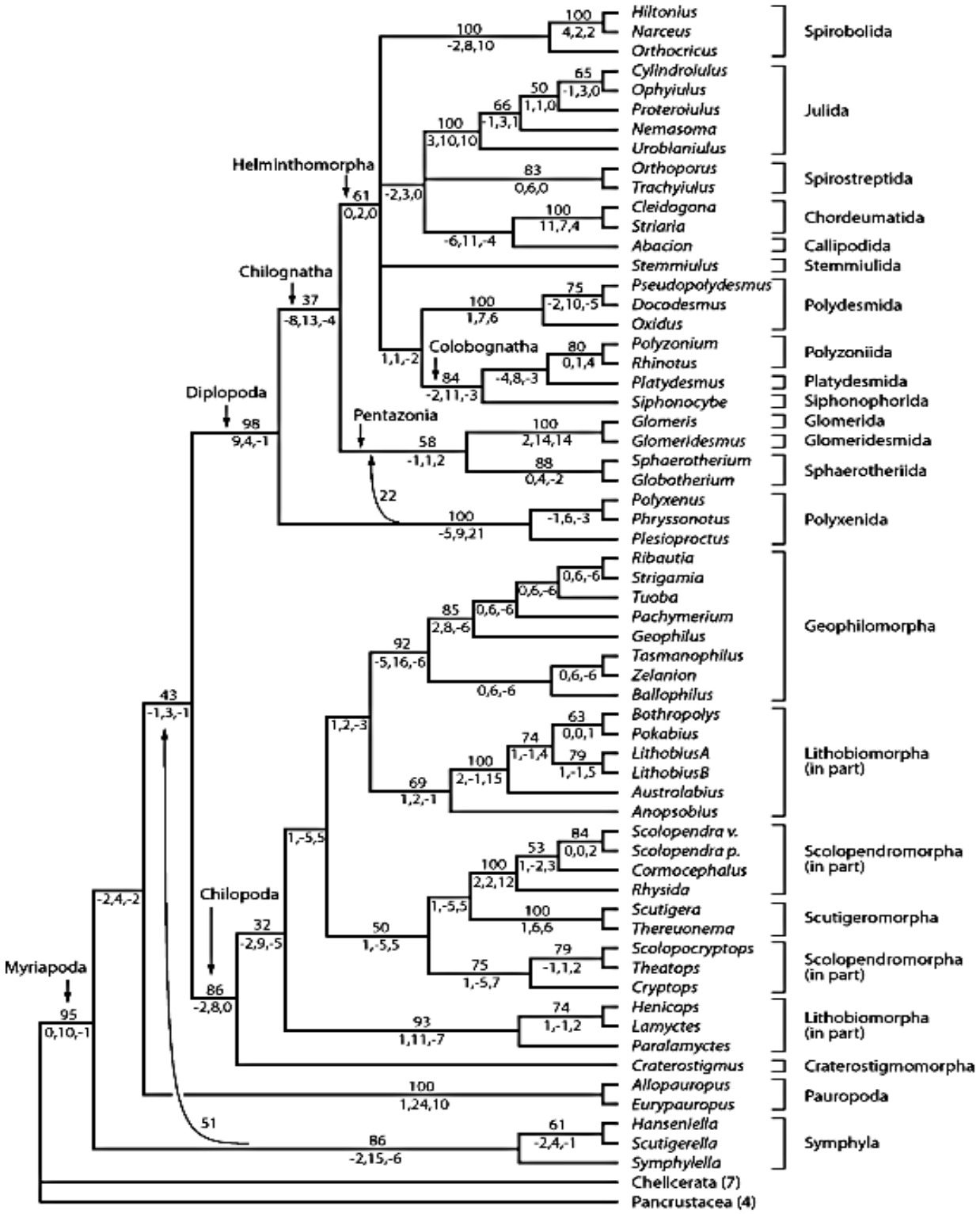

Figure 2.2: Phylogenetic representation of Myriapod relationships derived from parsimony analysis of three nuclear protein-encoding genes; elongation factor-1 (EF-1), elongation factor-2 (EF-2) and RNA polymerase II (Pol II) (Regeir *et al*., 2005).

Myriapods were originally considered to be closely related to hexapods based on similarities in developmental and morphological features (Edgecombe and Giribet, 2002). DNA molecular based studies have since debunked this conclusion and demonstrated hexapods to share great affinity with crustaceans (Paulus, 2000; Edgecombe and Giribet, 2002). Recent DNA molecular studies suggests that myriapoda is the sister group to pancrustacea, within a monophyletic mandibulata (Kenny *et al*., 2015; Robertson *et al*., 2015). Myriapoda is a diverse subphylum, with approximately 12 000 species described amongst the four classes (Sierwald and Bond, 2007). Despite this diversity, only a single myriapod genome has been released (*Strigamia maritime*), limiting the number of inferences in myriapod biology which can be made (Chipman *et al*., 2014; Kenny *et al*., 2015). Diplopods have great diversity amongst arthropods and are the most important consumers of detritus in most terrestrial ecosystems (Golovatch and Kime, 2009). The oldest fossil of a millipede dates back to 428 million years ago. To date, eight orders of millipedes are considered to be extinct (Nelson and Artomova, 2009; Kenny *et al*., 2015). The taxonomic productivity of millipedes has varied throughout history with as many as 300 species described in a single year (Brewer *et al*., 2012; Golovatch and Kime, 2009).

Characterizing species is of great scientific significance because it influences comparative, phylogenetic and diversification studies. It is therefore, common to hear of on-going debates regarding the classification of species (Bond and Sierwald, 2002). To date, morphological data (including genitalia/ gonopods) have been commonly used to delimit species (Wiens and Penkrot, 2002) and is still used currently in delimiting arthropods. Species rich taxa are known to display significant divergence in traits which are related to reproduction (Wojcieszek and Simmons, 2013). Although very little is known about variation in genitalia, it continues to play a vital role as a species-specific diagnostic tool for numerous arthropod groups (Bond *et al*., 2003). According to Miley (1925) the gonopod of male millipedes are of major diagnostic importance in the identification of species. Genitalia are not considered to be adaptive, however, they can evolve as a secondary consequence of ecological divergence, resulting from variation in environmental conditions which can have a direct effect on mating preferences and signals (Miley, 1925; Tanabe *et al*., 2001; Bond *et al*., 2003).

In male millipedes, legs 8, 8 and 9 or 9 and 10 are responsible for the transfer of sperm, known as the spermatopositors (Golovatch and Kime, 2009). The minor details in gonopods reflect differences between species (Golovatch and Kime, 2009; Wojcieszek and Simmons, 2013). Due to millipedes almost exclusively being identified and described based on the morphology of male genitalia, Diplopoda prove to be useful for exploring links between speciation and genital evolution (Wojcieszek and Simmons, 2013). The low vagility displayed by millipedes is a driving force toward geographic isolation, which ultimately contributes to genetic divergence and speciation (Tanabe *et al*., 2001; Wojcieszek and Simmons, 2013). Morphometric characterization is now being correlated with

molecular DNA methodologies as gonopod morphology alone may be unsuccessful in differentiating between genetically distinct species (Mwabvu *et al*., 2013).

2.2 Distribution of millipedes globally

Millipedes have a key role when it comes to the decomposition of organic matter in the soil (Reboleira and Enghoff, 2014). The population density of millipedes in forest soils can exceed 1000 individuals per square metre and it is estimated that 10–15 % of the leaf litter present is consumed by them (Golovatch and Kime, 2009). Millipedes are abundant in most habitats and can be found on six of the seven continents, except for Antarctica (Shelley and Golovatch, 2011). Numerous species around the world have adapted to subterranean life, with individuals being found in caves at almost 2000 m depth (Shelley and Golovatch, 2011; Reboleira and Enghoff, 2014; Kenny *et al*., 2015). European countries host the most relatively well-studied millipede faunas (Reboleira and Enghoff, 2014), however, there is still a paucity of scientific information about them.

Large areas on all continents except for Europe are in essence unsampled, as indicated by having considerably fewer samples or lacking a study collection altogether (Minelli, 1990; Shelley and Golovatch, 2011). The lack of studies associated with these areas, directly corresponds to unfavourable environmental conditions and in some instances a lack of funding, as is the case with most invertebrate studies. Conditions such as freezing arctic regions or arid deserts, which are undesirable for humans, will also not be desirable for millipedes to inhabit (Shelley and Golovatch, 2011; Reboleira and Enghoff, 2014; Donaldson *et al*., 2016). Indigenous millipedes from a number of areas still remain unstudied scientifically (Shelley, 2009). Millipedes have shown to be an ideal model for microendemic studies, as numerous species are confined to small geographic locations near other populations of closely related taxa. These observed distribution patterns are due to the inability of millipedes to travel long distances (Loria *et al*., 2011).

Due to their susceptibility to water deficit, millipedes are known to inhabit humid environments and are able to burrow into deeper soil during drier periods. It is of great concern that the destruction of forest areas will endanger millipede species. Those with smaller populations and limited distribution will be at a greater risk (Tole, 2002; Galanes and Thomlinson, 2011). Substantial changes in abiotic and biotic environmental factors along elevation gradients were found to influence the abundance, diversity and distribution of millipedes (Alagesan and Ramanathan, 2013; Golovatch and Wesener, 2016).

2.2.1 The distribution of millipedes in North and South America

In North America, millipedes are recognized as being the fourth largest group of troglobiotic (cave obligate) invertebrates (Culver *et al*., 2000; Loria *et al*., 2011). Such millipedes inhabit caves in both temperate and tropical areas. Orders including; Glomerida, Callipodida, Chordeumatida, Julida, Polydesmida and Spirostreptidae are known to be troglobiotic millipedes (Culver and Pipan, 2010).

Like most cave dwelling organisms, troglobiotic millipedes are fully dependent on resources being carried into caves from the surface (Culver and Pipan, 2010). Such resources which they may feed on include, bat guano, leaf litter and even scraping bacteria off rocks (Culver and Pipan, 2010; Loria *et al*., 2011). It has been reported that some species of troglobiotic millipedes are amphibious, with the ability to filter particulate matter from water sources (Culver and Pipan, 2010).

Iniesta and Ferreira (2015) reported that the troglobiotic fauna of Brazil is of great importance to cave conservation. Caves, which contain at least one species of troglobiotic species, are considered to be of maximum relevance and therefore, cannot be destroyed (Iniesta and Ferreira, 2015). Seven troglobiotic species are currently known from Brazil, among these seven, two species are known to be found in iron ore caves and limestone caves (Iniesta and Ferrerira, 2013). These species exhibit troglomorphic traits, which is recognized by a reduced number of ocelli, strong depigmentation and a relative decrease in body size compared to non-troglobiotic species (Iniesta and Ferreira, 2015). In Puerto Rico, millipedes have been well studied and characterized. Approximately 50 species have been described, with 39 of these species exhibiting endemism (Galanes and Thomlinson, 2011). Millipedes are abundant in the soil and litter in moist forests throughout the island of Puerto Rico, with a constant population year round (Galanes and Thomlinson, 2011).

2.2.2 The distribution of millipedes in Asia

Asia may be the largest continent, however the diplopoda fauna are still poorly described in this part of the world. Many studies carried out in Asia have been revisited, with more samples being collected (Likhitrakarn *et al*., 2011). The revisiting of such studies has demonstrated that the diversity of diplopoda found in Southern China is extraordinarily high. Similar to North and South America, records of fauna (including all arthropod groups) in China are strongly biased towards caves, with most species expected to be troglobiotic (Deharveng *et al*., 2008). Approximately five to six species, mainly members of the orders Chordeumatida and Callipodida can be found per a cave (Deharveng *et al*., 2008; Golovatch, 2015). The oriental region is the only global biogeographic realm to host all 16 orders of diplopoda, with 14 orders encountered in China and its adjacent parts (Golovatch, 2015). This leads to the belief that China may support no less than one thousand species of millipedes from various origins (Golovatch, 2015).

In the Asian part of Russia, the first diplopod to have been encountered was *Oxidus gracilis* on the Sakhalin Island and believed to have been of Southeast Asian origin (Nefediev *et al*., 2015). In the mainland of Russia, this species has previously been reported in green or hothouses in Khabarovsk and in the South West of Siberia (Nefediev *et al*., 2015). *Oxidus gracilis* is considered to be indigenous to East Asia, but can now be found free-living throughout tropical, subtropical and warm temperate areas (Nefediev *et al*., 2015). The species, *Julus ghilarovi* was found to have a widespread distribution in the following areas; Kemerovo, Novosibirsk, the Altai province, Republic of Altai,

Republic of Khakassia as well as the south of the Krasnoyarsk province (Nefedieva *et al*., 2015). This species can be found in various environments due to its high ecological plasticity, which allows it to inhabit different areas like small leaf litter areas, coniferous forests, meadows and stony tundras (Nefedieva *et al*., 2015).

Figure 2.3: Image of diplopod species *Oxidus gracilis* (Nguyen *et al*., 2017). A: full length body image. B: body segments.

Golovatch and Wesener (2016) reported that the millipede fauna of India is largely diverse, with 270 nominate species from at least 90 genera, 25 families and 11 orders. Although such diversity exists, knowledge of Indian millipedes still remain fragmented and scattered, as it has been reported that numerous millipede species have been inaccurately recorded in India and there is an absence of appropriate scientific information available on the identification and diversity of millipedes in forest ecosystems (Alagesan and Ramanathan, 2013; Chezhian and Prabakaran, 2016; Golovatch and Wesener, 2016). Alagesan and Ramanathan (2013) identified only five different species of millipedes in the study area of Algar Hills of the Eastern Ghats. Species richness of millipedes in the South Eastern Ghats of Tamil Nadu were investigated by Chezhian and Prabakaran (2016), however an in depth diversity profile was unable to be obtained due to lack of information regarding diversity and distribution in South Indian. Ten species of millipede were identified in the Yelagiri hills of the Eastern Ghats (Chezhian and Prabakaran, 2016).

2.2.3 The distribution of millipedes in Europe

Kime (1992) reported that there is a sound understanding of the abundance of various millipede species that inhabit Western Europe. Besides knowing their geographical region, their preferred biotopes and likely population densities in such habitats is also known. Since then, the knowledge of European millipede fauna has greatly increased. It was believed that only a small number of millipede species survived during the last ice age, in the area now known as Eastern Germany (Kime, 1992). The large number of species found in that area today are thought to have migrated there in the last 10 000 years (Hauser and Voigtländer, 2009). Eastern Germany is situated midway between the

South-West and South-East of the Alps. Species from either side are presumed to have migrated to the central point, with the same number of migrants coming from the east and west (Hauser and Voigtländer, 2009).

In the temperate regions of Europe, millipedes are commonly associated and found in either nemoral or deciduous forests. Oak forest habitats provide the best model for tracing intraspecies diversity and phylogeography (Wytwer *et al*., 2009). It is not uncommon for millipedes to inhabit European urban areas. Activities such as; gardening, cultivation of plants and the transportation of soil are just some of the ways in which species exchanges of millipedes occur between distant locations (Vilisics *et al*., 2012). Species diversity was observed to increase from the north to the south of Europe (Hauser and Voigtländer, 2009). A study reported that the millipede diversity in three Swiss cities (Zurich, Lucerne and Lugano) was found to be relatively low compared to described millipede fauna of Switzerland. In Europe, the average number of urban species range from 14 to 26 species (Vilisics *et al*., 2012).

2.2.4 The distribution of millipedes in Australia

Millipedes are amongst the poorly understood and studied groups in terrestrial Australia, despite their high levels of diversity and abundance in soil and leaf habitats (CSIRO, 2002). Most taxonomic studies available were based on small opportunistic collections and are not a true representation of the fauna which can be found in Australia (CSIRO, 2002). The Australian millipede fauna comprises of 2000 species which are in 20 families and 9 orders, however only 300 of these species have being described (Paoletti *et al*., 2007). Compared to the world fauna which consists of more than 12 000 species, it can be said that the millipede fauna in Australia is poorly represented. This is not unexpected as Australia is a relatively dry continent and millipedes are intolerant of low moisture environments (Paoletti *et al*., 2007).

Western Australia is regarded as a global biodiversity hotspot with great biological and biogeographical interest. The majority of the Diplopoda species has been collected predominantly within the high rainfall zones of South-Western Australia (Main *et al*., 2002). The forest region of south west Australia was reported to have rich invertebrate fauna, with the pill millipede genus (*Cynotelopus*) being located in a number of areas throughout such forests (Main *et al*., 2002). Although observed at numerous locations along river banks and under logs, it does retain a restricted range, only occurring along the 115 km stretch of the Southern coast of Western Australia (Main *et al*., 2002; Paoletti *et al*., 2007; Edward and Harvey, 2010). *Antichiropus* is a genus of short-range endemic millipedes occurring in south western and western South Australia (Wojcieszek *et al*., 2010). These millipedes are the centre of research based on behavioural ecology, biogeography, sexual selection and speciation and are vital for conservation planning in Western Australia (Wojcieszek *et al*., 2010).

Figure 2.4: Images of *Antichiropus variabilis,* the Western Australian millipede of the genus *Antichiropus* (Wojcieszek *et al*., 2010). **A:** Adult male in captivity. **B:** feeding on eucalypt leaves in captivity.

2.2.5 The distribution of millipedes in Africa

Over the years, numerous studies have been conducted in Africa regarding its invertebrate fauna, however, gaps in knowledge remain even after numerous efforts. A lack of taxonomic knowledge and a shortage of expertise in the region must be considered as contributing factors to the lack of comprehensive studies (Redman and Hamer, 2003). The orders Polydesmida, Siphononphorida, Polyzoniida, Sphaerotheriida, Stemmiulida, Spirobolida and Spirostreptida have been recorded in Africa (Redman and Hamer, 2003). The Spirostreptida is the most common and conspicuous (Mwabvu *et al*., 2015). The genera in the Spirostreptida that have been revised include *Archispirostreptus, Doratogonus, Cacuminostreptus, Spirostreptus, Plagiotaphrus* and *Bicoxidens* (Hamer, 2000; Mwabvu *et al*., 2009; Mwabvu *et al*., 2010; Mwabvu *et al*., 2015).

Of the nine African orders found south of the equator, eight of these orders occur in South Africa whilst two are restricted to the southern part of the country (Shelley and Golovatch, 2011). The millipede fauna of North Africa is better known than that of Sub-Saharan Africa, however, it still remains unsatisfactorily studied as information is still scattered in various taxonomic papers (Akkar *et al*., 2009). The richest and most diverse order of millipede in North Africa is the Julida, with 58 species belonging to 12 genera and 3 families, found in and around Algeria, Egypt, Morocco, Tunisia and Libya (Akkar *et al*., 2009). Hamer (1999) reported that Southern Africa has rich diplopod fauna with approximately 552 species from 71 genera. Spirostreptidans are the most familiar, noticeable and commonly encountered group of diplopods in urban areas of Southern Africa (Hamer, 1999).

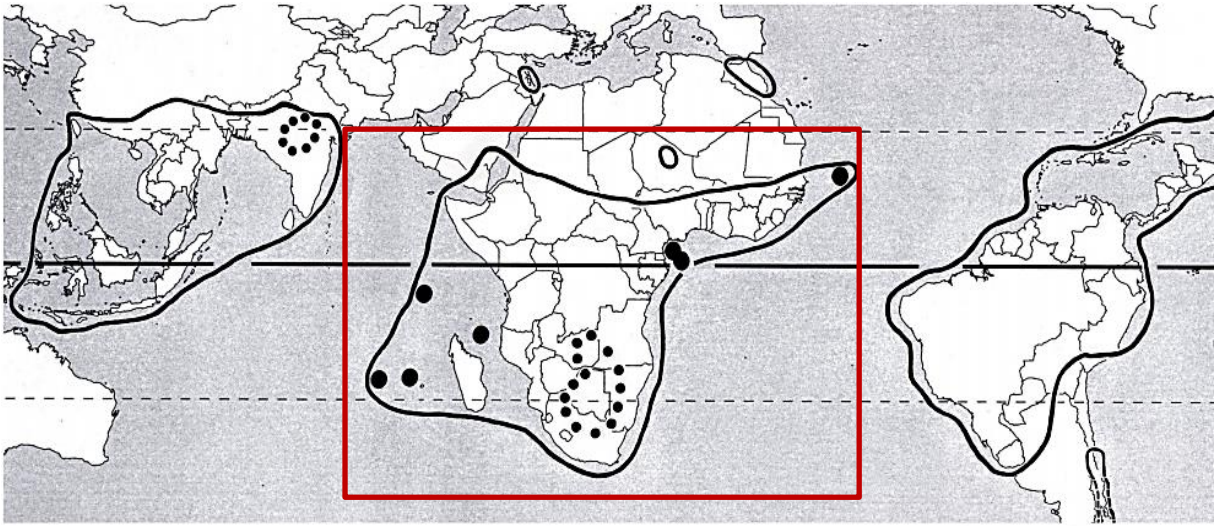

Figure 2.5: Distribution of the family Spirostreptidae. Solid lines represent known distribution and dash lines represent projected distribution (Shelley and Golovatch, 2011). The red rectangle highlights the distribution of Spirostreptidae in the Southern regions of Africa as earlier described by Hamer (1999).

The millipede fauna of South Africa is known to comprise of 458 species from seven orders and exhibits high level of endemism (Hamer *et al*., 2006). In South African biomes, approximately 17 % of species occur in the savanna and 52 % of species occur in forests (Hamer and Slotow, 2009). Although the grassland biome covers 28 % of South Africa, its millipede fauna has not been studied extensively (Hamer *et al*., 2006; Hamer and Slotow, 2009). A study of millipede fauna in the grassland areas of Drakensberg, reported a species richness of 5.7 %, considering only the described species in South Africa, this is indicative of the lack of species found within the area or rather, the lack of species one is able to identify due to insufficient knowledge (Hamer and Slotow, 2009). The low species richness can be attributed to the large proportion of undescribed species collected, emphasizing the poor state of knowledge of millipede fauna in South Africa (Hamer and Slotow, 2009). More emphasis must be placed on the identification and description of unknown millipede species in Africa, in order to gain better insight of the millipede diversity on the continent.

2.2.5.1 Genus *Bicoxidens* (Spirostreptidae)

Mwabvu *et al*. (2009) reported that the *Bicoxidens* is endemic to Southern Africa and inhabits areas such as the riverine vegetation, savanna woodlands and forests south of the Zambezi River. Nine species are known; *B. aridis, B. brincki, B. flavicollis, B. friendi, B. gokwensis, B. grandis, B. matopoensis, B. nigerrimus* and *B. nyathi* (Mwabvu *et al*., 2009, Tinago *et al*., 2017). Of the nine described species, *B. brincki* extends from Zimbabwe into the Kruger National Park, South Africa; and *B. flavicollis* which extends eastward from Zimbabwe to Maguge, Mozambique (Mwabvu *et al*., 2007). The genus occurs predominantly in central, eastern and southern regions of Zimbabwe (Mwabvu *et al*., 2009).

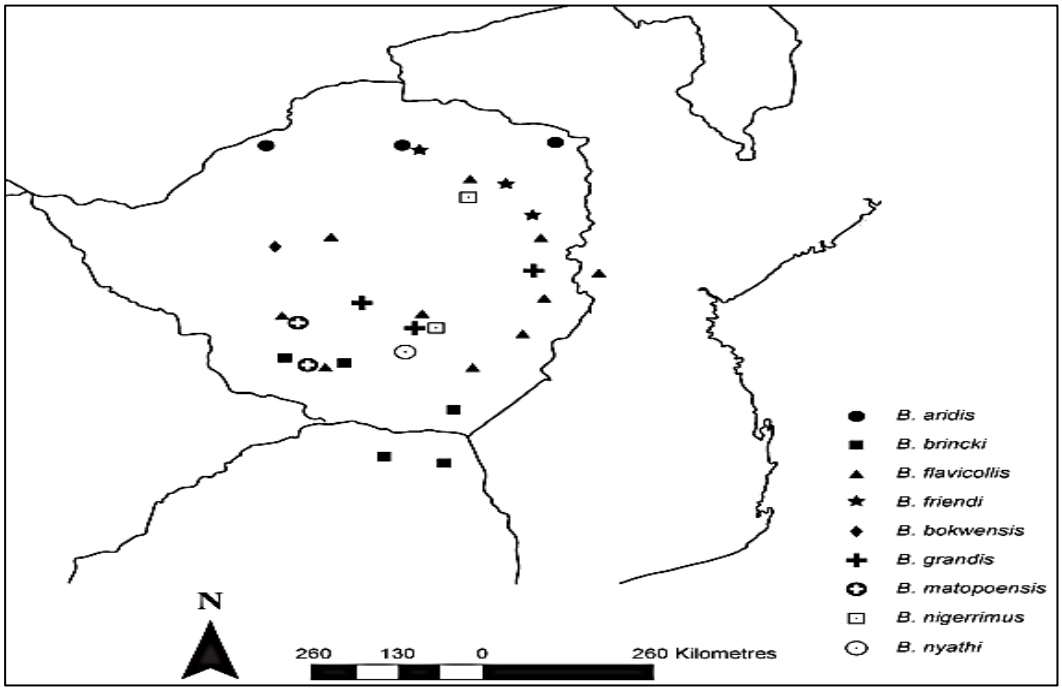

Figure 2.6: Distribution of *Bicoxidens* species in Southern Africa (Mwabvu *et al*., 2013).

The *Bicoxidens* is defined by a characteristic gonopod telopodite, which does not have a femoral process or for torsion of the femur (Mwabvu *et al*., 2007). However, the gonopods have one or two lobes situated at the femur or terminally (Mwabvu *et al*., 2007; Mwabvu *et al*., 2009). According to Mwabvu *et al*. (2007) species range from medium to large in size, and display intra-species variations, such as, colour and number of bodily rings based on their adaptations to different habitats. Individuals with an orange-yellow body colour have been observed in the highlands east of Zimbabwe and individuals north east of Harare are green-black in colour (Mwabvu *et al*. 2015). This demonstrates the vast physical differences expressed by the genus, although the most commonly observed colours are either brown or black (Mwabvu, 2000; Mwabvu *et al*., 2007; Mwabvu *et al*., 2009).

B. flavicollis and *B. nigerrimus* were the first two species to be described to the genus by Attems (1928) (Mwabvu *et al*., 2007). Seven more species have since been described based on gonopod morphology and external physical body characteristics (Mwabvu *et al*., 2007; Mwabvu *et al*., 2009). However, the species composition, taxonomic validity and distribution of some species within *Bicoxidens* still remain uncertain (Mwabvu *et al*., 2007). Species described from samples collected from a distance less than 300 km apart suggested that species may be synonymous, requiring further investigation (Mwabvu *et al*., 2007). Mwabvu (2000) cited gonopod descriptions and drawings of *B. polyptychus* as unclear. It was later determined that *B. polyptychus* was incorrectly described to the genus *Bicoxidens* based on morphological characteristics and was since moved to the genus *Brevitibius* (Mwabvu *et al*., 2007). The probability of discovering new species in remote locations is considerably high seeing that numerous areas in Zimbabwe and Southern Africa have yet to be

sampled. This leaves room for more studies to be conducted on *Bicoxidens* as current information is incomplete.

Table 2.1: Key to identifying species belonging to the genus *Bicoxidens* in Southern Africa.

Species	Description	Reference/s
B. flavicollis	Long and slender lobe of gonopod coxite, with a crossing lobe from the opposite side.	(Mwabvu, 2000)
B. nigerrimus	Short and round inwardly directed lobe found at the extreme end of gonopod coxite, no crossing over of lobes.	(Mwabvu, 2000)
B. aridis	Telopodite is knee produced and not L-shaped, clockwise coil after femoral lobes with no loops at the extremity.	(Mwabvu *et al*, 2009)
B. brincki	Telocoxite is "bird-head" like in appearance, with a pointed extension resembling a bird beak.	(Mwabvu, 2000)
B. grandis	Telocoxite has both and outwardly and inwardly directed lobe, in a downward direction yet overlapping the opposite lobe.	(Mwabvu, 2000)
B. friendi	T-shaped gonopod, with three median processes of the telocoxite. Short process is peg-like, median and longer process overlaps process from opposite gonopod. Incomplete loop of telopodite formed towards extremity.	(Mwabvu, 2000; Mwabvu *et al*., 2007)
B. nyathi	Apical telocoxite with oval shape median and lateral process. Incomplete loop formed distally by telopodite. Lateral leaf lobe extends from telocoxite.	(Mwabvu *et al*., 2007)
B. gokwensis	Apical telocoxite folded inward into an ear-shaped lobe. Inner margin of telocoxite has short median process and rounded proximal median lobe, separated by U-shaped groove. Only species within the genus to have yellow antennae.	(Mwabvu *et al*., 2007)
B. matopoensis	Apical telocoxite with rounded, broad lateral process and two median processes which form a V-shape.	(Mwabvu *et al*., 2007)

2.4 Habitats of Millipedes

Millipedes are found in temperate, tropical and subtropical regions, they inhabit leaf litter, the interface of leaf litter and soil, the surface/ uppermost soil and dead wood (Golovatch and Kime, 2009). Millipedes are also adapted to living in deeper soil layers, in the desert and in environments that may be considered extreme for arthropod groups (Kime and Golovatch, 2000; Golovatch and Kime, 2009).

2.4.1 Deserts

Millipedes found in dry environments are extremely scarce, but some are known to inhabit some of the harshest deserts, such as the Kalahari and Sonara (Crawford, 2012). *Orthoporus ornatus*, a species belonging to the family Spirostreptidae, is able to survive in arid conditions (Crawford, 2012). It is believed that this relatively large species has a slow rate of water loss due to numerous factors, namely; a waxy/impermeable cuticle, ability to water uptake from unsaturated air and the use of cracks, stones and burrows as shelter to buffer the temperature fluctuations experienced in the desert (Golovatch and Kime, 2009; Crawford, 2012). According to Crawford (2012) *Orthoporus ornatus* remains dormant deep in the soil for most of its life and the most active periods above ground, coincide with periods of rain, as individuals may alternately bask in the sun and then retreat, allowing their bodies to thermoregulate. Behavioural regulation of water balance takes place once individuals feed on the surface (Crawford, 2012).

In Africa, half of the land surface is covered by savanna grasslands therefore, making it the most extensive biome on the African continent (Scholes and Walker, 2004). Savannas cannot be defined as either forest or grassland, but rather a dynamic combination of both trees and grasses (Scholes and Walker, 2004, Beerling and Osborne, 2006). The savanna biome in Southern Africa is dominated largely by the Spirostreptida, unlike the Kalahari and Sonara, the diplopod fauna is far greater and richer in this area (Hamer *et al*., 2006). The Kalahari is notably one of the harshest deserts that millipedes are found to occupy (Mwabvu, 2017). Burrowing allows spirostreptid species to escape the extreme conditions in the Kalahari (Mwabvu, 2017).

Body size was found to play a pivotal and advantageous role in surviving desert conditions, as water loss along with heat loss and gain takes place at a slower rate in larger bodied diplopods (Golovatch and Kime, 2009). However, a definitive factor influencing the aforementioned is soil. The soil must always be loose enough to allow active burrowing (Minelli and Golovatch, 2001; Golovatch and Kime, 2009). Shear and Shelley (2007) reported that the most common deserticolous diplopods belong to the juloid morphotype and only a few small to tiny polydemoids thrive in deserts. Almost 17 % of the total described diplopod fauna can be found in Southern Africa savanna (Golovatch and Kime, 2009). Thus making Southern Africa deserts far richer in diplopod species than any other extreme environments (Hamer *et al*., 2006).

2.5 Gonopod morphology in taxonomic identification

Morphology based taxonomy is widely used to delimit taxa amongst groups which are difficult to identify, it is therefore a common method of classification amongst millipedes (Mwabvu *et al*., 2013). It is widely known that male genitalia evolve at a more rapid rate, making it efficient in taxonomic identification as they exhibit species specific characteristics (Song and Bucheli, 2010; Mwabvu *et al*., 2013). However, not all groups of millipedes have useful taxonomic information in the structure of

gonopods. Only orders belonging to Eugnatha (Chordeumatida, Callipodida, Polydesmida, Stemmiulida, Siphoniulida, Julida, Spirobolida and Spirostreptida), have a modified eighth leg pair which serves as a functional gonopod and can be used as a diagnostic tool at species level (Brewer *et al*., 2012; Minelli, 2015).

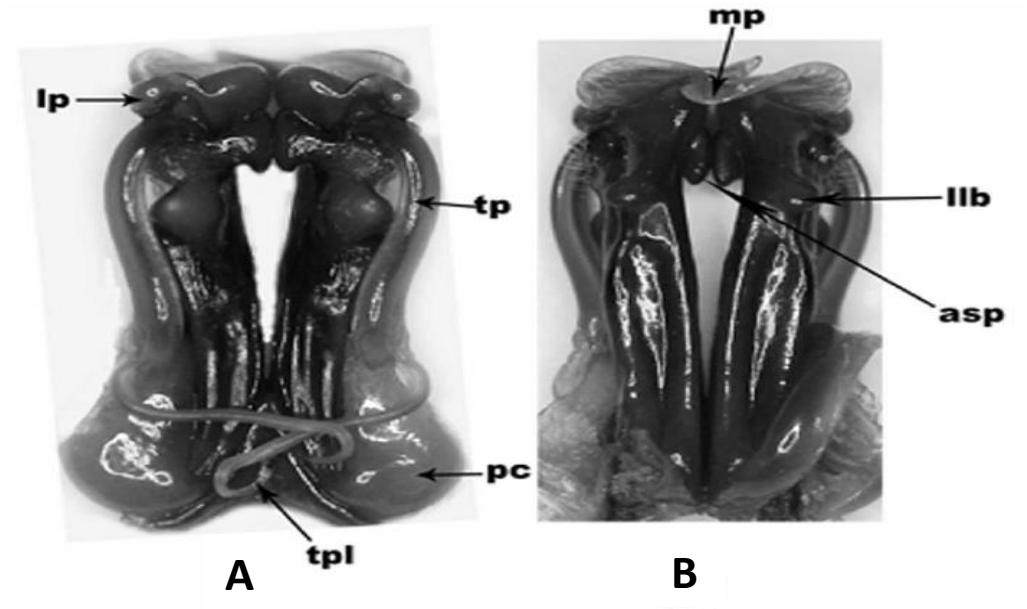

Figure 2.7: Gonopod imagery of *Bicoxidens flavicollis* (Mwabvu *et al*., 2007). **A:** Oral view (lp = lateral process, tp = telopodite, pc = paracoxite, tpl = telopodite loop). **B:** Aboral view (mp = median process, llb = lateral leaf lobe, asp = axe-shaped process).

Identification of spirostreptid is still largely dependent on gonopod morphology, however, morphology based classification is too inclusive and can potentially hide significant variation (Mwabvu *et al*., 2013; Mwabvu *et al*., 2015). Using the diverse structure of gonopods for identification has been hindered due to a lack of descriptive terminology which is deemed appropriate for significant comparisons to be made (Hoffman, 2008). Recent evidence has also come to light which suggests that genetic divergence might occur at a faster rate than genital divergence, indicating that speciation can occur without changes in gonopod morphology (Mwabvu *et al*., 2013; Tinago *et al*., 2017). As such, one cannot rely on gonopod morphology alone but rather verify diagnoses with molecular techniques.

2.6 Methods of studying phylogeny and phylogeography of millipedes

Several authors have used DNA methods that are used to study the taxonomy and delimiting millipedes (Pires and Marinoni, 2010; Mwabvu *et al*., 2013). Pires and Marinoni (2010) cited that due to DNA sequences of each species being unique, they can be regarded as genetic barcodes and are likely to fill gaps in knowledge that morphology was unable to address. Recent studies have used sequence data to test the taxonomic and phylogenetic hypotheses that are based on morphology

(Wiens and Penkrot, 2002). Mitochondrial DNA (mtDNA) is frequently used to analyse variation in wide-ranging polytypic animal species (Wiens and Penkrot, 2002).

2.6.1 DNA Barcoding

DNA barcoding is a relatively new method of identification which involves the extraction of DNA from a tissue sample of an organism (Kress and Erickson, 2012). DNA sequences have assisted in the accumulation of new information for our advancement and understanding of genetic and evolutionary relationships (Hajibabaei *et al*., 2007). This method aids in identification of species and enables diagnosis of all life stages of an organism (Kress and Erickson, 2012). It also assists in identifying species which are potentially new and have not yet been extensively studied. The process of DNA barcoding is relatively simple and consists of two steps; building a DNA barcode library of known species, and matching the barcode sequence of the unknown against the known library (Spelda *et al*., 2011; Kress and Erickson, 2012). The DNA barcoding process aims to use information from a single gene region to identify various species. Thus the information generated from barcoding and sequence analysis is apparent in all areas of biological science (Hajibabaei *et al*., 2007).

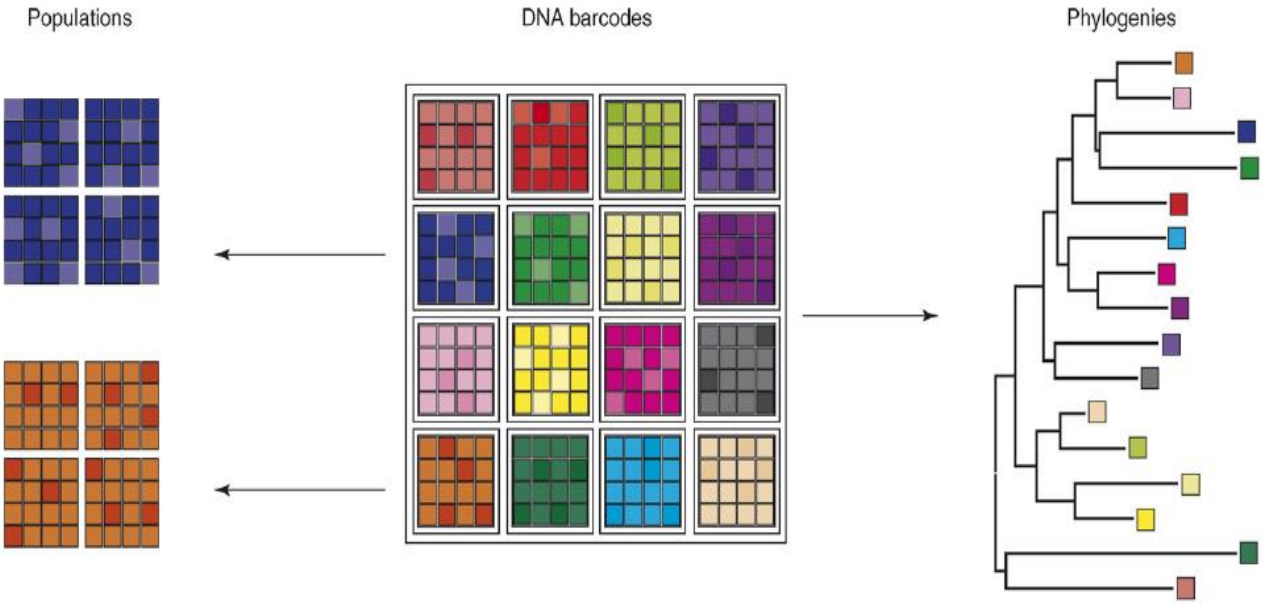

Figure 2.8: Diagrammatic representation of DNA barcoding in relation to phylogenetic and population genetic studies. Each colour represents a species and each square is representative of an individual. Within species variation is depicted by difference in colour shading (Hajibabaei *et al*., 2007).

For a gene sequence to be considered for use as a DNA barcode, it must satisfy the following; contain species level genetic variability and divergence, possess conserved flanking regions for the development of universal primers to be used in wide taxonomic application, and finally, the gene should be of short sequence length to facilitate extraction and amplification (Kress and Erickson,

2008; Spelda *et al*., 2011). A large number of markers are used to study population structure, however the number of species studied still remains low (Hajibabaei *et al*., 2007). Population genetic studies typically examine the variation occurring within a population of an individual species. The information gathered is used in geographical studies of populations, investigating factors such as genetic drift and migration (Hajibabaei *et al*., 2007).

Spelda *et al*. (2011) reported that the barcoding of myriapods is growing, especially in diplopods. With the identification of most millipedes, only the males specimens are able to be identified based on gonopod morphology. However, barcoding allows for the identification of all stages of life of both males and females (Spelda *et al*., 2011; Mwabvu *et al*., 2013). Wesener (2015) carried out a DNA barcode study of the pill millipede, *Glomeris malmivaga*, which is endemic to Germany. The study was based on the COI mitochondrial gene and found that *Glomeris malmivaga* and *Glomeris ornate* showed remarkable genetic and morphological similarities to each other, despite the large geographic distance between them (Wesener, 2015). The information which DNA barcoding provides insight into the genomic diversity patterns existing within a species, however it not sufficient to thoroughly address population level questions (Moritz and Cicero, 2004; Hajibabaei *et al*., 2007).

2.6.2 Mitochondrial DNA

The most widely used genetic marker in population genetic studies of animals is mtDNA cytochrome oxidase subunit I (COI). Mitochondrial DNA is often exploited due to its haploid nature, allowing it to be easily amplified among numerous taxa with no cloning necessary (Hurst and Jiggins, 2005). The mitochondrial genome structure and sequence provides researchers with evolutionary and comparative information as well as information based on gene flow, phylogenetics and molecular evolution (Mandal *et al*., 2014). Although mitochondrial DNA makes up a small fraction of the entire organisms genome, it continues to be the most popular marker used to study molecular genetic diversity (Galtier *et al*., 2009).

Using mtDNA in phylogenetic studies has many advantages. The mitochondrial region is strongly conserved across organisms, with limited duplications and short intergenic regions. Genetic material is strictly of maternal transmission with a high mutation rate, making it highly variable amongst natural populations (Galtier *et al*., 2009). The high mutation rate is due to its conserved simple structure and limited repair system. Such properties of mtDNA allow for the use of universal primers and recovery of genetic information from minute or degraded biological material (Galtier *et al*., 2009; Mandal *et al*., 2014). This allows for deductions based on the populations' history over a short period of time to be drawn, making it the first choice for population genetic studies (Wiens and Penkrot, 2002; Galtier *et al*., 2009; Mandal *et al*., 2014).

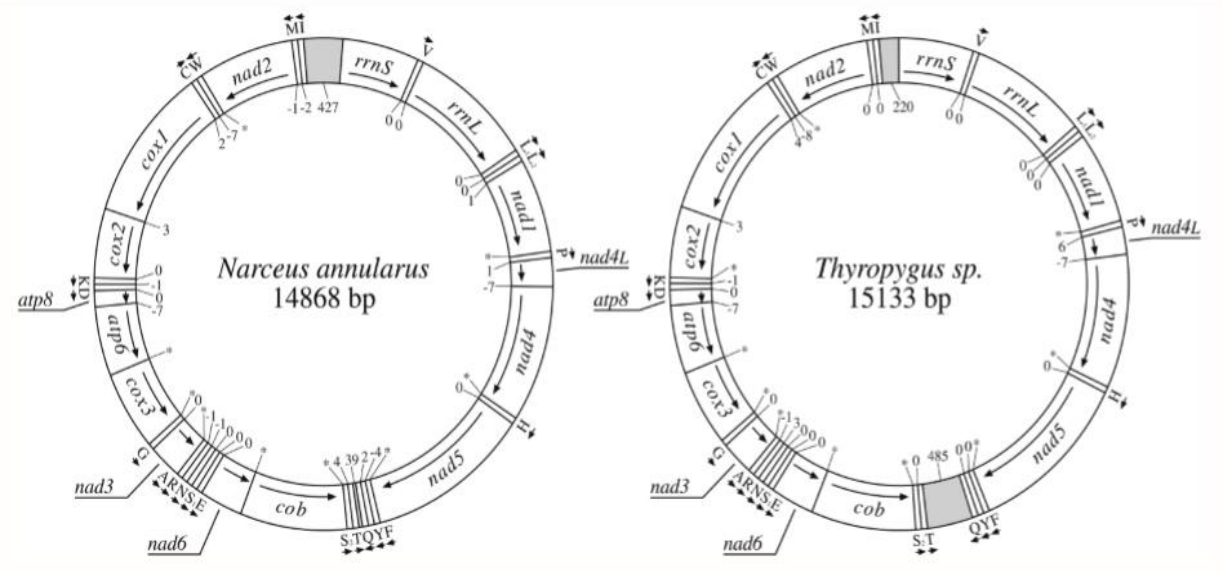

Figure 2.9: Gene maps of *N. annularus* and *Thyropygus* sp. generated from mitochondrial DNA (Lavrov *et al*., 2002).

Cryptic species require the use of molecular techniques to distinguish between morphologically identical species (Friedheim, 2016). DNA aids traditional taxonomy and morphological approaches, as each species has a unique DNA sequence (Pires and Marinoni, 2010). The use of DNA removes morphologically based disputes amongst researchers and allows for the generation of large amounts of data from small samples (Friedheim, 2016). DNA Barcoding has therefore, become an extremely useful tool with the potential to solve challenges of characterization and allows for the swift identification of species (Pires and Marinoni, 2010; Stepanović *et al*., 2016). Comparison of mitochondrial genetic information has become a powerful tool for reconstruction ambiguous phylogenies. It has gained popularity especially in arthropod studies (Mandal *et al*., 2014). Lavrov *et al*. (2002) sequenced the complete mitochondrial genome of two millipede species; *Narceus annularus* (Spirobolida) and *Thyropygus* sp. (Spirostreptida). Both these genomes displayed an identical gene arrangement pattern, which is usually not observed amongst metazoans. The genome is also divided in two gene clusters, with two control regions between them (Lavrov *et al*., 2002; Woo *et al*., 2007).

Cytochrome c oxidase (CO) is found in the mitochondrial region and is commonly used for phylogenetic studies (Tobe *et al*., 2010). The two subunits COI and COII sequences are applied to phylogenetic problems ranging from closely related species, families and orders (Tobe *et al*., 2010). The popularity of COI in such studies is mainly due to its slow evolving nature in comparison to other mitochondrial genes (Tobe *et al*., 2010; Patwardhan *et al*., 2014). It is therefore, useful in the recovery of trees. COI used in combination with 12S rRNA is suitable for distinguishing taxa at different taxonomic levels. Although COI is reported to provide substantial phylogenetic signal, cytochrome b

is reported as being the most useful marker in determining phylogenetic relationships (Tobe *et al*., 2010; Patwardhan *et al*., 2014).

2.6.2.1 Ribosomal RNA genes

Over the last six decades, ribosomal RNA encoding genes and related elements has been well studied, with researchers obtaining functional, structural and evolutionary information regarding these molecules (Noller, 2005; Gillespie *et al*., 2006). Their impact and use in phylogenetic studies has grown due its high gene copy per cell, which assists in gene amplification and sequencing (Gillespie *et al*., 2006). Evolutionary rates of base substitutions, variable regions and expansion segments can be contained to the conserved rRNA sequences within the rRNA gene (Gillespie *et al*., 2006). High levels of divergence can be found within closely related species, as the rates of substitution occur faster within transcribed spacer regions and non- coding spacer regions of the gene (Gillespie *et al*., 2006). It is for this reason that the conserved and variable regions of rRNA genes are a useful tool in uncovering phylogenetic relationships (Noller, 2005; Gillespie *et al*., 2006).

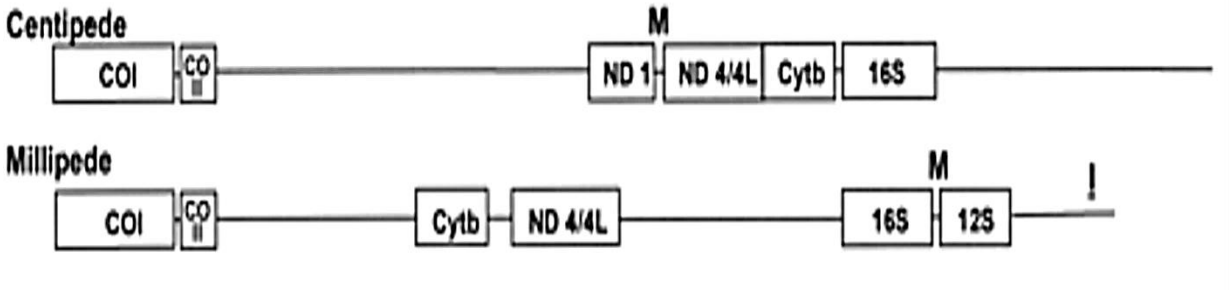

Figure 2.10: Map of mitochondrial gene locations in two Myriapod classes. **A:** Centipede (Chilopoda), **B:** Millipede (Diplopoda). The alphabets "M" and "I" represent the known positions of the tRNAs, Met and Ile respectively (Roehrdanz *et al*., 2002).

Carapelli *et al* (2000) cited the 12S rRNA gene as a small mitochondrial subunit composed of four domains, all of which hold single strand conserved regions. The ribosomal RNA form secondary structure motifs when molecules fold (Carapelli *et al*., 2000). The conservation of these motifs is advantageous for establishing homologies based on their position across sites and sequence alignment before phylogenetic analysis (Carapelli *et al*., 2000). The loops of the 12S rRNA secondary structure evolve at a faster rate than that of the stems (Wang *et al*., 2000). Stem regions compromise of strands with complementary base pairs and loop regions are not defined by complementary base changes (Pitz and Sierwald, 2010). Loop regions have transitional substitutions which accumulated faster than 10 million years after the divergence event had occurred, whereas transversional substitutions in both loops and stems remain unsaturated for more than 100 million years thus making 12S rRNA genes better suited and comprehensive for phylogenetic reconstruction studies (Wang *et al*., 2000).

However, the high mutation and variability rates in loop regions compared with stem regions can obscure the phylogenetic signals (Pitz and Sierwald, 2010).

The 16S rRNA gene is a ribosomal subunit, whose conserved secondary structure in combination with proteins forms the large mitochondrial RNA subunit (Schubart *et al*., 2000). Calo-Mata *et al* (2009) cited that the 16S rRNA region exhibits low evolutionary rates, making it useful in studies of interspecific differentiation among various species rather than studies of intraspecific differentiation among specimens of the same species. Conserved and variable regions present in the same gene are the likely reason 16S rRNA is favoured for phylogenetic research pertaining to separation events and phylogenetic reconstruction (Schubart *et al*., 2000). The genes, 12S and 16S rRNA is often used in conjunction with each other in various phylogenetic studies of different species (Kuznetsova *et al*., 2002). Due to the highly accelerated evolutionary rate of mtDNA, large amounts of variation between sequences of closely related species can be observed (Yang *et al*., 2014). The 12S and 16S rRNA sequences exhibit inter and intraspecfic nucleotide variations, however, both these genes share similar function and structure in organisms (Yang *et al*., 2014).

2.6.3 Nuclear DNA

The rate of evolution of nuclear DNA is significantly slow and has therefore, been considered a limitation with regards to intraspecific studies. Nuclear based markers have not being widely appreciated in phylogenetic reconstruction as haplotypes created with mtDNA. Such haplotypes are likely to become monophyletic four times quicker than nuclear markers (Wiens and Penkrot, 2002; Zhang and Hewitt, 2003). This suggests that species should show distinction in their mtDNA haplotypes long before showing distinction in nuclear based markers (Wiens and Penkrot, 2002; Siriwut *et al*., 2015). Polymorphisms exist widely in the nuclear DNA of eukaryotic organisms, which allows for unlimited opportunities for studying the mechanisms of evolution (Wiens and Penkrot, 2002). However, along with the opportunity come numerous challenges that must be overcome. Such challenges include; selection, recombination, gene-specific variation in rate and history and sequencing difficulties (Zhang and Hewitt, 2003). Most of these challenges do not occur with mtDNA and are specific to nuclear markers (Wiens and Penkrot, 2002; Zhang and Hewitt, 2003).

From the above mentioned hindrances, recombination occurs regularly in nuclear genomes and rates of recombination vary between loci (Posada, 2001). Evolutionary information generated by nuclear DNA haplotypes will be distorted by recombination events and produce false inferred history. A single tree will not be generated where recombination has occurred, as the sequences will be split into a group of trees (Posada, 2001; Zhang and Hewitt, 2003). There are two known strategies to deal with recombination events. The first would be to identify recombination events from the data set and to incorporate recombination in evolutionary models and second, to use nuclear regions with known low recombination rates (Posada, 2001; Posada, 2002). Nuclear ribosomal DNA which encodes for rRNA

is frequently used in phylogenetic approaches. The spacer regions of nuclear rDNA evolve much faster than the nuclear rRNA coding regions (Hwang and Kim, 1999). Spacer regions conduct neutral mutations with hardly any constraints, which results in fast evolutionary rates of the region. Due to the various rates of evolution among the different regions of nuclear rRNA, the nucleotide sequences can be used to infer broad spectrum phylogenetic relationships. Studies between closely related species or populations and basal lineages can also be investigated (Hwang and Kim, 1999; Patwardhan *et al*., 2014).

2.6.4 Microsatellites

It has been suggested that microsatellites have the potential to be used in studies of evolutionary relationships between groups which have evolved independently from each other for up to several million years (Richard and Thorpe, 2001). Microsatellites are widely distributed in the genome and are variable in length, they are short tandemly arrayed di, tri or tetranucleotide repeat sequences, ranging from one to five base pairs (Abdul-Muneer, 2014). This characteristic makes microsatellites a popular choice for molecular studies and to assess the genetic structure of populations (Richard and Thorpe, 2001; Abdul-Muneer, 2014). Data obtained from the evaluation of microsatellites is vital, given that most population genetic studies are based on data obtained from mitochondrial and nuclear markers (Abdul-Muneer *et al*., 2008).

Microsatellite sequences can be classified based on size, into minisatellites and microsatellites. The speed of processing, the potential to isolate a large number of loci and high variability levels associated with microsatellites, make them suitable in the detecting closely related populations (Abdul-Muneer *et al*., 2008). Microsatellites are abundantly distributed throughout eukaryotic genomes and their alleles are inherited in a mendelian fashion. Each locus is characterized by a known DNA sequences, consisting of both unique and repetitive DNA (Richard and Thorpe, 2001; Abdul-Muneer *et al*., 2008). Microsatellite loci have relatively high mutations rates which allows for very powerful kinship analyses to be conducted (Webster and Reichart, 2005). As large number of alleles will exist at a single locus, it is unlikely that unrelated individuals would share alleles (Webster and Reichart, 2005).

Most microsatellite primers are species specific which can be disadvantageous, as primers need to be isolated for each species of study. This can be a very time consuming and laborious task especially if the taxa in question have a considerably low microsatellite frequency (Zane *et al*., 2002; Abdul-Muneer *et al*., 2008). New techniques are being made available to facilitate the process of microsatellite isolation (Zane *et al*., 2002; Abdul-Muneer *et al*., 2008). Wojcieszek and Simmons (2009) isolated 25 novel microsatellite markers for the millipede, *Antichiropus variabilis,* to establish patterns in paternity. The markers used in this study represent the first microsatellite loci isolated from a millipede. Eleven of the 25 loci were found to be polymorphic, with eight loci successfully

amplifying in other species of *Antichiropus* (Wojcieszek and Simmons, 2009). Hasegawa *et al*. (2011) studied thirteen polymorphic microsatellite loci, newly isolated for *Brachycybe nodulosa.* Of the thirteen loci, only two showed heterozygosity lower than what was expected, the new loci will be useful in conservation efforts for this species as well as other species within the taxon (Hasegawa *et al*., 2011).

2.7 Statistical methods used in phylogenetic research

For many years the importance of natural selection in terms of molecular evolution has raised many questions, however it is generally accepted by researchers with regards to morphological, physiological and behavioural traits (Yang and Bielawski, 2000). Observed molecular variation such as polymorphism occurring within species and genetic divergence between species can be attributed to the random fixation of selectively neutral mutations (Yang and Bielawski, 2000). Over time we have witnessed the development of powerful statistical tools which detect molecular evolution based on environmental adaption, comparing substation rates in protein coding genes (Nei and Kumar, 2000; Yang and Bielawski, 2000). Statistical methods are crucial in determining evolutionary distance and divergence and have a fundamental role in phylogenetic studies (Nei and Kumar, 2000). The most commonly used methods will be highlighted below.

2.7.1 Bayesian inference of phylogeny

Bayesian inference is based on posterior probability of a tree and finds the phylogenetic tree which exhibits the maximum posterior probability, whilst evaluating the common features among sampled trees (Huelsenbeck, 2001). Posterior probability can be regarded as the probability of the tree being correct. Trees with high posterior probability values can therefore be considered as the best estimate of phylogeny (Huelsenbeck, 2001). Determining the posterior probability of a tree, requires the summation of all possible trees, the substitution model parameter values for each tree and the integration over all branch length combinations (Huelsenbeck *et al*., 2002). Markov chain Monte Carlo (MCMC) is the best numerical method used for the approximation of posterior probability (Huelsenbeck, 2001; Huelsenbeck *et al*., 2002). MCMC works in the following way; the Markov chain is started with a randomly chosen tree which describes the data well, the tree is then designated branch lengths. Secondly, a new tree is designated and proposed which must satisfy three conditions (Huelsenbeck, 2001; Huelsenbeck *et al*., 2002).

1. The proposed mechanism must be stochastic
2. Every tree must be accessible via repeated application of the proposed mechanism
3. The chain must be aperiodic (Huelsenbeck *et al*., 2002).

2.7.2 Maximum likelihood method of phylogeny

There is a need for phylogenetic reconstruction which can be performed at a faster speed without compromising accuracy. The Maximum likelihood (ML) method is found to be a good median to

meet both these requirements (Guindon and Gascuel, 2003). Maximum likelihood programs have the ability to recover the best or most correct tree from simulated data sets more often than other methods can, whilst comparing evolutionary models and different phylogenetic trees within a statistical framework (Whelan *et al*., 2001; Guindon and Gascuel, 2003). The statistical method behind ML can be described as, the likelihood of a hypothesis being equal to the probability of observing data if the hypothesis were to be correct. The statistical method chooses amongst hypotheses for the one which renders the data most acceptable (maximizes the likelihood) (Whelan *et al*., 2001). The attractive feature of the maximum likelihood approach is the knowledge that the phylogenetic estimates are statistically consistent and will always give the correct tree topology if given an adequate model and enough working data (Whelan *et al*., 2001; Guindon and Gascuel, 2003). Bielawski and Yang (2004) cited the ML method as useful in measuring divergent select pressures and identifying amino acid sites affected by selection. The only drawback of this method is the huge number of possible tree topologies generated when examining large numbers of sequences. It tree topology should be assessed individually which is a tedious task (Whelan *et al*., 2001).

2.7.3 Neighbour Joining method of phylogeny

The rise in phylogenetic studies requires many researchers to infer increasing large phylogenies. The Neighbour-joining (NJ) algorithm is a popular choice for scientists as it clusters taxa according to estimated pairwise distances (Shenemen *et al*., 2006). It offers a combination of computational efficiency along with reasonable accuracy (Simonsen *et al*., 2008). Phylogenies are reconstructed by iteratively joining pairs of nodes until a single node remains. As implied by the name, NJ begins with a star-shaped tree and iteratively joins neighbouring nodes until the construction of a bifurcating tree remains (Shenemen *et al*., 2006). Distance between a pair of nodes and the average distance to the rest of the nodes is the criterion used to assess the merger of nodes (Simonsen *et al*., 2008). NJ searches the distance matrix, identifying and joining the pair of nodes with the global minimum transformed distance, providing both topology and branch lengths of the final tree (Shenemen *et al*., 2006; Simonsen *et al*., 2008). Bryant (2005) cited the three criteria necessary for tree building based on distances, all of which the NJ method possesses:

1. Selection criterion should be linear
2. No a priori significance should be given to any taxa, meaning that any varying of the input orders will have no effect on the relative ordering of pairs.
3. The criterion should be consistent, the criterion will pick neighbouring taxa when given tree-like data.

A study by Jokobbson *et al* (2008) made use of the NJ method along with other statistical methods to infer genetic similarity in the human population worldwide by analysing single nucleotide polymorphism. The use of this algorithm in such a large scale study signifies its increasing reliability

as many selected parameters ensure the best tree topology. The NJ tree was constructed based on allele sharing distance with all internal branches supported by a bootstrap of 1000 replicates (Jokobbson *et al*., 2008).

2.8 Conclusion

Against this background this study aimed to determine genetic divergence present between populations of *B. flavicollis*, as gonopod morphology has been found to be too inclusive for a species which exhibits intraspecies diversity. The study aimed to evaluate the diversity present via phylogenetic and statistical analyses and determine the polymorphic or cryptic nature of the species.

2.9 References

Abdul Muneer, P., Gopalakrishnan, A., Musammilu, K., Mohindra, V., Lal, K., Basheer, V. and Lakra, W. (2008). Genetic variation and population structure of endemic yellow catfish, *Horabagrus brachysoma* (Bagridae) among three populations of Western Ghat region using RAPD and microsatellite markers. *Molecular Biology Reports*, 36(7): 1779–1791.

Abdul Muneer, P. (2014). Application of Microsatellite Markers in Conservation Genetics and Fisheries Management: Recent Advances in Population Structure Analysis and Conservation Strategies. *Genetics Research International*, 2014: 1–11.

Akkar, N., Stoev, P., Enghoff, H. and Nouira, S. (2009). The millipede order Julida (Myriapoda: Diplopoda) in Tunisia, with an overview of the North African species. *Soil Organisms*, 81(3): 453–488.

Alagesan, P. and Ramanathan, B. (2013). Diversity of Millipedes in Alagar Hills Reserve Forest in Tamil Nadu, India. *International Journal of Biodiversity*, 2013: 1–5.

Beerling, D. and Osborne, C. (2006). The origin of the savanna biome. *Global Change Biology*, 12(11): 2023–2031.

Bielawski, J. and Yang, Z. (2004). A Maximum Likelihood Method for Detecting Functional Divergence at Individual Codon Sites, with Application to Gene Family Evolution. *Journal of Molecular Evolution*, 59(1): 121–132.

Bond, J. and Sierwald, P. (2002). Cryptic speciation in the *Anadenobolus excisus* millipede species complex on the island of Jamaica. *Evolution*, 56(6): 1123–1135.

Bond, J., Beamer, D., Hedin, M. and Sierwald, P. (2003). Gradual evolution of male genitalia in a sibling species complex of millipedes (Diplopoda : Spirobolida : Rhinocricidae : Anadenobolus). *Invertebrate Systematics*, 17(6): 711–717.

Brewer, M., Sierwald, P. and Bond, J. (2012). Millipede Taxonomy after 250 Years: Classification and Taxonomic Practices in a Mega–Diverse yet understudied Arthropod Group. *PLoS ONE*, 7(5): 1–12.

Brewer, M., Swafford, L., Spruill, C. and Bond, J. (2013). Arthropod Phylogenetics in Light of Three Novel Millipede (Myriapoda: Diplopoda) Mitochondrial Genomes with Comments on the Appropriateness of Mitochondrial Genome Sequence Data for Inferring Deep Level Relationships. *PLoS ONE*, 8(7): 1–11.

Bryant, D. (2005). On the Uniqueness of the Selection Criterion in Neighbor–Joining. *Journal of Classification*, 22(1): 3–15.

Calo–Mata, P., Pascoal, A., Fernández–No, I., Böhme, K., Gallardo, J. and Barros–Velázquez, J. (2009). Evaluation of a novel 16S rRNA/tRNAVal mitochondrial marker for the identification and phylogenetic analysis of shrimp species belonging to the superfamily Penaeoidea. *Analytical Biochemistry*, 391(2): 127–134.

Carapelli, A., Frati, F., Nardi, F., Dallai, R. and Simon, C. (2000). Molecular phylogeny of the apterygotan insects based on nuclear and mitochondrial genes. *Pedobiologia*, 44(3–4): 361–373.

Chezhian, Y. and Prabakaran, S. (2016). Diversity of millipedes (Myriapoda: Diplopoda) In Yelagiri hills, Eastern Ghats, Vellore district, Tamil Nadu. *International Journal of Fauna and Biological Studies*, 3(2): 91–97.

Chipman, A. (2014). The First Myriapod Genome Sequence Reveals Conservative Arthropod Gene Content and Genome Organisation in the Centipede *Strigamia maritima*. *PLoS Biology*, 12(11), p.e1002005.

Crawford, C. (2012). *Biology of desert invertebrates*. Berlin: Springer: 1–314.

CSIRO (2002). *Invertebrate systematics: continuing invertebrate taxonomy*. 1st ed. Collingwood. Vic: CSIRO Publishing:1–17.

Culver, D. and Pipan, T. (2010). *The biology of caves and other subterranean habitats*. 1st ed. New York: Oxford University Press.

Culver, D., Master, L., Christman, M. and Hobbs, H. (2000). Obligate Cave Fauna of the 48 Contiguous United States. *Conservation Biology*, 14(2): 386–401.

Deharveng, L., Bréhier, F., Bedos, A., Tian, M., Li, Y., Feng, Z., Wengeng, Q. and Xuefeng, T. (2008). Mulun and surrounding karsts (Guangxi) host the richest cave fauna of China. *Subterranean Biology*, 6: 75–79.

Donaldson, M., Burnett, N., Braun, D., Suski, C., Hinch, S., Cooke, S. and Kerr, J. (2016). Taxonomic bias and international biodiversity conservation research. *Facets*, 1(1): 105–113.

Edgecombe, G. and Giribet, G. (2010). Myriapod phylogeny and the relationships of chilopoda. *Department of Organismic and Evolutionary Biology, Harvard University*: 143–168.

Edward, K. and Harvey, M. (2010). A review of the Australian millipede genus *Atelomastix* (Diplopoda: Spirostreptida: Iulomorphidae). *Zootaxa*, 2371: 1–63.

Friedheim, S. (2016). Comparison of Species Identification Methods: DNA Barcoding versus Morphological Taxonomy. *Manoa Horizons*, 1: 74–86.

Fusco, G. (2005). Trunk segment numbers and sequential segmentation in myriapods. *Evolution and Development*, 7(6): 608– 617.

Galanes, I. and Thomlinson, J. (2011). Soil millipede diversity in tropical forest patches and its relation to landscape structure in northeastern Puerto Rico. *Biodiversity and Conservation*, 20(13): 2967–2980.

Galtier, N., Nabholz, B., Glémin, S. and Hurst, G. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology*, 18(22): 4541–4550.

Gillespie, J., Johnston, J., Cannone, J. and Gutell, R. (2006). Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta:

Hymenoptera): structure, organization, and retrotransposable elements. *Insect Molecular Biology*, 15(5): 657–686.

Golovatch, S. (2015). Cave Diplopoda of southern China with reference to millipede diversity in Southeast Asia. *ZooKeys*, 510: 79–94.

Golovatch, S. and Kime, R. (2009). Millipede (Diplopoda) distributions: A review. *Soil organisms*, 81(3): 565–597.

Golovatch, S. and Wesener, T. (2016). A species checklist of the millipedes (Myriapoda, Diplopoda) of India. *Zootaxa*, 1(1): 1–75

Guindon, S. and Gascuel, O. (2003). A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology*, 52(5): 696–704.

Hajibabaei, M., Singer, G., Hebert, P. and Hickey, D. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, 23(4): 167–172.

Hamer, M. (1999). An illustrated key to the Spirostreptidan (Diplopoda: Spirostreptida) genera of Southern Africa. *Annals of the Natal Museum*, 40(1): 1–22.

Hamer, M. (2000). Review of the millipede genus *Doratogonus*, with description of fifteen new species from Southern Africa (Diplopoda, Spirostreptida, Spirostreptidae). *Annals of the Natal Museum*, 41: 1–76.

Hamer, M. and Slotow, L. (2009). A comparison and conservation assessment of the high–altitude grassland and forest–millipede (Diplopoda) fauna of the South African Drakensberg. *Soil organisms*, 81(3): 701–717.

Hamer, M., Slotow, R. and Lovell, S. (2006). The South African savanna millipede (Diplopoda) fauna: taxonomic diversity, endemism and spatial and temporal effects on conservation assessments. *Norwegian Journal of Entomology*, 53: 321–324.

Hasegawa, E., Yao, I., Futami, K., Yagi, N., Kobayashi, K. and Kudo, S. (2011). Isolation of microsatellite loci from the millipede, *Brachycybe nodulosa* Verhoeff. *Conservation Genetics Resources*, 4(1): 89–91.

Hauser, H. and Voigtländer, K. (2009). Zoogeography of the millipedes (Diplopoda) of eastern Germany. *Soil organisms*, 81(3): 617–633.

Hoffman, R. (2008). Two new genera of spirostreptid millipeds from central Africa, with a new terminology for male genitalia in the family Spirostreptidae (Diplopoda Spirostreptida). *Tropical Zoology*, 21: 167–186.

Huelsenbeck, J. (2001). Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. *Science*, 294(5550): 2310–2314.

Huelsenbeck, J., Larget, B., Miller, R.E. and Ronquist, F. (2002). Potential applications and pitfalls of Bayesian inference of Phylogeny. *Systems Biology*, 51(5): 673–688.

Hurst, G. and Jiggins, F. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B: Biological Sciences*, 272(1572): 1525–1534.

Hwang, U. and Kim, W. (1999). General properties and phylogenetic utilities of nuclear ribosomal DNA and mitochondrial DNA commonly used in molecular systematics. *The Korean Journal of Parasitology*, 37(4): 215–228

Iniesta, L. and Ferreira, R. (2013). Two new species of *Pseudonannolene* Silvestri, 1895 from Brazilian limestone caves (Spirostreptida: Pseudonannolenidae): synotopy of a troglophilic and a troglobiotic species. *Zootaxa*, 3702(4): 357–369.

Iniesta, L. and Ferreira, R. (2015). Pseudonannolene lundi n. sp., a new troglobitic millipede from a Brazilian limestone cave (Spirostreptida: Pseudonannolenidae). *Zootaxa*, (1): 123–128.

Jakobsson, M., Scholz, S., Scheet, P., Gibbs, J., VanLiere, J. and *et al*., (2008). Genotype, haplotype and copy–number variation in worldwide human populations. *Nature*, 451: 998 – 1003.

Kenny, N., Shen, X., Chan, T., Wong, N., Chan, T., Chu, K., Lam, H. and Hui, J. (2015). Genome of the Rusty Millipede, *Trigoniulus corallinus*, Illuminates Diplopod, Myriapod, and Arthropod Evolution. *Genome Biology and Evolution*, 7(5): 1280–1295.

Kime, R. (1992). On Abundance of West–European Millipedes. *Ber. nat. –med*, (10): 393–399.

Kime, R. and Golovatch, S. (2000). Trends in the ecological strategies and evolution of millipedes (Diplopoda). *Biological Journal of the Linnean Society*, 69(3):333–349.

Kress, W. and Erickson, D. (2012). DNA Barcodes: Methods and Protocols. *Methods of molecular biology*, 858: 3–8.

Kress, W. and Erickson, D. (2008). DNA barcodes: Genes, genomics, and bioinformatics. *PNAS*, 105(8): 2761–2762.

Kuznetsova, M., Kholodova, M. and Luschekina, A. (2002). Phylogenetic Analysis of Sequences of the 12S and 16S rRNA Mitochondrial Genes in the Family Bovidae: New Evidence. *Russian Journal of Genetics*, 38(8): 942–950.

Lavrov, D., Boore, J. and Brown, W. (2002). Complete mtDNA Sequences of Two Millipedes Suggest a New Model for Mitochondrial Gene Rearrangements: Duplication and Nonrandom Loss. *Molecular Biology and Evolution*, 19(2): 163–169.

Likhitrakarn, N., Golovatch, S. and Panha, S. (2011). Revision of the Southeast Asian millipede genus *Orthomorpha* Bollman, 1893, with the proposal of a new genus (Diplopoda, Polydesmida, Paradoxosomatidae). *ZooKeys*, 131: 1–161.

Loria, S., Lewis, J. and Zigler, K. (2011). Molecular phylogeography of the troglobiotic millipede *Tetracion* Hoffman, 1956 (Diplopoda, Callipodida, Abacionidae). *International Journal of Myriapodology*, 5: 35–48.

Main, B., Harvey, M. and Waldock, J. (2002). The distribution of the Western Australian pill millipede, *Cynotelopus notabilis* Jeekel (Sphaerotheriidae). *Records of the Western Australian Museum*, 20: 383–385.

Mandal, S., Chhakchhuak, L., Gurusubramanian, G. and Kumar, N. (2014). Mitochondrial markers for identification and phylogenetic studies in insects – A Review. *DNA Barcodes*, 2(1).

Miley, H. (1925). Development of the male gonopods and life history studies of a polydesmid millipede. *Ohio Journal of Science*, 27(1): 25–43.

Minelli, A. (1990). *Proceedings of the 7th International Congress of Myriapodology*. 1st ed. Leiden: E. J. Brill.

Minelli, A. (2015). *Treatise on Zoology – Anatomy, Taxonomy, Biology. The Myriapoda, Volume 2*. Leiden: Brill: 353–362.

Minelli, A. and Fusco, G. (2013). *Arthropod Biology and Evolution: Molecules, Development and Morphology*. Berlin, Heidelberg: Springer Berlin Heidelberg: 91–122.

Minelli, A. and Golovatch, S. (2001). Myriapods. *Encyclopedia of Biodiversity*, 4: 291–303.

Miyazawa, H., Ueda, C., Yahata, K. and Su, Z. (2014). Molecular phylogeny of Myriapoda provides insights into evolutionary patterns of the mode in post–embryonic development. *Scientific Reports*, 4: 1–9

Moritz, C. and Cicero, C. (2004). DNA Barcoding: Promise and Pitfalls. *PLoS Biology*, 2(10), p.e354.

Mwabvu, T. (2000). Two new species *Bicoxidens* Attems, 1928, from Zimbabwe (Diplopoda: Spirostreptida: Spirostreptidae). *African Zoology*, 35(1): 115–119.

Mwabvu, T. (2017). Descriptions of movement and burrow morphology of a tropical millipede, *Spirostreptus heros* Porat, 1872 (Diplopoda: Spirostreptida: Spirostreptidae) in the Kalahari desert, South Africa. *Journal of Entomology and Zoology Studies*, 5(3): 327–329.

Mwabvu, T., Hamer, M. and Slotow, R. (2009). A new species of *Bicoxidens* Attems, 1928 (Diplopoda, Spirostreptida, Spirostreptidae) from northern Zimbabwe. *ZooKeys*, 7: 75–81.

Mwabvu, T., Hamer, M. and Slotow, R. (2007). A taxonomic review of the southern African millipede genus, *Bicoxidens* Attems, 1928 (Diplopoda: Spirostreptida: Spirostreptidae), with the description of three new species and a tentative phylogeny. *Zootaxa*, 1452: 1–23.

Mwabvu, T., Hamer, M., Slotow, R. and Barraclough, D. (2010). A revision of the taxonomy and distribution of *Archispirostreptus* Silvestri 1895 (Diplopoda, Spirostreptida, Spirostreptidae), and description of a new spirostreptid genus with three new species. *Zootaxa*, (2567): 1–49.

Mwabvu, T., Lamb, J., Slotow, R., Hamer, M. and Barraclough, D. (2013). Is Millipede Taxonomy Based on Gonopod Morphology Too Inclusive? Observations on Genetic Variation and Cryptic Speciation in *Bicoxidens flavicollis* (Diplopoda: Spirostreptida: Spirostreptidae). *African Invertebrates*, 54(2): 349–356.

Mwabvu, T., Lamb, J., Slotow, R., Hamer, M. and Barraclough, D. (2015). Do cytochrome c oxidase 1 Gene Sequences Differentiate Species of Spirostreptid Millipedes (Diplopoda: Spirostreptida: Spirostreptidae)?. *African Invertebrates*, 56(3): 651–661.

Nefediev, P., Nefedieva, J. and Dyachkov, Y. (2014). A review of the anthropochore millipede fauna of Asian Russia, with new records from the Altai Province, Siberia (Diplopoda). *Arthropoda selecta*, 23(4): 337–345.

Nefedieva, J., Nefediev, P., Sakhnevich, M. and Dyachkov, Y. (2015). Distribution of millipedes along an altitudinal gradient in the south of Lake Teletskoye, Altai Mts, Russia (Diplopoda). *ZooKeys*, 510: 141–161.

Nei, M. and Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. New York: Oxford University Press: 17– 39.

Nelson, M. and Artamova, I. (2009). *Organic farming*. 1st ed. New York: Nova Science Publishers: 60–91.

Nguyen, A., Korsós, Z., Jang, K. and Hwang, U. (2017). A revision and phylogenetic analysis of the millipede genus *Oxidus* Cook, 1911 (Polydesmida, Paradoxosomatidae). *European Journal of Taxonomy*, (293): 1–22.

Nistelberger, H., Byrne, M., Coates, D. and Roberts, J. (2014). Strong Phylogeographic Structure in a Millipede Indicates Pleistocene Vicariance between Populations on Banded Iron Formations in Semi–Arid Australia. *PLoS ONE*, 9(3), p.e93038.

Noller, H. (2005). RNA Structure: Reading the Ribosome. *Science*, 309(5740): 1508–1514.

Paoletti, M., Osler, G., Kinnear, A., Black, D., Thomson, L., Tsitsilas, A., Sharley, D., Judd, S., Neville, P. and D'Inca, A. (2007). Detritivores as indicators of landscape stress and soil degradation. *Australian Journal of Experimental Agriculture*, 47(4): 412–423.

Patwardhan, A., Ray, S. and Ray, A. (2014). Molecular Markers in Phylogenetic Studies–A Review. *Journal of Phylogenetics & Evolutionary Biology*, 02(02).

Paulus, H. (2000). Phylogeny of the Myriapoda – Crustacea – Insecta: a new attempt using photoreceptor structure. *Journal of Zoological Systematics and Evolutionary Research*, 38(3): 189–208.

Pires, A. and Marinoni, L. (2010). DNA barcoding and traditional taxonomy unified through Integrative Taxonomy: a view that challenges the debate questioning both methodologies. *Biota Neotropica*, 10(2): 339–346.

Pitz, K. and Sierwald, P. (2010). Phylogeny of the millipede order Spirobolida (Arthropoda: Diplopoda: Helminthomorpha). *Cladistics*, 26(5): 497–525.

Posada, D. (2001). Unveiling the Molecular Clock in the Presence of Recombination. *Molecular Biology and Evolution*, 18(10): 1976–1978.

Posada, D. (2002). Evaluation of Methods for Detecting Recombination from DNA Sequences: Empirical Data. *Molecular Biology and Evolution*, 19(5): 708–717.

Reboleira, A. and Enghoff, H. (2014). Millipedes (Diplopoda) from Caves of Portugal. *Journal of Cave and Karst Studies*, 76(1): 20–25.

Redman, G. and Hamer, M. (2003). The distribution of southern African *Harpagophoridae* Attems, 1909 (Diplopoda: Spirostreptida). *African Invertebrates*, 44(1): 213–226.

Regier, J., Wilson, H. and Shultz, J. (2005). Phylogenetic analysis of Myriapoda using three nuclear protein–coding genes. *Molecular Phylogenetics and Evolution*, 34(1): 147–158.

Richard, M. and Thorpe, R. (2001). Can Microsatellites Be Used to Infer Phylogenies? Evidence from Population Affinities of the Western Canary Island Lizard (Gallotia galloti). *Molecular Phylogenetics and Evolution*, 20(3): 351–360.

Robertson, H., Lapraz, F., Rhodes, A. and Telford, M. (2015). The Complete Mitochondrial Genome of the Geophilomorph Centipede *Strigamia maritima*. *PLOS ONE*, 10(3), p.e0121369.

Roehrdanz, R., Degrugillier, M. and Black, W. (2002). Novel Rearrangements of Arthropod Mitochondrial DNA Detected with Long–PCR: Applications to Arthropod Phylogeny and Evolution. *Molecular Biology and Evolution*, 19(6): 841–849.

Scholes, R. and Walker, B. (2004). *An African savanna*. Cambridge [England]: Cambridge University Press: 2–16.

Schubart, C., Neigel, J. and Felder, D. (2000). The use of the mitochondrial 16S rRNA gene for phylogenetic and biogeographic studies of Crustacea. *Crustacean Issues*, 12: 817–830.

Shear, W. and Edgecombe, G. (2010). The geological record and phylogeny of the Myriapoda. *Arthropod Structure & Development*, 39(2–3): 174–190.

Shear, W. and Shelley, R. (2007). The milliped genus *Tidesmus* Chamberlin, 1943 (Polydesmida: Macrosternodesmidae). *Zootaxa*, 1656: 61–68.

Shearer, T., van Oppen, M., Romano, S. and Wörheide, G. (2008). Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Molecular Ecology*, 11(12): 2475–2487.

Shelley, R. (2009). The millipede families Spirostreptidae (Spirostreptida) and Paradoxosomatidae (Polydesmida) in the Middle East; first records of the Diplopoda from Saudi Arabia. *Insecta Mundi*, 83: 1–6.

Shelley, R. and Golovatch, S. (2011). Atlas of myriapod biogeography. I. Indigenous ordinal and supra–ordinal distributions in the Diplopoda: perspectives on taxon origins and ages, and a hypothesis on the origin and early evolution of the class. *Insecta Mundi*, 0158(1): 1–134.

Sheneman, L., Evans, J. and Foster, J. (2006). Clearcut: a fast implementation of relaxed neighbor joining. *Bioinformatics*, 22(22): 2823–2824.

Sierwald, P. and Bond, J. (2007). Current Status of the Myriapod Class Diplopoda (Millipedes): Taxonomic Diversity and Phylogeny. *Annual Review of Entomology*, 52(1): 401–420.

Simonsen, M., Muilund, T. and Pedersen, C. (2008). Rapid Neighbour–Joining. *Algorithms in Bioinformatics*, 5251: 113–122.

Siriwut, W., Edgecombe, G., Sutcharit, C. and Panha, S. (2015). The Centipede Genus *Scolopendra* in Mainland Southeast Asia: Molecular Phylogenetics, Geometric Morphometrics and External Morphology as Tools for Species Delimitation. *PLOS ONE*, 10(8): 1–56.

Song, H. and Bucheli, S. (2010). Comparison of phylogenetic signal between male genitalia and non–genital characters in insect systematics. *Cladistics*, 26(1): 23–35.

Spelda, J., Reip, H., Oliveira Biener, U. and Melzer, R. (2011). Barcoding Fauna Bavarica: Myriapoda – a contribution to DNA sequence–based identifications of centipedes and millipedes (Chilopoda, Diplopoda). *ZooKeys*, 156: 123–139.

Stepanović, S., Kosovac, A., Krstić, O., Jović, J. and Toševski, I. (2016). Morphology versus DNA barcoding: two sides of the same coin. A case study of *Ceutorhynchus erysimi* and *C. contractus* identification. *Insect Science*, 23(4): 638–648.

Tanabe, T., Katakura, H. and Mawatari, S. (2001). Morphological difference and reproductive isolation: morphometrics in the millipede *Parafontaria tonominea* and its allied forms. *Biological Journal of the Linnean Society*, 72(2): 249–264.

Tinago, T., Mwabvu, T. and MacDonald, A. (2017). Evidence of multiple divergent mitochondrial lineages within the southern African diplopod genus *Bicoxidens* Attems, 1928 (Spirostreptida). *African Zoology*, 52(4): 229–235.

Tobe, S., Kitchener, A. and Linacre, A. (2010). Reconstructing Mammalian Phylogenies: A Detailed Comparison of the Cytochrome b and Cytochrome Oxidase Subunit I Mitochondrial Genes. *PLoS ONE*, 5(11), p.e14156.

Tole, L. (2002). Habitat loss and anthropogenic disturbance in Jamaica's Hellshire Hills area. *Biodiversity and Conservation*, 11(4): 575–598.

Vilisics, F., Bogyó, D., Sattler, T. and Moretti, M. (2012). Occurrence and assemblage composition of millipedes (Myriapoda, Diplopoda) and terrestrial isopods (Crustacea, Isopoda, Oniscidea) in urban areas of Switzerland. *ZooKeys*, 176: 199–214.

Walker, M., Stockman, A., Marek, P. and Bond, J. (2009). Pleistocene glacial refugia across the Appalachian Mountains and coastal plain in the millipede genus Narceus: Evidence from population genetic, phylogeographic, and paleoclimatic data. *BMC Evolutionary Biology*, 9(25): 1–11.

Wang, H., Tsai, M., Tu, M. and Lee, S. (2000). Universal Primers for Amplification of the Complete Mitochondrial 12S rRNA Gene in Vertebrates. *Zoological Science*, 39(1): 61–66.

Webster, M. and Reichart, L. (2005). Use of Microsatellites for Parentage and Kinship Analyses in Animals. *Methods in Enzymology*, 395:222–235.

Wesener, T. (2015). No millipede endemics north of the Alps? DNA–Barcoding reveals *Glomeris malmivaga* Verhoeff, 1912 as a synonym of G. ornata Koch, 1847 (Diplopoda, Glomerida, Glomeridae). *Zootaxa*, 3999(4): 571–580.

Whelan, S., Liò, P. and Goldman, N. (2001). Molecular phylogenetics: state–of–the–art methods for looking into the past. *Trends in Genetics*, 17(5): 262–272.

Wiens, J. and Penkrot, T. (2002). Delimiting Species Using DNA and Morphological Variation and Discordant Species Limits in Spiny Lizards (Sceloporus). *Systematic Biology*, 51(1): 69–91.

Wojcieszek, J. and Simmons, L. (2013). Divergence in genital morphology may contribute to mechanical reproductive isolation in a millipede. *Ecology and Evolution*, 3(2): 334–343.

Wojcieszek, J. and Simmons, L. (2009). Isolation and characterization of 11 polymorphic microsatellite loci in the millipede *Antichiropus variabilis* Attems (Diplopoda: Polydesmida: Paradoxosomatidae). *Molecular Ecology Resources*, 9(4): 1208–1211.

Wojcieszek, J., Harvey, M. and Rix, M. (2010). Optimised captive husbandry conditions for the Western Australian 'Marri Millipede' *Antichiropus variabilis* (Diplopoda: Polydesmida: Paradoxosomatidae), with notes on natural history and tissue preservation techniques. *Records of the Western Australian Museum*, 26(1): 87–93

Woo, H., Lee, Y., Park, S., Lim, J., Jang, K., Choi, E., Choi, Y. and Hwang, U. (2007). Complete Mitochondrial Genome of a Troglobite Millipede *Antrokoreana gracilipes* (Diplopoda, Juliformia, Julida), and Juliformian Phylogeny. *Molecules and cells*, 23(2): 182–191.

Wytwer, J., Golovatch, S. and Penev, L. (2009). Variation in millipede (Diplopoda) assemblages in oak woodlands of the Eastern European Plain. *Soil organisms*, 81(3): 791–813.

Yang, L., Tan, Z., Wang, D., Xue, L., Guan, M., Huang, T. and Li, R. (2014). Species identification through mitochondrial rRNA genetic analysis. *Scientific Reports*, 4(1): 1–10.

Yang, Z. and Bielawski, J. (2000). Statistical methods for detecting molecular adaptation. *Trends in Ecology & Evolution*, 15(12) :496–503.

Zane, L., Bargelloni, L. and Patarnello, T. (2002). Strategies for microsatellite isolation: A review. *Mol. Ecol*, 11: 1–16.

Zhang, D. and Hewitt, G. (2003). Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, 12(3): 563–584.

CHAPTER 3

Genetic relatedness of *Bicoxidens flavicollis* from different localities in Zimbabwe based on 16S rRNA gene

Abstract

Southern Africa is host to the endemic millipede genus, *Bicoxidens*. Individuals belonging to this genus can be found inhabiting areas ranging from forests to woodlands. Previous classification of *Bicoxidens* based on the characterization of gonopod morphology was found to be too inclusive, hiding significant variation which exists between individuals belonging to a single species. Recent evidence suggests that genital divergence occurs at a slower rate than genetic divergence, thus making molecular techniques better suited to verify speciation. As such, the mitochondrial 16S rRNA gene was used to determine the genetic variation of *B. flavicollis* specimens from different localities in and around Zimbabwe. The genetic relationship between the *B. flavicollis* sequences inferred by the neighbour joining (NJ) heuristic search (1000 replicates) algorithm in PAUP generated a phylogenetic tree which suggested the presence of both admixed and genetically distinct populations. Twenty–two haplotype networks supported the tree findings, with each haplotype exhibiting high levels of genetic variation (H_d > 0.9). AMOVA analysis determined that variation among the populations was significantly greater (> 80 %) than the variation occurring within populations (< 12 %). A high fixation index (F_{ST} = 0.882) indicated a high level of population genetic differentiation. Principal coordinates analysis calculated using Euclidean similarity index, ordinated the individual sequences in accordance to observed haplotype networks with 11.9 % of the distance variation located to the right of the axis and 88.1% located to the left. Poisson correction model genetic divergence estimates were calculated based on haplotype data, p > 0.1 indicated high levels of divergence between haplotypes and p < 0.05 indicated low levels of divergence. Specimens sampled from the same locality did not exhibit identical genetic structure (p = 0.00). *Bicoxidens flavicollis* individuals demonstrated both distinctive phylogeographic diversity and genetic similarity for specific localities as confirmed by the genetic diversity and population clustering analyses. Observed phenotypic variation between individuals is presumed to be attributed to individuals acting under natural selection, suggesting *B. flavicollis* to be a genetically diverse species, with moderate variation between geographic localities. This study provides a significantly better understanding of the evolutionary relationship shared between geographically distant individuals belonging to a single species. This information can be further implemented in additional studies of the same species or utilized in comparative studies along with other species belonging to the genus.

Key words: *Bicoxidens flavicollis*; Divergence; Gonopod, Haplotype; Mitochondrial; Phylogeographic; Polymorphic; 16S rRNA

3.1 Introduction

Millipedes (Diplopoda) are extremely understudied arthropods (Brewer *et al*., 2012). Individuals belonging to this class occur in most environments, however, there a literature dearth to support this assertion (Cardosa *et al*., 2011). These terrestrial invertebrates undertake an important role in soil environments, they are among the most diverse group of decomposers partaking in the transformation of dead organic matter and movement of mineral soil constituents (Sierwald and Bond, 2007; Stašiov *et al*., 2012). The species are found in both tropical and temperate areas, the taxonomic productivity of millipedes have varied over the years with more than 12 000 species having been described to date (Sierwald and Bond, 2007; Golovatch and Kime, 2009; Brewer *et al*., 2012).

The characterization of species is of great importance as it is the driving force behind comparative and phylogenetic studies of diversity (Bond and Sierwald, 2002). Previously, before the rapid onset of molecular DNA technologies, morphometric data analysis was the most frequently used method to define species (Wiens and Penkrot, 2002). The genitalia of male millipedes is not considered adaptive, however, they can evolve as a result of environmental variation (Tanabe *et al*., 2001; Bond *et al*., 2003). The change in genitalia has a direct impact on mating signals and preferences and is considered to be a secondary consequence of ecological divergence (Miley, 1925; Tanabe *et al*., 2001; Bond *et al*., 2003). The minor details observed in gonopods are used by morphologists to reflect differences amongst species (Golovatch and Kime, 2009; Wojcieszek and Simmons, 2013). There is a need to re–evaluate morphological characterization against DNA based methods because morphology alone may not differentiate species which are genetically distinct (Mwabvu *et al*., 2013).

DNA based methods of delimiting of millipedes have become increasingly popular because each species has a unique DNA sequence which can potentially assist researchers in achieving more accurate information and conclusions, unlike morphological studies (Pires and Marinoni, 2010; Mwabvu *et al*., 2013). DNA barcoding assists in identification and the gathering of information which can be implemented in phylogenetic studies for the progression and understanding of evolutionary relationships and reconstruction (Hajibabaei *et al*., 2007; Kress and Erickson, 2012). During the study of polytypic species, mitochondrial DNA (mtDNA) is widely used to assess the level of genetic variation (Wiens and Penkrot, 2002). In comparison to the entire genome of an organism, mtDNA makes up a small fraction (Galtier *et al*., 2009). Due to the simple conserved structure of the mitochondrial region and limited repair system, mtDNA has a high mutation rate which makes it extremely variable in populations (Galtier *et al*., 2009). These attributes allow for the recovery of genetic information from degraded samples and to be used as universal primers (Galtier *et al*., 2009; Mandal *et al*., 2014). Cytochrome c oxidase (CO) and cytochrome b are often used in many phylogenetic studies as both these markers provide substantial phylogenetic signal (Tobe *et al*., 2010; Patwardhan *et al*., 2014). However, the conserved secondary structure of ribosomal RNA genes harbour both conserved and variable regions within the loops and stems of the structure, making such

genes suitable in unpacking phylogenetic relationships (Noller, 2005; Gillespie *et al*., 2006; Pitz and Sierwald, 2010). The 16S rRNA mitochondrial gene is routinely used in phylogenetic studies focusing on species differentiation because a great deal of variation between the sequences of closely associated species can be detected (Kuznetsova *et al*., 2002; Yang *et al*., 2014). Both the 12S rRNA and 16S rRNA sequences display inter and intraspecific nucleotide variations, yet, both these genes demonstrate similar structural functions in organisms (Yang *et al*., 2014). The 16S rRNA gene is commonly used in microbial phylogenetic studies because 16S rRNA phylogenies relate well to trends in the overall content of the gene (Caporaso *et al*., 2011). The low dispersal ability displayed by millipedes promotes geographic isolation, which over time may contribute to genetic divergence and speciation, making the 16S rRNA gene suitable for phylogeographic studies (Tanabe *et al*., 2001; Caporaso *et al*., 2011; Wojcieszek and Simmons, 2013).

The Spirostreptida is the most abundant and common millipede order in Africa, with approximately 71 genera recorded (Hamer, 1999; Mwabvu *et al*., 2015). Half of these genera occur south of the Zambezi and Kunene rivers in Southern Africa (Hamer, 1999). *Bicoxidens* is among the most studied genus in Africa. *Bicoxidens* is endemic to southern parts of the continent, inhabiting savanna woodlands, forests and riverine vegetation (Mwabvu *et al*., 2009; Mwabvu *et al*., 2010; Mwabvu *et al*., 2015). The genus occurs primarily in eastern, southern and central regions of Zimbabwe (Mwabvu *et al*., 2009).Nine species are currently described in this genus, with *Bicoxidens flavicollis* exhibiting phenotypic variation among populations (Tinago *et al*., 2017). Most specimens are either black or brown, some individuals from the east highlands of Zimbabwe and northeast Harare are orange–yellow and green–black, respectively (Mwabvu, 2000; Mwabvu *et al*., 2007; Mwabvu *et al*., 2009; Mwabvu *et al*. 2015). Individuals exhibit intra–species disparity, ranging from medium to large in size and vary in the number of body rings based probably on environmental adaptations (Mwabvu *et al*., 2007). Against this background, this study set out to determine the phylogenetic background of *B. flavicollis* based on its distribution across Zimbabwe using the 16S rRNA gene.

3.2 Materials and Methods

3.2.1 Sample collection

Bicoxidens flavicollis samples were collected from various localities in Zimbabwe (Table 3.1), resulting in 12 different population sample groups. Mid–body legs of collected specimens were removed and preserved in 100 % ethanol prior to DNA extraction. For this study, representative specimens from each locality were selected from a larger sample pool. Of the selected 42 samples, 37 successfully amplified the 16S rRNA gene with banding suitable for gel extraction. Five sequences from known localities (Chihota and Muterere) were obtained from Genbank with the following accession numbers; KF057753.1, KF057754.1, KF057755.1, KF057756.1 and KF057757.1 to make

the total sample size for this study 42, including the two outgroups, *Spirostreptus kruegeri* (*S. kruegeri*) and *Spirostreptus sebae* (*S. sebae*).

Table 3.1: Locality coordinates from where *B. flavicollis* specimens were collected in Zimbabwe.

Locality	Latitude	Longitude	Number of specimens
Charangwa	S 18 º 25.7174'	E 32 º 58.217'	4
Chihota	S 18 º 15.730'	E 31 º 05.274'	2
Chitombo	S 18 º 32.869'	E 32 º 47.168'	2
Hot Springs	S 19 º 39'	E 32 º 28'	2
Marange/ Bocha*	S 19 º 10'	E 32 º 18'	3
Mazowe	S 17 º 28.832'	E 30 º 59.206'	3
Muterere/ Dombomupunga*	S 18 º 25.732'	E 32 º 57.521'	5
Muzinga	S 18 º 25'	E 32 º 58'	3
Nyanga	S 18 º 10.662'	E 32 º 45.181'	4
Nyika	S 20.0792°	E 30.8384°	3
Sahumani	S 18 º 24.594'	E 32 º 58.208'	2
Shumba	S 20 º 0792'	E 30 º 8384'	2

* signifies samples collected from different spots within the same locality

3.2.2 DNA extraction

The mid–body legs were ground using mortar and pestle. Total genomic DNA was extracted using the Zymo Research Quick- DNA™ Miniprep kit (manufactured in USA). Ground mid–body legs were suspended in 500µl genomic lysis buffer, thereafter supernatant was transferred to a spin column where it was treated with 200µl DNA pre-wash buffer, 500µl g–DNA wash buffer and 50µl DNA elution buffer in accordance with the manufacturer's instructions. The extracted DNA was then quantified using the Thermo Scientific NanoDrop 2000 Spectrophotometer. High DNA concentrations were diluted and then stored for PCR usage at - 20°C.

3.2.3 Molecular amplication of the 16S rRNA gene using the Polymerase Chain Reaction

The 16S rRNA gene fragment was amplified by the universal primers 16Sar (5'CGCCTGTTTTTCAAAAACAT3') and 16Sbr (5'CCGGTTTGAACTCAGATCATGT3') (Nistelberger *et al*., 2014). All PCR amplifications were performed in a 25µl mixture which consisted of 12.5µl Green Taq PCR master mix, 1.5µl of each primer, 5.0µl DNA and 4.5µl dH_2O. The reaction mixture was subjected to 38 amplification cycles in the T100™ Thermal Cycler (BIO RAD) under the following conditions; initial denaturation at 94 °C for 3 minutes, denaturation at 94 °C for 30 seconds, annealing at 53 °C for 30 seconds, extension at 72 °C for 30 seconds and final extension at 72 °C for 2 minutes. Following amplification, the PCR products were analysed by gel electrophoresis on a 1.5 % gel run at 70 V for 60 minutes to view the product size of 500 bp and viewed on a ChemiDoc™ MP Imaging system (BIO RAD) to assess successful amplification. Gel extraction of the desired product band was performed at Inqaba Biotec (Pretoria, South Africa), thereafter the product was sequenced within the same facility. The sequencing procedure carried out at Inqaba, as cited by Mkize *et al* (2016); Sequencing is conducted using the ABI V3.1 Big dye kit following the manufacturer's instructions on an ABI 377 automated sequencer. The Zymo Seq clean–up kit is used to clean labelled

products, thereafter products are injected in the ABI 3500XL Genetic Analyzer (with a 50 cm array) using POP7.

3.2.4 Phylogenetic and Statistical analysis

The programme BioEdit was used to edit and align the sequences using ClustalW and mutations were checked against raw sequence data. The structuring of mitochondrial DNA diversity amongst the sample populations were tested using base–pair differences. The relationship between the sequences was then inferred by a phylogenetic tree, which was constructed using the (NJ) heuristic search (1000 replicates) algorithm in PAUP. Minimum spanning haplotype networks were constructed using NETWORK software and the appropriate polymorphism data tests based on the haplotype network were conducted using DnaSP v.6.12.01. The following tests were carried out; haplotype diversity, nucleotide diversity, Fu and Li's test and Tajima's test. Genetic and population structure was analysed using Arlequin 3.1. The Principal Coordinate's Analysis (PCoA) was conducted using the euclidean similarity index and a transformation component of c = 2. Genetic divergence estimates were calculated using the Poisson corrected model in MEGA7.

3.3 Results

3.3.1 Genetic diversity

The 16S rRNA mitochondrial DNA gene was successfully sequenced for 35 *B. flavicollis* individuals and two *Spirostreptus* individuals that were used as outgroups. Analysis of all 42 sequences yielded a total of 22 haplotypes. Seven haplotypes represented individuals from more than two populations. The DNA polymorphism data (Table 3.2) were generated based on the haplotype network. A high haplotype diversity (H_d = 0.959) and low nucleotide diversity (π_n = 0.13888) was observed, with an average number of nucleotide differences of k = 13.05459. The large haplotype diversity value and low nucleotide diversity value is indicative of a high measure of uniqueness between the haplotypes and the presence of genetic variation of the 16S rRNA gene, respectively.

The appropriate statistical tests (Tajima's test, Fu and Li's D, F statistic and Fu's F statistic) were conducted to gain insight regarding the frequency of mutations of the 16S rRNA gene. The statistical tests however, were not found to be significant. Tajima's D and Fu and Li statistics were used to assess if the sequence data showed deviation from neutrality. The output value for Tajima's D test was negative (-0.70968), signifying an excess of low frequency polymorphisms relative to the expectation. The Fu and Li's neutrality tests were not significant (p values > 0.10).

Table 3.2: Gene polymorphism data based on the 16S rRNA haplotype network of *B. flavicollis*.

Nucleotide Analysis	Parameter Estimate
Number of Haplotypes	22
Haplotype (gene) diversity	0.959
Variance of haplotype diversity	0.00016
Standard deviation of haplotype diversity	0.013
Nucleotide diversity per site (π)	0.13888
Sampling variance π	0.0012964
Standard deviation of π	0.03601
Average number of nucleotide differences, k	13.05459
Theta (per sequence) from Eta	16.26797
Theta (per site) from Eta	0.17306
Tajima's D	-0.70968 (p > 0.10)
Fu and Li's D* statistic	1.13497 (p > 0.10)
Fu and Li's F* statistic	0.57907 (p > 0.10)
Fu's Fs statistic	-1.297

* P values less than 0.10 were considered to be significant

3.3.2 Population structure

The analysis of molecular variance (AMOVA) (Table 3.3) was consistent with the haplotype polymorphism data, indicating genetic diversity amongst individuals. Individuals were grouped based on their sample localities and *Spirostreptus* individuals as the outgroup. The variation was perceived as follows; among populations within groups (88.23 %) and within populations (11.77 %). The genetic variation between *B.flavicollis* populations based on their locality is considerably high, indicating that each locality exhibits a great level of genetic diversity from each other. The variation observed between *B. flavicollis* individuals within the same population is relatively low, however, it was still significant. The fixation index value (F_{ST} = 0.88229), which is a measure of differentiation based on the population structure, approaches 1, supporting the calculated percentage variance and indicating high levels of heterozygosity within each population.

Table 3.3: AMOVA results for *B. flavicollis* 16S rRNA gene.

	d.f	Sum of squares	Variance components	Percentage of variance (%)	Fixation index
Among populations (within groups)	17	248.452	5.98592	88.23	0.88229(F_{ST})
Within populations	24	19.167	0.79861	11.77	

The Principal Coordinate's Analysis (PCoA) (Figure 3.1) supported the observed haplotype results. The GenBank obtained sequences ordinated to the far right of the axis making up 11.9 % of the distance variation whilst the remaining 88.1 % of the individuals ordinated to the left of the axis. Close clustering of individuals is observed in the top left axis, indicating low genetic distance between those individuals.

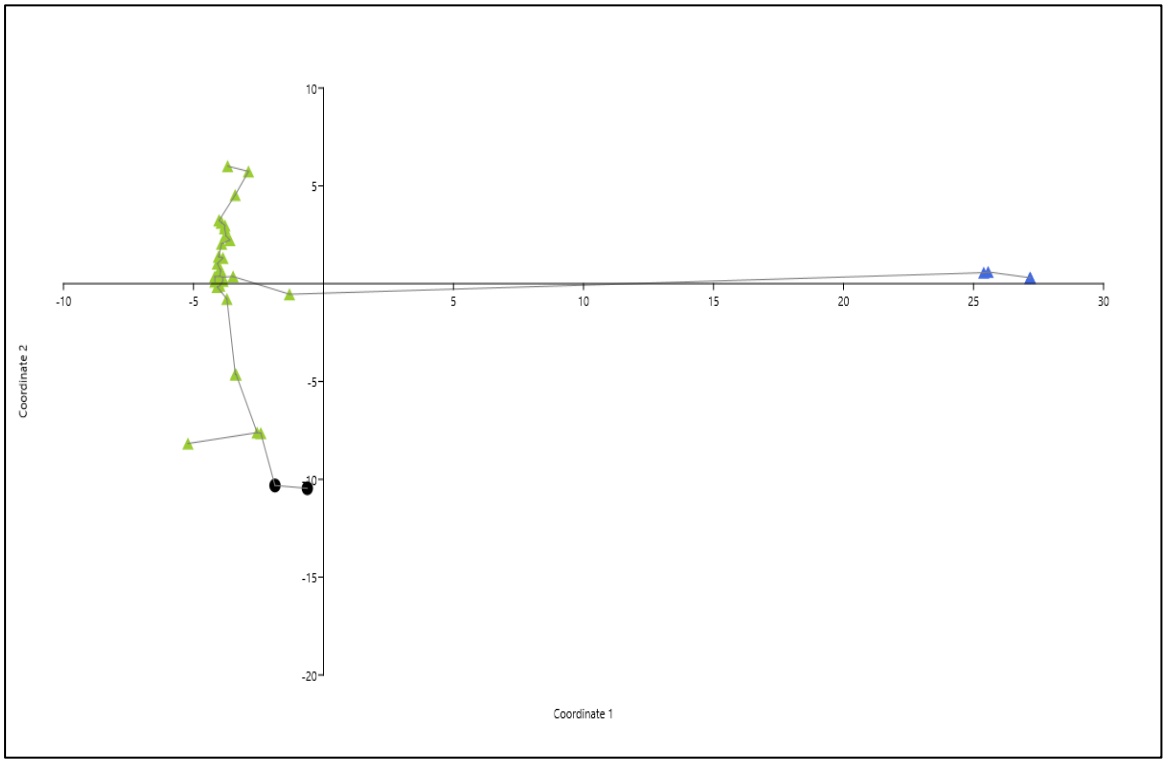

Figure 3.1: Principal coordinates analysis of *B. flavicollis* species based on 16S rRNA mitochondrial gene. Green triangles = *B. flavicollis* individuals obtained for study; Blue triangles = Genbank obtained sequences; Black dots = *Spirostreptus* sp. (outgroup).

Haplotype sharing was evident for 7 of the 22 haplotypes (Figure 3.2). Three of the shared haplotypes are supported by the Neighbour-joining phylogenetic tree generated by PAUP, showing the relationship between individuals sampled from Chitombo and Nyika (H17), Chitombo, Nyika and Nyanga (H16) and Muzinga and Dombomupunga (H4). The haplotypes H7 and H13 showed an admixture of the Charangwa and Marange populations, although not in close proximity on the phylogenetic tree. Both the aforementioned haplotypes shared similarity with individuals from Nyanga (H7) and Shumba (H13). The Chihota and Muterere sequences retrieved from GenBank appeared to have a strong genetic distinction from individuals sampled for this study, forming unique haplotypes and separate clades from individuals of the same locality. A total of 11 unique haplotypes were observed.

Estimates of evolutionary divergence between haplotypes (Table 3.4) were conducted using the Poisson correction model (Zuckerland and Pauling, 1965). All estimates greater than 0.1 are shaded in grey. A high estimate is indicative of greater genetic diversity and distinct population structure between said haplotypes. Haplotypes 20–22 had estimates ranging from 0.16 – 1.70, this is representative of the vast genetic difference of the Chihota and Muterere sequences obtained from GenBank in comparison to the sampled individuals for this study. None of the haplotypes are genetically identical to each other ($p \neq 0.00$) and it can be said that individuals or populations forming unique haplotypes exhibit genetic diversity from each other based on locality.

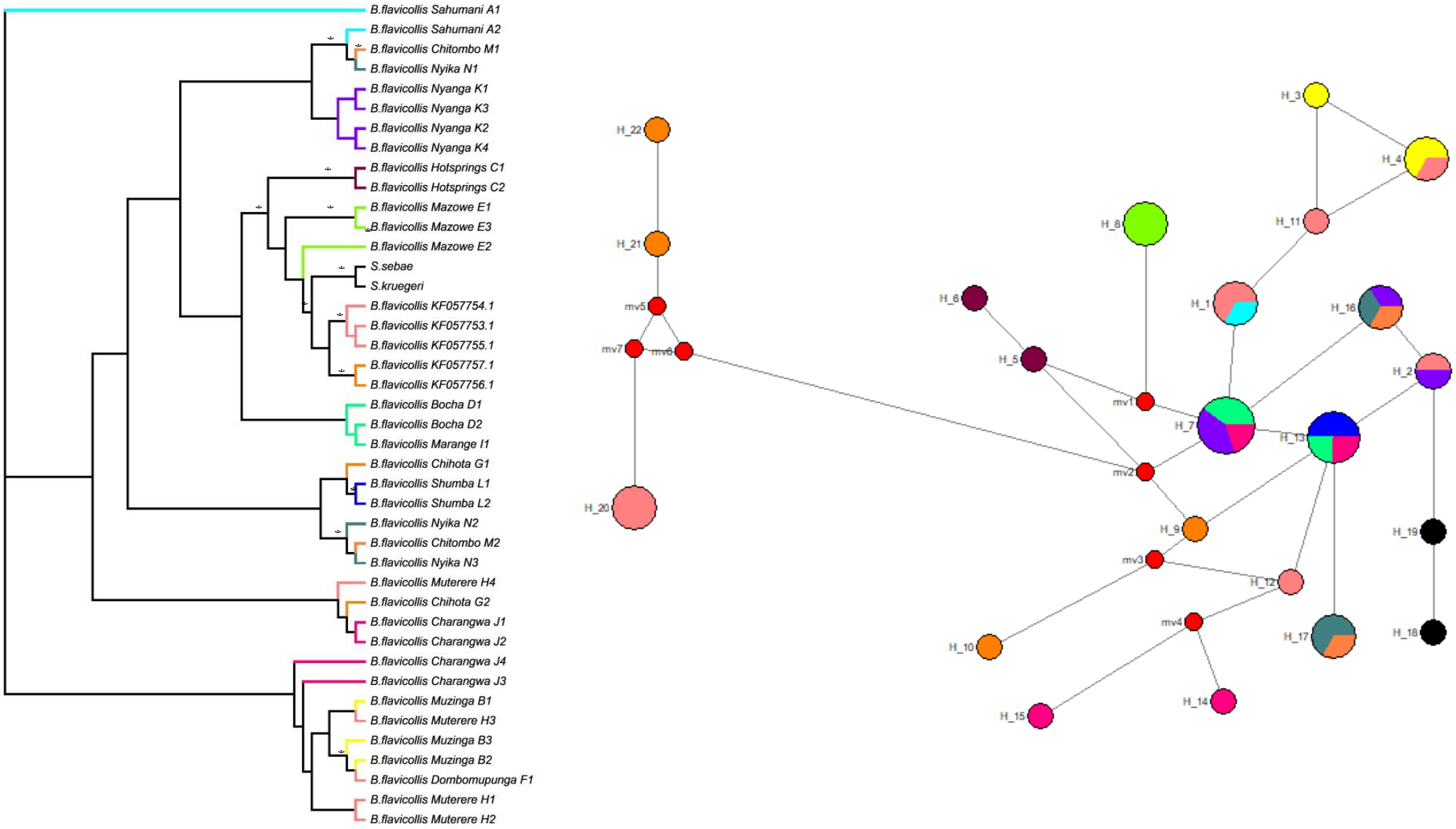

Figure 3.2: Heuristic search (1000 replicates) tree and haplotype network showing genetic diversity of *B. flavicollis* individuals using 16S rRNA gene. **A**: PAUP generated Neighbour joining tree, all clades with strong bootstrap values (> 70) are denoted by an asterisk *. **B**: Haplotype network, red circles indicate median vectors.

Table 3.4: Pairwise estimates of genetic distance based on 16S rRNA mitochondrial gene haplotype data of *B. flavicollis*. Genetic distances are shown below the diagonal and the standard error(s) are shown above the diagonal obtained by bootstrap procedure (1000 replicates).

	H_1	H_2	H_3	H_4	H_5	H_6	H_7	H_8	H_9	H_10	H_11	H_12	H_13	H_14	H_15	H_16	H_17	H_18	H_19	H_20	H_21	H_22
H_1		0.03	0.02	0.02	0.03	0.04	0.02	0.04	0.03	0.05	0.02	0.03	0.02	0.04	0.05	0.02	0.03	0.05	0.04	0.26	0.24	0.25
H_2	0.05		0.04	0.04	0.04	0.04	0.02	0.04	0.02	0.04	0.04	0.02	0.02	0.03	0.04	0.02	0.02	0.05	0.02	0.28	0.25	0.26
H_3	0.03	0.07		0.02	0.04	0.05	0.03	0.04	0.03	0.05	0.02	0.03	0.03	0.04	0.05	0.04	0.03	0.06	0.04	0.28	0.26	0.28
H_4	0.03	0.07	0.02		0.04	0.05	0.03	0.04	0.03	0.05	0.02	0.03	0.03	0.04	0.05	0.04	0.03	0.06	0.04	0.28	0.26	0.28
H_5	0.07	0.09	0.11	0.11		0.01	0.03	0.02	0.03	0.05	0.04	0.04	0.03	0.05	0.06	0.04	0.04	0.05	0.04	0.29	0.25	0.27
H_6	0.09	0.11	0.12	0.12	0.02		0.03	0.03	0.03	0.05	0.04	0.04	0.04	0.05	0.06	0.04	0.04	0.05	0.04	0.31	0.26	0.28
H_7	0.02	0.03	0.05	0.05	0.05	0.07		0.03	0.02	0.04	0.02	0.02	0.02	0.03	0.04	0.02	0.02	0.05	0.04	0.27	0.24	0.25
H_8	0.07	0.09	0.11	0.11	0.03	0.05	0.05		0.04	0.06	0.04	0.04	0.03	0.05	0.05	0.04	0.04	0.06	0.04	0.33	0.25	0.26
H_9	0.05	0.03	0.07	0.07	0.05	0.07	0.03	0.09		0.04	0.03	0.02	0.02	0.03	0.04	0.03	0.02	0.05	0.04	0.27	0.24	0.25
H_10	0.12	0.11	0.14	0.14	0.12	0.14	0.11	0.16	0.07		0.05	0.03	0.04	0.04	0.05	0.05	0.04	0.06	0.05	0.26	0.24	0.25
H_11	0.02	0.07	0.02	0.02	0.09	0.11	0.03	0.09	0.07	0.14		0.03	0.03	0.04	0.05	0.03	0.03	0.06	0.04	0.27	0.25	0.26
H_12	0.05	0.03	0.07	0.07	0.09	0.11	0.03	0.09	0.03	0.07	0.07		0.02	0.02	0.04	0.03	0.02	0.05	0.03	0.28	0.25	0.26
H_13	0.03	0.02	0.05	0.05	0.07	0.09	0.02	0.07	0.02	0.09	0.05	0.02		0.03	0.04	0.02	0.02	0.05	0.03	0.28	0.25	0.26
H_14	0.09	0.07	0.11	0.11	0.12	0.14	0.07	0.12	0.07	0.11	0.11	0.03	0.05		0.04	0.04	0.03	0.05	0.04	0.30	0.27	0.28
H_15	0.12	0.11	0.14	0.14	0.16	0.18	0.11	0.16	0.11	0.14	0.14	0.07	0.09	0.07		0.05	0.04	0.06	0.05	0.30	0.26	0.28
H_16	0.03	0.02	0.07	0.07	0.07	0.09	0.02	0.07	0.05	0.12	0.05	0.05	0.03	0.09	0.12		0.03	0.05	0.03	0.27	0.24	0.25
H_17	0.05	0.03	0.07	0.07	0.09	0.11	0.03	0.09	0.03	0.11	0.07	0.03	0.02	0.07	0.11	0.05		0.05	0.04	0.28	0.26	0.27
H_18	0.16	0.12	0.18	0.18	0.14	0.16	0.16	0.18	0.12	0.16	0.18	0.12	0.14	0.12	0.20	0.14	0.12		0.04	0.32	0.28	0.30
H_19	0.09	0.03	0.11	0.11	0.09	0.11	0.07	0.09	0.07	0.14	0.11	0.07	0.05	0.11	0.14	0.05	0.07	0.09		0.32	0.28	0.31
H_20	1.46	1.61	1.61	1.61	1.61	1.70	1.53	1.79	1.53	1.53	1.53	1.61	1.61	1.70	1.70	1.53	1.61	1.79	1.79		0.06	0.06
H_21	1.39	1.46	1.53	1.53	1.46	1.53	1.39	1.46	1.39	1.39	1.46	1.46	1.46	1.53	1.53	1.39	1.53	1.61	1.61	0.16		0.02
H_22	1.46	1.53	1.61	1.61	1.53	1.61	1.46	1.53	1.46	1.46	1.53	1.53	1.53	1.61	1.61	1.46	1.61	1.70	1.70	0.18	0.02	

3.4 Discussion

The mitochondrial 16S rRNA is a first–rate phylogenetic marker, which is used in numerous studies to establish ecological characteristics and biodiversity (Caporaso *et al*., 2011). In comparison to the nuclear genome, mitochondrial DNA evolves at a much faster rate and is therefore, used to examine relationships occurring at family, genus, species and population level (Hwang and Kim, 1999; Caporaso *et al*., 2011). The 16S rRNA expresses more variability (Hwang and Kim, 1999), making it a good marker for this particular study. Using the 16S rRNA gene, *B. flavicollis* expressed levels of intra–species variation based on population locality.

In the generated polymorphism data (Table 3.2), a strong haplotype diversity of H_d = 0.959 is observed. This highlights the presence of unique haplotypes and suggests that each haplotype, both unique and admixed are highly distinct from each other, with a low level of nucleotide differences observed. It is not uncommon for such distinction within a species to occur and can be attributed to geographic isolation based on distance. The sequences obtained from GenBank exhibited a higher level of genetic distance from the rest of the specimens sampled from the same localities as well as a high level of distance from each other based of PCoA results (Figure 3.1). Such can be representative of an evolved *B. flavicollis* population for said localities which has taken place to the time of the present study. A large clustering of sequences to the left of the axis is indicative of the low frequency polymorphisms present between individuals which resulted in the unique haplotypes.

Over time, restricted gene flow will result in genetic divergence due to environmental adaptations (Nistelberger *et al*., 2014). This is evident for the populations at Charangwa, Chihota, Muzinga, Mazowe, Hot springs and Dombomupunga, predominantly forming unshared haplotypes with single individuals (Figure 3.2B). Mwabvu *et al*. (2007) reported that the genetic variability of *B. flavicollis* is predictable as individuals are small–bodied, making dispersal difficult. This supports the AMOVA analysis (Table 3.3) which determined that variation among the populations was significantly greater (> 80 %) than the variation occurring within populations (< 12 %). However, seven haplotypes with admixed populations represent possible overlapping of populations. Hurst and Jiggins (2005) cited population and biogeographic studies using mtDNA to be compromised in certain instances by the existence of inherited symbionts, which happen to be common in invertebrates. A population infected by symbionts will exhibit mitochondrial polymorphism based on natural selection acting upon the microorganisms (Hurst and Jiggins, 2005). It is possible for the migration of an infected individual into an uninfected population to homogenize the haplotypes of the two populations based on mtDNA alone, whereas nuclear genes remain un-homogenized (Hurst and Jiggins, 2005).

Dombomupunga, Muzinga, Nyanga, Chitombo, Marange and Hot springs make up part of the eastern populations sampled in Zimbabwe. Both Dombomupunga and Muzinga form a single clade in the phylogenetic tree supporting their close proximity and shared a haplotype (H4) . Specimens collected

from Shumba and Nyika from the southern province of Masvingo appeared in the same clade, however, individuals from Chitombo shared a closer genetic identity to individuals from Nyika despite the geographic distance (H16 – 17). Environmental barriers contribute immensely to intra–species variation, as it further inhibits the movement of a species of low vagility (Mwabvu *et al*., 2007). A recent study conducted by Tinago *et al*. (2017) using the cytochrome c oxidase subunit 1 (CO1) to assess the phylogenetic relationship of *B. flavicollis* along with *B. brincki* and *B. friendi* determined that genetic divergence does not only occur amongst the widely distributed *B. flavicollis* but also amongst individuals within the same genus. It was demonstrated that genetic divergence was also present in individuals from the same locality.

Similar results were demonstrated in the present study using 16S rRNA. A large number of high ranging divergence estimates (> 0.1) (Table 3.4) of haplotypes were obtained, bearing in mind that certain individuals formed unique haplotypes, this is suggestive of the diversity present among individuals within populations. Interestingly, the outgroup *S. sebae* (H19) showed low genetic distance from haplotypes 2, 13 and 16 ($p = 0.03 – 0.05$). These haplotypes contain mixed populations, making it highly plausible that individuals within the above mentioned haplotypes share mixed ancestry with *S. sebae*. The distance estimates for *S. kruegeri* (H18) meet the expectation of what is probable from an outgroup, as all estimates range above 0.1 ($p > 0.1$).

The results of this study express both strongly distinct and structured *B. flavicollis* populations, as well as, high similarity of individuals from opposing localities. These results are not in line and expected of a phylogeographic study sampling from isolated populations. Nistelberger *et al*. (2014) had studied the semi–arid Australian millipede, *Atelomastix bamfordi*, from the Yilgarn Banded Iron Formations which are isolated islands, the findings showed strong genetic divergence based on mitochondrial markers. The strong divergence was attributed to genetic drift acting on each population over time. Cryptic species account for high 16S rRNA variation within a single species which was previously described by morphological methods alone (Bond and Sierwald, 2002). This is not the case in our study as more admixed haplotypes suggesting a level of genetic similarity which was observed across geographic locations.

3.5 Conclusion

B. flavicollis individuals exhibited both distinct phylogeographic diversity and genetic similarity for certain localities evident from shared haplotypes. One or more individuals from the following populations; Bocha, Charangwa, Chitombo, Marange, Muterere, Muzinga, Nyanga, Nyika, Sahumani and Shumba made up the seven shared haplotypes in this study. This was not the expected outcome as isolated populations in most circumstances host individuals with high genetic variation although from the same morphological species. Based on our results, *B. flavicollis* can be regarded to be a genetically diverse species, exhibiting phenotypic differences with low frequency polymorphisms

between individuals but moderate variation between geographic localities. The current study provided evidence of genetic divergence existing within a single species. This leaves room for further research using an array of available markers to be utilized on other diplopod species or arthropod groups which are geographically distant. Such information will provide researchers with a better evolutionary understanding and likely assist with conservation efforts where necessary.

3.6 References

Bond, J. and Sierwald, P. (2002). Cryptic speciation in the *Anadenobolus excisus* millipede species complex on the island of Jamaica. *Evolution*, 56(6): 1123–1135.

Bond, J., Beamer, D., Hedin, M. and Sierwald, P. (2003). Gradual evolution of male genitalia in a sibling species complex of millipedes (Diplopoda : Spirobolida : Rhinocricidae : Anadenobolus). *Invertebrate Systematics*, 17(6): 711–717.

Brewer, M., Sierwald, P. and Bond, J. (2012). Millipede Taxonomy after 250 Years: Classification and Taxonomic Practices in a Mega–Diverse yet understudied Arthropod Group. *PLoS ONE*, 7(5): p.e37240.

Caporaso, J., Lauber, C., Walters, W., Berg–Lyons, D., Lozupone, C., Turnbaugh, P., Fierer, N. and Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, 108(1): 4516–4522.

Galtier, N., Nabholz, B., Glémin, S. And Hurst, G. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology*, 18(22): 4541–4550.

Gillespie, J., Johnston, J., Cannone, J. and Gutell, R. (2006). Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements. *Insect Molecular Biology*, 15(5):657–686.

Golovatch, S. and Kime, R. (2009). Millipede (Diplopoda) distributions: A review. *Soil organisms*, 81(3): 565–597.

Hajibabaei, M., Singer, G., Hebert, P. and Hickey, D. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, 23(4): 167–172.

Hamer, M. (1999). An illustrated key to the spirostreptidan (Diplopoda: Spirostreptida) genera of Southern Africa. *Annals of the Natal Museum*, 40(1): 1–22.

Hurst, G. and Jiggins, F. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B: Biological Sciences*, 272(1572): 1525–1534.

Hwang, U. and Kim, W. (1999). General properties and phylogenetic utilities of nuclear ribosomal DNA and mitochondrial DNA commonly used in molecular systematics. *The Korean Journal of Parasitology*, 37(4): 215 – 228.

Kress, W. and Erickson, D. (2012). DNA Barcodes: Methods and Protocols. *Methods of molecular biology*, 858:3–8.

Kuznetsova, M., Kholodova, M. and Luschekina, A. (2002). Phylogenetic Analysis of Sequences of the 12S and 16S rRNA Mitochondrial Genes in the Family Bovidae: New Evidence. *Russian Journal of Genetics*, 38(8): 942–950.

Mandal, S., Chhakchhuak, L., Gurusubramanian, G. and Kumar, N. (2014). Mitochondrial markers for identification and phylogenetic studies in insects – A Review. *DNA Barcodes*, 2(1).

Miley, H. (1925). Development of the male gonopods and life history studies of a polydesmid millipede. *Ohio Journal of Science*, 27(1): 25–43.

Mkize, L., Mukaratirwa, S. and Zishiri, O. (2016). Population genetic structure of the freshwater snail, *Bulinus globosus*, (Gastropoda: Planorbidae) from selected habitats of KwaZulu–Natal, South Africa. *Acta Tropica*, 161: 91–99.

Mwabvu, T. (2000). Two new species *Bicoxidens* Attems, 1928, from Zimbabwe (Diplopoda: Spirostreptida: Spirostreptidae). *African Zoology*, 35(1): 115–119.

Mwabvu, T., Hamer, M. and Slotow, R. (2007). A taxonomic review of the southern African millipede genus, *Bicoxidens* Attems, 1928 (Diplopoda: Spirostreptida: Spirostreptidae), with the description of three new species and a tentative phylogeny. *Zootaxa*, (1452): 1–23.

Mwabvu, T., Hamer, M. and Slotow, R. (2009). A new species of *Bicoxidens* Attems, 1928 (Diplopoda, Spirostreptida, Spirostreptidae) from northern Zimbabwe. *ZooKeys*, 7: 75–81.

Mwabvu, T., Hamer, M., Slotow, R. and Barraclough, D. (2010). A revision of the taxonomy and distribution of *Archispirostreptus* Silvestri 1895 (Diplopoda, Spirostreptida, Spirostreptidae), and description of a new spirostreptid genus with three new species. *Zootaxa*, (2567): 1–49.

Mwabvu, T., Lamb, J., Slotow, R., Hamer, M. and Barraclough, D. (2013). Is Millipede Taxonomy Based on Gonopod Morphology Too Inclusive? Observations on Genetic Variation and Cryptic Speciation in *Bicoxidens flavicollis* (Diplopoda: Spirostreptida: Spirostreptidae). *African Invertebrates*, 54(2): 349–356.

Mwabvu, T., Lamb, J., Slotow, R., Hamer, M. and Barraclough, D. (2015). Docytochrome c oxidase 1 Gene Sequences Differentiate Species of Spirostreptid Millipedes (Diplopoda: Spirostreptida: Spirostreptidae)?. *African Invertebrates*, 56(3):651–661.

Nistelberger, H., Byrne, M., Coates, D. and Roberts, J. (2014). Strong Phylogeographic Structure in a Millipede Indicates Pleistocene Vicariance between Populations on Banded Iron Formations in Semi–Arid Australia. *PLoS ONE*, 9(3): p.e93038.

Noller, H. (2005). RNA Structure: Reading the Ribosome. *Science*, 309(5740): 1508–1514.

Patwardhan, A., Ray, S. and Ray, A. (2014). Molecular Markers in Phylogenetic Studies–A Review. *Journal of Phylogenetics & Evolutionary Biology*, 02(02): 1–9.

Pires, A. and Marinoni, L. (2010). DNA barcoding and traditional taxonomy unified through Integrative Taxonomy: a view that challenges the debate questioning both methodologies. *Biota Neotropica*, 10(2): 339–346.

Pitz, K. and Sierwald, P. (2010). Phylogeny of the millipede order Spirobolida (Arthropoda: Diplopoda: Helminthomorpha). *Cladistics*, 26(5): 497–525.

Sierwald, P. and Bond, J. (2007). Current Status of the Myriapod Class Diplopoda (Millipedes): Taxonomic Diversity and Phylogeny. *Annual Review of Entomology*, 52(1): 401–420.

Stašiov, S., Stašiová, A., Svitok, M., Michalková, E., Slobodník, B. and Lukáčik, I. (2012). Millipede (Diplopoda) communities in an arboretum: Influence of tree species and soil properties. *Biologia*, 67(5): 945–946.

Tanabe, T., Katakura, H. and Mawatari, S. (2001). Morphological difference and reproductive isolation: morphometrics in the millipede *Parafontaria tonominea* and its allied forms. *Biological Journal of the Linnean Society*, 72(2): 249–264.

Tinago, T., Mwabvu, T. and MacDonald, A. (2017). Evidence of multiple divergent mitochondrial lineages within the southern African diplopod genus *Bicoxidens* Attems, 1928 (Spirostreptida). *African Zoology*, 52(4):2 29–235.

Tobe, S., Kitchener, A. and Linacre, A. (2010). Reconstructing Mammalian Phylogenies: A Detailed Comparison of the Cytochrome b and Cytochrome Oxidase Subunit I Mitochondrial Genes. *PLoS ONE*, 5(11): p.e14156.

Wiens, J. and Penkrot, T. (2002). Delimiting Species Using DNA and Morphological Variation and Discordant Species Limits in Spiny Lizards (Sceloporus). *Systematic Biology*, 51(1): 69–91.

Wojcieszek, J. and Simmons, L. (2013). Divergence in genital morphology may contribute to mechanical reproductive isolation in a millipede. *Ecology and Evolution*, 3(2): 334–343.

Yang, L., Tan, Z., Wang, D., Xue, L., Guan, M., Huang, T. and Li, R. (2014). Species identification through mitochondrial rRNA genetic analysis. *Scientific Reports*, 4(1): 1–10.

Zuckerkandl, E. and Pauling, L. (1965). *Evolutionary divergence and convergence in proteins*. New York: Academic Press: 97–166.

CHAPTER 4

Phylogenetic structure of an endemic southern African millipede, *Bicoxidens flavicollis* based on the analysis of the mitochondrial 12S rRNA gene

Abstract

Bicoxidens flavicollis is one of the nine species belonging to the genus *Bicoxidens* which is endemic to Southern Africa. Individuals belonging to *B. flavicollis* can be found spread across Zimbabwe, even extending towards the western regions of Mozambique. Morphological methods of millipede characterization, has shown to have less diminishing power in comparison to molecular methods and hides significant amounts of variation existing between individuals of geographically distant populations. The mitochondrial 12S rRNA gene was used to determine the genetic divergence of populations of *B. flavicollis*. The relationship between the *B. flavicollis* sequences inferred by the (NJ) heuristic search (1000 replicates) algorithm in PAUP generated a phylogenetic tree which suggested the presence of both genetically distinct populations with two regions sharing close identity to outgroups. Nineteen haplotype networks supported the phylogenetic tree findings, with each haplotype exhibiting distinct genetic variation (H_d = 1.000). AMOVA analysis determined that variation among the populations was less (< 40 %) than the variation occurring within populations (> 60 %), although both are quite high. A low fixation index (F_{ST} = 0.37466) suggests a predominantly homozygous population genetic differentiation. Principal coordinates analysis calculated using Euclidean similarity index, ordinated the individual sequences in accordance to observed haplotype networks with 52.6 % of the distance variation located to the right of the axis and 47.4 % located to the left. Genetic divergence estimates based on individual sequences were calculated under the Poisson correction model. Divergence estimates greater than p > 0.1 indicated high levels of divergence between sequences and p < 0.05 indicated low levels of divergence. No identical sequences were observed (p = 0.00). *Bicoxidens flavicollis* individuals revealed distinctive phylogenetic diversity for each of the localities as confirmed by the genetic diversity and population structure analyses of the mitochondrial marker used. Further studies can be conducted utilizing the same mitochondrial marker within a larger sample size of *B. flavicollis* or amongst other genera. Comparative phylogenetic studies can be carried out using varying sample sizes and determining the effect of sample size on the observed population diversity.

Key words: *Bicoxidens flavicollis*; Genetic divergence; Haplotype network; Millipedes; Mitochondrial; Molecular; Phylogeographic

4.1 Introduction

Ecosystem engineering processes are largely dependent on soil invertebrates which are the key mediators in soil function (Lavelle *et al*., 2006). Activities such as burrowing, nesting and casting performed by these small–bodied taxonomic groups inadvertently contribute to aggregation and incorporation of leaf litter, as well as, structural porosity of soil (Lavelle *et al*., 2006; Golovatch and Kime, 2009). Millipedes, which are macro–invertebrate detrivores, take on a crucial role in the decomposition of organic matter (Reboleira and Enghoff, 2014; Tinago *et al*., 2017). With a high population density in forest soil (>1000 individuals per square metre) millipedes can consume approximately 10 % of the leaf litter present on the forest floor (Golovatch and Kime, 2009). These detrivores can inhabit caves, grasslands and deserts and change assemblage composition (Bogyó *et al*., 2015).

Over the years, the identification and description of millipede species has relied on the study of male gonopod morphology. Gonopods are responsible for the transfer of sperm (Bond and Sierwald, 2002; Tinago *et al*., 2017). Wojcieszek and Simmons (2013) reported that species rich taxa display substantial divergence in reproduction related traits that can be used to delimit species. Although variation amongst genitalia has a significant role as a species–specific diagnostic tool, very little is actually known about the mechanism of variation (Bond *et al*., 2003). Despite the frequent usage of morphological methods in delimiting species, gonopod structure may not be of taxonomic value in all groups. However, in the Orders Eugnatha; Callipodida, Chordeumatida, Julida, Polydesmida, Stemmiulida, Siphoniulida, Spirobolida and Spirostrepida gonopods can be used as a diagnostic tool at species level because the legs on the eigth body ring are modified for sperm transfer (Brewer *et al*., 2012; Minelli, 2015).

New evidence suggests that speciation may occur without a significant change in gonopod morphology because genetic divergence may occur at a faster rate than divergence of male genitalia (Bond *et al*., 2003; Wojcieszek and Simmons, 2013; Tinago *et al*., 2017). Therefore, genetic divergence along with an increase in evolutionary diversity cannot be completely determined based on morphological characterization of the gonopod alone (Bond and Sierwald, 2002; Mwabvu *et al*., 2013; Tinago *et al*., 2017). Morphological methods of characterization must therefore be reassessed against DNA based methodology, as genetic distinction and divergence cannot be determined based on morphology alone (Mwabvu *et al*., 2013).

Mitochondrial DNA is often preferred for phylogenetic studies due to its haploid nature which allows it to be easily amplified in numerous taxa without the tedious task of cloning (Hurst and Jiggins, 2005). The structure and sequence of the mitochondrial genome provides researchers with sufficient evolutionary and comparative information to draw conclusions on gene flow, phylogenetics and molecular evolution (Mandal *et al*., 2014). The use of mtDNA in phylogenetic studies has numerous

advantages. The most imperative being that the mitochondrial region is strongly conserved across organisms, with short intergenic regions and limited duplications. Genetic material is strictly maternally inherited with a high mutation rate, making it highly variable amongst natural populations and perfect for phylogenetic studies (Galtier *et al*., 2009).

The genus *Bicoxidens* is among the most studied genera in Africa, predominantly inhabiting southern parts of the continent (Mwabvu *et al*., 2009; Mwabvu *et al*., 2010; Mwabvu *et al*., 2015). The genus occurs primarily in Zimbabwe, South Africa and Mozambique with nine described species to date (Mwabvu *et al*., 2009; Tinago *et al*., 2017). *Bicoxidens flavicollis* has been found to exhibit phenotypic variation, expressing different colours ranging from orange–yellow and green–black in the easter highlands of Zimbabwe and north eastern Harare, respectively (Mwabvu, 2000; Mwabvu *et al*., 2007; Mwabvu *et al*., 2009; Mwabvu *et al*. 2015). Intra–species disparities is common as some individuals range from medium to large in size and vary in the number of body rings (Mwabvu *et al*., 2007). Against this background, this study set out to use the mtDNA 12S rRNA gene to establish the phylogenetic nature of *B. flavicollis* based on its wide spread distribution across Zimbabwe.

4.2 Materials and Methods

4.2.1 Sample collection

Bicoxidens flavicollis samples were collected from various localities in Zimbabwe (Table 4.1), resulting in 13 different population sample groups. About 10 mid–body legs of collected specimens were removed and preserved in 100 % ethanol. Representative specimens from each locality were selected from a larger sample pool for PCR, however only one sample from each locality was successfully amplified by PCR. A total of 19 samples were used in this study, which included four *Doratogonus* sp. specimens as the outgroup.

Table 4.1: Coordinates of localities from where *B. flavicollis* was collected in Zimbabwe.

Locality	Latitude	Longitude	Number of specimens
Charangwa	S 18 º 25.7174'	E 32 º 58.217'	1
Chihota	S 18 º 15.730'	E 31 º 05.274'	1
Chitombo	S 18 º 32.869'	E 32 º 47.168'	1
Chegutu	S 18 º 08.1380'	E 30 º 08. 5887'	1
Hot Springs	S 19 º 39'	E 32 º 28'	1
Marange/ Bocha*	S 19 º 10'	E 32 º 18'	2
Mazowe	S 17 º 28.832'	E 30 º 59.206'	1
Muterere/ Dombomupunga*	S 18 º 25.732'	E 32 º 57.521'	2
Muzinga	S 18 º 25'	E 32 º 58'	1
Nyanga	S 18 º 10.662'	E 32 º 45.181'	1
Nyika	S 20.0792°	E 30.8384°	1
Sahumani	S 18 º 24.594'	E 32 º 58.208'	1
Shumba	S 20 º 0792'	E 30 º 8384'	1

* signifies samples collected from different spots within the same locality

4.2.2 DNA extraction

The mid–body legs were ground using a mortar and pestle. Total genomic DNA was extracted using the Zymo Research Quick-DNA™ Miniprep kit (manufactured in USA). Following the manufacturer's instructions, the ground mid–body legs were firstly suspended in 500µl genomic lysis buffer, thereafter the supernatant was transferred to a spin column where it was treated with 200µl DNA pre–wash buffer, 500µl g–DNA wash buffer and 50µl DNA elution buffer. The extracted DNA was then quantified using the Thermo Scientific NanoDrop 2000 Spectrophotometer. High DNA concentrations were diluted and then stored for PCR usage at - 20°C.

4.2.3 Molecular amplification of the 12S rRNA gene using Polymerase Chain Reaction

The 12S rRNA gene fragment was amplified by the universal primers 12S F1069 (5'ACTGGGATTAGATACCCCACTATG3') and 12S R1219 (5'ATCGATTATAGAACAGGCTCC TC3') (Foran *et al*., 2015). All PCR amplifications were performed in a 25 µl mixture which consisted of 12.5 µl Green Taq PCR master mix, 1.5 µl of each primer, 5.0 µl DNA and 4.5 µl dH_2O. Following gradient PCR, an annealing temperature range of 51– 55 °C amplified the gene successfully. The temperature producing the best amplicon band was selected for conventional PCR. The reaction mixture was subjected to 38 amplification cycles in the T100™ Thermal Cycler under the following conditions; initial denaturation at 94 °C for 3 minutes, denaturation at 94 °C for 30 seconds, annealing at 55 °C for 30 seconds, extension at 72 °C for 30 seconds and final extension at 72 °C for 2 minutes. Following amplification, the PCR products were analysed by gel electrophoresis on a 1.5 % gel run at 50 V for 70 minutes to view the product size of 150 bp. The unpurified PCR product was sequenced at Inqaba Biotec (Pretoria, South Africa). Mkize *et al* (2016) cited the sequencing procedure at Inqaba as follows; the sequencing is conducted using the ABI V3.1 Big dye kit following the manufacturer's instructions on an ABI 377 automated sequencer. The Zymo Seq clean–up kit is used to clean products, thereafter products are inserted into the ABI 3500XL Genetic Analyzer (with a 50 cm array) using POP7.

4.2.4 Phylogenetic and Statistical analysis

The programme BioEdit was used to edit and align the sequences using ClustalW and mutations were checked against raw sequence data. The structuring of mitochondrial DNA diversity amongst the sample populations were tested using base pair differences. The relationship between the sequences was then inferred by a phylogenetic tree, which was constructed using the (NJ) heuristic search (1000 replicates) algorithm in PAUP. Minimum spanning haplotype networks were constructed using NETWORK software and the appropriate polymorphism data tests based on the haplotype network were conducted using DnaSP v.6.12.01. The following tests were carried out; haplotype diversity, nucleotide diversity, Fu and Li's test and Tajima's test. Genetic and population structure was analysed using Arlequin 3.1. The principal coordinates analysis (PCoA) was carried out using the euclidean

similarity index and a transformation component of c = 2. Genetic divergence estimates were calculated using the Poisson corrected model in MEGA7.

4.3 Results

4.3.1 Genetic diversity

The 12S rRNA mitochondrial DNA gene was successfully sequenced in 15 selected *B. flavicollis* individuals with at least one representative from each of the 13 populations. Four individuals were used as the outgroup species, *D. flavifilis* and *D. uncinatus*. Analysis of all 19 sequences yielded a total of 19 unique haplotypes. The DNA polymorphism data (Table 4.2) were generated based on the haplotype network. A high haplotype diversity (H_d = 1.000) and low nucleotide diversity (π_n = 0.16438) were observed, with an average number of nucleotide differences (k = 15.12281) between sequences. The large haplotype diversity value suggests a strong definitive measure of uniqueness between the haplotypes and low nucleotide diversity value is indicative of the presence of genetic variation of the 12S rRNA gene between sampled populations.

The appropriate statistical tests such as Tajima's test, Fu and Li's D, F statistic and Fu's F statistic were calculated to gain insight regarding the frequency of mutations of the 12S rRNA gene. The statistical tests however, were not significant. Tajima's D and Fu and Li statistics were used to assess if the sequence data show deviation from neutrality. The output value for Tajima's D test was negative (- 0.36288), signifying an excess of low frequency polymorphisms relative to the expectation. The Fu and Li's neutrality tests were not considered to be significant as p values greater than 0.10 were observed.

Table 4.2: Gene polymorphism data based on the 12S rRNA haplotype network of *B. flavicollis*.

Nucleotide Analysis	**Parameter Estimates**
Haplotype (gene) diversity	1.000
Variance of haplotype diversity	0.00029
Standard deviation of haplotype diversity	0.017
Nucleotide diversity per site (π)	0.16438
Sampling variance π	0.0000899
Standard deviation of π	0.00948
Average number of nucleotide differences, k	15.12281
Theta (per sequence) from Eta	16.59462
Theta (per site) from Eta	0.18038
Tajima's D	- 0.36288 (p > 0.10)
Fu and Li's D* statistic	0.44481 (p > 0.10)

Fu and Li's F* statistic	0.23888 (p > 0.10)
Fu's Fs statistic	- 8.358

* P values less than 0.10 were considered to be significant

4.3.2 Population genetic differentiation

The analysis of molecular variance (AMOVA) (Table 4.3) was consistent with the haplotype polymorphism data, indicating genetic diversity amongst individuals. Individuals were grouped based on their sample localities, *D. flavillis* and *D.uncinatus* individuals were the outgroup. The percentage of variation was perceived as follows; among groups (1.26 %), among populations within groups (36.20 %) and within populations (62.53 %). The genetic variation between *B. flavicollis* populations based on their locality is considerably high, indicating that each locality exhibits a great level of genetic diversity from each other thus supporting the unique haplotypes for each sequence. The variation observed between *B. flavicollis* individuals within the same population is relatively high and significant as it supports the haplotype diversity. The fixation index value (F_{ST} = 0.37466), used to measure of differentiation based on the population structure, is considerably low, suggesting higher levels of homozygosity within each population.

Table 4.3: AMOVA results for *B. flavicollis* 12S rRNA gene.

	d.f	**Sum of squares**	**Variance components**	**Percentage of variance (%)**	**Fixation index**
Among groups	14	111.605	0.09581	1.26	0.37466 (F_{ST})
Among populations (within groups)	2	15.000	2.75000	36.20	
Within populations	2	9.500	4.75000	62.53	

d.f = degrees of freedom

The principal coordinates analysis (PCoA) (Figure 4.1) supported the observed haplotype results. Ordinated to the far right of the axis was 52.6 % of the variation which included the outgroup, *D. flavifilis*, whilst the remaining 47.4 % of the individuals ordinated to the left of the axis with outgroup *D. uncinatus*. Close clustering is observed with individuals from Charangwa and Dombomupunga.

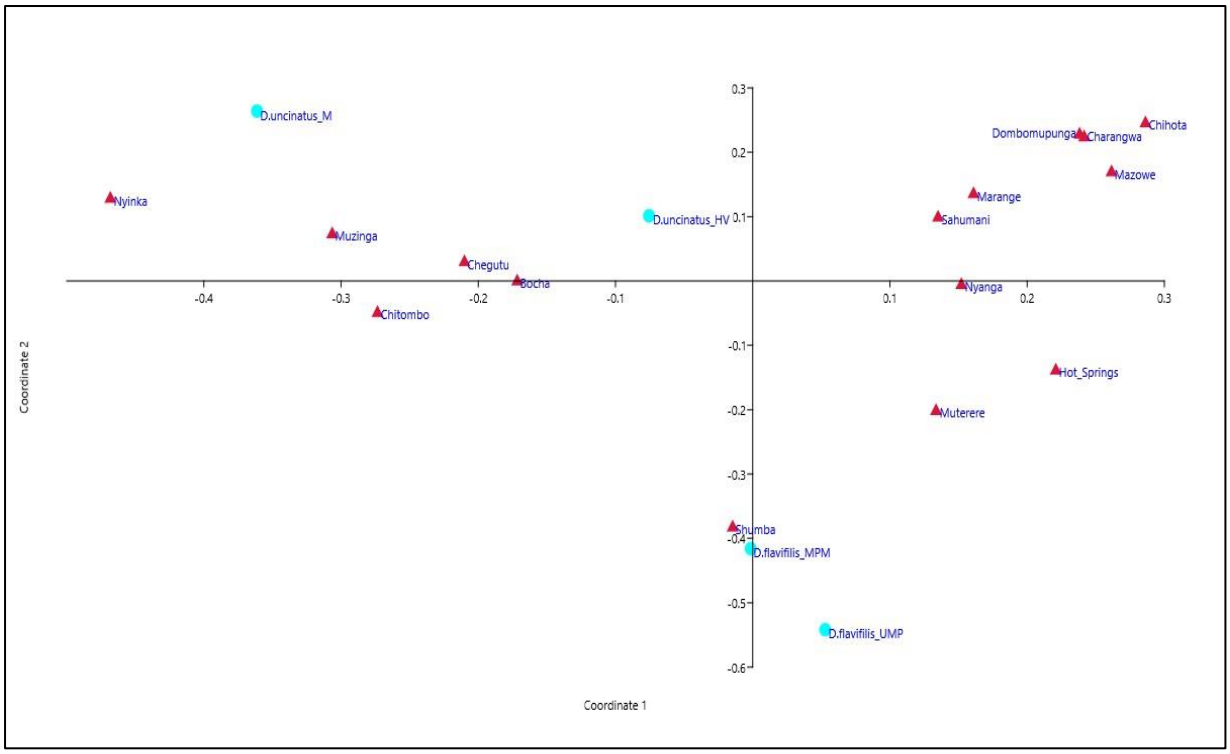

Figure 4.1: Principal coordinates analysis of *B. flavicollis* species based on the 12S rRNA mitochondrial gene. Red triangles = *B. flavicollis* individuals; Blue dots = *D. flavifilis* and *D. uncinatus* (Outgroup).

No haplotype sharing was observed for the 12S rRNA gene (Figure 4.2). All 19 sequences formed individual and unique haplotypes with 31 median vectors. Although no admixed populations were observed, individuals from Charangwa (H1) and Dombomupunga (H15) share a close relationship based on sequence analysis, which is supported by their close proximity on the phylogenetic tree and PCoA scatter plot. High sequence similarity is also observed between Shumba (H11) and the *D. flavifilis* individuals, again supported by phylogenetic and PCoA analysis.

Estimates of evolutionary divergence between each sequence (Table 4.4) were conducted using the Poisson correction model (Zuckerland and Pauling, 1965). All estimates less than 0.1 are shaded in grey. A low estimate is indicative of greater genetic similarity between sequences of said locality. The sequences representatives of Charangwa (A1) and Dombomupunga (O1) expresses a low pairwise estimate value ($p = 0.01$), this is representative of the strong sequence identity between the two regions for the 12S rRNA gene. Charangwa (A1) and Mazowe (C1), as well as Dombomupunga (O1) and Mazowe (C1) have low sequence divergence estimates of $p = 0.04$ and $p = 0.03$, respectively, signifying sequence similarity. However, none of the sequences are genetically identical to each other ($p \neq 0.00$) and it can be said that individuals exhibit genetic diversity from each other based on locality. ($p \neq 0.00$) and it can be said that individuals exhibit genetic diversity from each other based on locality.

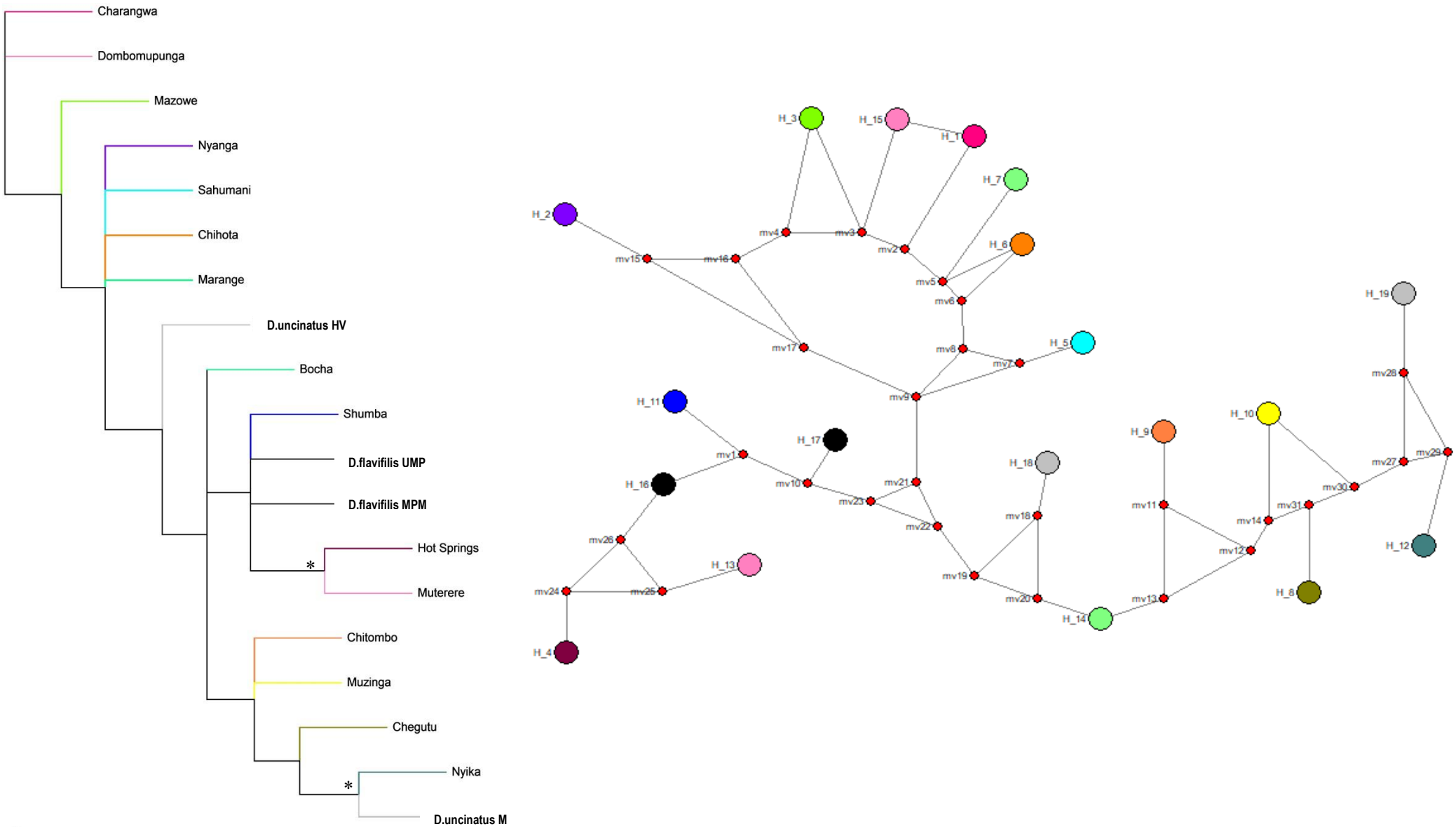

Figure 4.2: Heuristic search (1000 replicates) tree and haplotype network showing genetic diversity of *B. flavicollis* individuals using 12S rRNA gene. **A**: PAUP generated Neighbour joining tree, all clades with strong bootstrap values (>70) are denoted by an asterisk *. **B**: Haplotype network, red circles indicate median vectors.

Table 4.4: Pairwise estimates of genetic distance based on 12S rRNA mitochondrial gene individual sequence data of *B. flavicollis*. Genetic distances are shown below the diagonal and the standard error(s) are shown above the diagonal obtained by bootstrap procedure (1000 replicates).

	A1	B1	C1	D1	E1	F1	G1	H1	I1	J1	K1	L1	M1	N1	O1	P1	P2	Q1	Q2
A1		0.04	0.02	0.05	0.04	0.03	0.04	0.05	0.05	0.05	0.05	0.07	0.05	0.05	0.01	0.04	0.05	0.05	0.06
B1	0.13		0.03	0.05	0.03	0.04	0.04	0.05	0.05	0.05	0.05	0.06	0.04	0.05	0.04	0.05	0.05	0.04	0.06
C1	0.04	0.10		0.05	0.03	0.03	0.04	0.05	0.05	0.05	0.05	0.06	0.05	0.04	0.02	0.05	0.05	0.04	0.06
D1	0.18	0.19	0.19		0.05	0.04	0.05	0.05	0.05	0.06	0.04	0.07	0.03	0.05	0.05	0.04	0.04	0.05	0.06
E1	0.13	0.10	0.09	0.20		0.03	0.04	0.05	0.05	0.05	0.04	0.06	0.04	0.04	0.04	0.04	0.04	0.03	0.05
F1	0.07	0.12	0.07	0.17	0.07		0.03	0.05	0.05	0.06	0.05	0.06	0.05	0.04	0.03	0.04	0.04	0.04	0.05
G1	0.14	0.14	0.13	0.23	0.13	0.09		0.05	0.05	0.06	0.05	0.06	0.05	0.05	0.04	0.05	0.05	0.04	0.06
H1	0.20	0.22	0.20	0.23	0.20	0.20	0.22		0.04	0.04	0.05	0.04	0.05	0.04	0.05	0.05	0.04	0.04	0.04
I1	0.22	0.22	0.23	0.23	0.23	0.22	0.23	0.14		0.03	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.04	0.04
J1	0.25	0.23	0.23	0.27	0.23	0.26	0.26	0.12	0.07		0.05	0.04	0.05	0.03	0.05	0.05	0.05	0.04	0.03
K1	0.23	0.19	0.23	0.12	0.18	0.19	0.20	0.19	0.19	0.22		0.06	0.04	0.04	0.06	0.02	0.02	0.04	0.06
L1	0.32	0.32	0.32	0.32	0.29	0.30	0.32	0.12	0.19	0.14	0.27		0.06	0.05	0.07	0.06	0.06	0.05	0.04
M1	0.18	0.17	0.17	0.09	0.15	0.18	0.20	0.25	0.20	0.22	0.14	0.32		0.04	0.05	0.03	0.04	0.05	0.05
N1	0.20	0.18	0.17	0.20	0.15	0.17	0.19	0.14	0.12	0.10	0.15	0.20	0.15		0.05	0.04	0.04	0.03	0.04
O1	0.01	0.12	0.03	0.19	0.12	0.08	0.14	0.22	0.23	0.23	0.25	0.33	0.17	0.19		0.05	0.05	0.04	0.06
P1	0.17	0.18	0.18	0.12	0.15	0.17	0.20	0.20	0.19	0.22	0.06	0.29	0.10	0.15	0.18		0.02	0.04	0.06
P2	0.19	0.19	0.20	0.12	0.15	0.17	0.22	0.18	0.19	0.22	0.04	0.29	0.17	0.14	0.20	0.06		0.04	0.06
Q1	0.17	0.17	0.13	0.23	0.10	0.13	0.14	0.14	0.17	0.17	0.15	0.22	0.19	0.08	0.15	0.15	0.14		0.04
Q2	0.26	0.27	0.26	0.26	0.22	0.25	0.26	0.15	0.17	0.10	0.27	0.12	0.20	0.13	0.25	0.26	0.26	0.17	

A1 = Charangwa; B1 = Nyanga; C1 = Mazowe; D1 = Hot Springs; E1 = Sahumani; F1 = Chihota; G1 = Marange; H1 = Chegutu; I1 = Chitombo; J1 = Muzinga; K1 = Shumba; L1 = Nyika; M1 = Muterere; N1 = Bocha; O1 = Dombomupunga; P1 = *D. flavifilis* (UMP– University of Mpumalunga); P2 = *D. flavifilis* (MPM– Mbonisweni Mpumalunga); Q1 = *D. uncinatus* (HV – Honde valley); Q2 = *D. uncinatus* (M– Muterere)

4.4 Discussion

Phylogenetic investigation of millipedes most often utilizes mitochondrial DNA as it provides useful information when comparing phylogenetic relationships due to its significantly fast evolution rate (Hwang and Kim, 1999; Caporaso *et al*., 2011; Brewer *et al*., 2012). The mitochondrial protein coding 12S rRNA gene is among the less commonly used markers in phylogenetic studies or is often used in conjunction with other mtDNA markers (Kuznetsova *et al*., 2002; Shearer *et al*., 2008). In the present study, the 12S rRNA gene marker was used as a stand–alone marker and revealed distinct population structure of *B. flavicollis* based on phylogeography.

The generated polymorphism data (Table 4.2), indicates a haplotype diversity of H_d = 1.000. This is suggestive of unique haplotypes being present. Each individual exhibits enough low frequency polymorphisms to be considered genetically distinct. Geographic location should be considered as a contributing factor to the distinction observed because divergence is likely to occur over time due to environmental barriers (Nistelberger *et al*., 2014). Although 19 unique haplotypes (Figure 4.2B) were observed, phylogenetic analysis showed that Charangwa and Dombomupunga, and Mazowe and the outgroup *D. flavilis* populations expressed high levels of similarity in comparison to other populations, as the aforementioned populations formed linked haplotypes, protruding from a common median vector. The strong presence of median vectors in the haplotype network is representative of common ancestry. The high number of median vectors observed in the generated haplotype network is reflective of the lack of representative samples for each population group in this study. Increased taxon and sample size have a strong and positive effect on the accuracy and reliability of the phylogenetic analyses (Zwickl and Hillis, 2002). Data exploration is of high importance especially when using network methods to gain insight of phylogenetic diversity (Joly *et al*., 2007). Homoplasies, unsampled haplotypes, missing or insufficient data need to be paid further attention when constructing networks in order to obtain a reliable representation of the genetic variation (Joly *et al*., 2007).

AMOVA analysis (Table 4.3) identified a high variation occurring within populations (> 60 %) and the variation occurring among populations to be lower (< 40 %). These results could be strongly influenced by the experimental data set as only one representative individual per locality was selected for DNA sequencing. The high variation occurring within populations can therefore, be attributed to the outgroups, *D. flavifilis* and *D. uncinatus*, that had two sequence representatives. The observed F_{ST} value (F_{ST} = 0.37466) is suggestive of a predominantly homozygous populations. This is in concordance with the phylogenetic findings of this study because homozygous populations suggest the presence of barriers to gene flow and increased likelihood of unique haplotypes being formed. The low vagility of *B. flavicollis* supports the idea of genetic distinction among populations (Mwabvu *et al*., 2007).

The phylogenetic tree (Figure 4.2A) strongly supports and represents the findings of the haplotype network and PCoA analysis (Figure 4.1). The genetic closeness and clade forming of populations can be visualized on the haplotype network. The population of Nyika formed a clade with *D. uncinatus*, it is seen that both these populations share common median vectors and are ordinated closer together on the left of the axis in PCoA analysis. Definitive clustering and close ordination is observed between Charangwa and Dombomupunga, as well as, Shumba and *D. flavifilis*. The close clustering is supported by the genetic divergence estimates (Table 4.4) with estimates ranging between 0.01 – 0.04. Such estimates signify high similarity between individuals. It is interesting to note that both Nyika and Shumba are sampled localities from the province of Masvingo and share high infinity to *Doratogonus* based on phylogenetic analysis. One can assume that there is shared ancestry between the two genera.

Walker *et al.* (2009) conducted the first phylogeographic study of the *Narceus* to investigate its genetic structure and distribution using mtDNA. Northern populations of the millipede were expected to exhibit lower levels of genetic diversity as opposed to southern populations in the east of Mississippi (Walker *et al*., 2009). Findings expressed distinct genetic isolation for all sampled populations based on complex history of Pleistocene climate change, which influenced diversity in each population (Walker *et al*., 2009). Similarly, the present study expressed strong genetic divergence which is likely influenced by environmental stresses and adaptations (Mwabvu *et al*., 2007), which impacts the distribution of *B. flavicollis*. The divergence estimates strongly support the distinct genetic diversity observed across the large geographic distance of each population.

The results of this study express distinct genetic structure of *B. flavicollis* populations. These results are in line with the expectation of phylogeographic studies. After studying the cytochrome c oxidase (CO1) region of *Bicoxidens* species, Tinagao *et al.* (2017) determined strong divergence among the species *B. friendi, B. brincki* and *B. flavicollis*, as well as, divergence between individuals within the same population group. As single representatives for each population were available for this study, the conclusion of genetic divergence between *B. flavicollis* individuals within a population group cannot be drawn. Morphological methods alone will not be able to determine the genetic distinction and structure observed, hence the use of molecular methods (Bond and Sierwald, 2002) utilizing mtDNA for this study.

4.5 Conclusion

Bicoxdens flavicollis individuals exhibited distinct phylogenetic diversity for each locality which was evident from the unique haplotypes formed. This was the expected outcome as it is not uncommon for individuals to express high genetic variation within isolated populations although belonging to the same species, however the low sample size can likely be the reason for the strongly observed divergence. Based on this study, it is highly plausible that genetic makeup can contribute to the

phenotypic variation observed amongst populations, however, one must not rule out adaptation to environmental stress. *Bicoxidens flavicollis* based on analysis of the 12S rRNA gene can be said to be a genetic divergent species, exhibiting low frequency polymorphisms between geographic locations. This study suggested strong genetic diversity amongst populations of *B. flavicollis* across Zimbabwe. However, due to the low sample size and the inability to amplify the gene in more samples, there remains a gap in fully understanding the genetic diversity of the sampled populations. This allows for further research to be conducted using the same genetic marker in a larger sample pool or it can be further utilized to determine the population diversity of new genera within the myriapod subphylum.

4.6 References

Bogyó, D., Magura, T., D. Nagy, D. and Tothmeresz, B. (2015). Distribution of millipedes (Myriapoda, Diplopoda) along a forest interior – forest edge – grassland habitat complex. *ZooKeys*, 510: 181–195.

Bond, J. and Sierwald, P. (2002). Cryptic speciation in the *Anadenobolus excisus* millipede species complex on the island of Jamaica. *Evolution*, 56(6): 1123–1135.

Bond, J., Beamer, D., Hedin, M. and Sierwald, P. (2003). Gradual evolution of male genitalia in a sibling species complex of millipedes (Diplopoda : Spirobolida : Rhinocricidae : Anadenobolus). *Invertebrate Systematics*, 17(6): 711–717.

Brewer, M., Sierwald, P. and Bond, J. (2012). Millipede Taxonomy after 250 Years: Classification and Taxonomic Practices in a Mega–Diverse yet understudied Arthropod Group. *PLoS ONE*, 7(5): p.e37240.

Caporaso, J., Lauber, C., Walters, W., Berg–Lyons, D., Lozupone, C., Turnbaugh, P., Fierer, N. and Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, 108(1): 4516–4522.

Foran, D., Fischer, A. and Stoloff, M. (2015). A Comparison of Mitochondrial DNA Amplification Strategies for Species Identification. *Journal of Forensic Investigation*, 3(2): 1 – 7.

Galtier, N., Nabholz, B., Glémin, S. And Hurst, G. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology*, 18(22): 4541–4550.

Golovatch, S. and Kime, R. (2009). Millipede (Diplopoda) distributions: A review. *Soil organisms*, 81(3): 565–597.

Hurst, G. and Jiggins, F. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B: Biological Sciences*, 272(1572): 1525–1534.

Hwang, U. and Kim, W. (1999). General properties and phylogenetic utilities of nuclear ribosomal DNA and mitochondrial DNA commonly used in molecular systematics. *The Korean Journal of Parasitology*, 37(4): 215–228.

Joly, S., Stevens, M. and van Vuuren, B. (2007). Haplotype Networks Can Be Misleading in the Presence of Missing Data. *Systematic Biology*, 56(5): 857–862.

Kuznetsova, M., Kholodova, M. and Luschekina, A. (2002). Phylogenetic Analysis of Sequences of the 12S and 16S rRNA Mitochondrial Genes in the Family Bovidae: New Evidence. *Russian Journal of Genetics*, 38(8): 942–950.

Lavelle, P., Decaëns, T., Aubert, M., Barot, S., Blouin, M., Bureau, F., Margerie, P., Mora, P. and Rossi, J. (2006). Soil invertebrates and ecosystem services. *European Journal of Soil Biology*, 42: S3–S15.

Mandal, S., Chhakchhuak, L., Gurusubramanian, G. and Kumar, N. (2014). Mitochondrial markers for identification and phylogenetic studies in insects – A Review. *DNA Barcodes*, 2(1).

Minelli, A. (2015). *Treatise on Zoology – Anatomy Taxonomy Biology. The Myriapoda Volume 2*. 1st ed. Brill.303–304.

Mkize, L., Mukaratirwa, S. and Zishiri, O. (2016). Population genetic structure of the freshwater snail, Bulinus globosus, (Gastropoda: Planorbidae) from selected habitats of KwaZulu–Natal, South Africa. *Acta Tropica*, 161: 91–99.

Mwabvu, T. (2000). Two new species *Bicoxidens* Attems, 1928, from Zimbabwe (Diplopoda: Spirostreptida: Spirostreptidae). *African Zoology*, 35(1): 115–119.

Mwabvu, T., Hamer, M. and Slotow, R. (2007). A taxonomic review of the southern African millipede genus, *Bicoxidens* Attems, 1928 (Diplopoda: Spirostreptida: Spirostreptidae), with the description of three new species and a tentative phylogeny. *Zootaxa*, (1452): 1–23.

Mwabvu, T., Hamer, M. and Slotow, R. (2009). A new species of *Bicoxidens* Attems, 1928 (Diplopoda, Spirostreptida, Spirostreptidae) from northern Zimbabwe. *ZooKeys*, 7: 75–81.

Mwabvu, T., Hamer, M., Slotow, R. and Barraclough, D. (2010). A revision of the taxonomy and distribution of *Archispirostreptus* Silvestri 1895 (Diplopoda, Spirostreptida, Spirostreptidae), and description of a new spirostreptid genus with three new species. *Zootaxa*, (2567): 1–49.

Mwabvu, T., Lamb, J., Slotow, R., Hamer, M. and Barraclough, D. (2013). Is Millipede Taxonomy Based on Gonopod Morphology Too Inclusive? Observations on Genetic Variation and Cryptic Speciation in *Bicoxidens flavicollis* (Diplopoda: Spirostreptida: Spirostreptidae). *African Invertebrates*, 54(2): 349–356.

Mwabvu, T., Lamb, J., Slotow, R., Hamer, M. and Barraclough, D. (2015). Do cytochrome c oxidase 1 Gene Sequences Differentiate Species of Spirostreptid Millipedes (Diplopoda: Spirostreptida: Spirostreptidae)?. *African Invertebrates*, 56(3): 651–661.

Nistelberger, H., Byrne, M., Coates, D. and Roberts, J. (2014). Strong Phylogeographic Structure in a Millipede Indicates Pleistocene Vicariance between Populations on Banded Iron Formations in Semi–Arid Australia. *PLoS ONE*, 9(3): p.e93038.

Shearer, T., van Oppen, M., Romano, S. and Wörheide, G. (2008). Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Molecular Ecology*, 11(12): 2475–2487.

Tamura, K., Nei, M. and Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor–joining method. *Proceedings of the National Academy of Sciences*, 101(30): 11030–11035.

Tinago, T., Mwabvu, T. and MacDonald, A. (2017). Evidence of multiple divergent mitochondrial lineages within the southern African diplopod genus *Bicoxidens* Attems, 1928 (Spirostreptida). *African Zoology*, 52(4): 229–235.

Wiens, J. and Penkrot, T. (2002). Delimiting Species Using DNA and Morphological Variation and Discordant Species Limits in Spiny Lizards (Sceloporus). *Systematic Biology*, 51(1): 69–91.

Wojcieszek, J. and Simmons, L. (2013). Divergence in genital morphology may contribute to mechanical reproductive isolation in a millipede. *Ecology and Evolution*, 3(2): 334–343.

Zuckerkandl, E. and Pauling, L. (1965). *Evolutionary divergence and convergence in proteins*. New York: Academic Press: 97–166.

Zwickl, D. and Hillis, D. (2002). Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology*, 51(4): 588–598.

CHAPTER 5

GENERAL CONCLUSION

Being endemic to southern Africa, the genus *Bicoxidens* Attems, 1928 can be found in diverse vegetation, ranging from savannah to montane habitats (Mwabvu *et al*., 2013). The genus expresses a variety of phenotypic differences which include body size and colour (Mwabvu *et al*., 2007; Mwabvu *et al*., 2013). The most widely distributed species is *B. flavicollis*, which occurs in Zimbabwe even extending to western Mozambique. Interestingly, *B. flavicollis* individuals exhibited phenotypic differences similar to those observed in individuals within the genus (Mwabvu *et al*., 2007). Therefore, further assessment of this species was imperative, focusing particularly on the distribution of *B. flavicollis* (Mwabvu *et al*., 2013).

Traditional morphology–based characterization of gonopods is considered too inclusive and cannot identify the high levels of variation which exist between populations (Mwabvu *et al*., 2013; Mwabvu *et al*., 2015). With the evidence of genital divergence occurring at a slower rate than genetic divergence, taxonomy based on morphology alone cannot be considered reliable when conducting phylogeographic studies (Mwabvu *et al*., 2013; Tinago *et al*., 2017). Molecular techniques are considered more useful to distinguish between otherwise morphologically identical species because each species has a unique sequence which will assist in categorization and provides additional information (Pires and Marinoni, 2010; Friedheim, 2016). With this in mind, the implementation of molecular techniques is better suited to verify species divergence. The present study made use of the protein coding mtDNA genes; 12S rRNA and 16S rRNA to identify variation among *B. flavicollis* individuals from different populations.

5.1 Significant findings

This study determined the phylogeographic nature of *B. flavicollis* using two mitochondrial DNA genes. Analysis using the 16S rRNA gene suggested mixed ancestry and similarity between populations, as shown by shared haplotypes. Unique haplotypes were identified, which indicated strong genetic structure within certain isolated populations. The populations with shared haplotypes likely host individuals with conserved genes between them based on random sampling. Isolated populations in most circumstances host individuals with high genetic variation although from the same species. Analysis of the 12S rRNA gene provided evidence of strong genetic divergence between populations as shown by the unique haplotypes that were observed. However, each population had one representative, which could have contributed to bias, as replicates are more likely to detect significant similarity or differentiation if present. Although the results for both mitochondrial markers are different from each other, both are still significant for this polymorphic species. Sample size for each mitochondrial marker is the most likely factor for the variation in results observed for both genes. The 16S rRNA gene having had more success with amplification produced more

sequences for each locality to be utilized in the study, hence exhibiting admixed haplotypes that existed between localities.

5.2 Conclusions and recommendations

Morphology based taxonomy may not reveal the variation present in individuals (Mwabvu *et al*. 2013). Therefore, molecular methods must be used in conjunction with morphological methods to distinguish and reassess taxa (Bond and Sierwald, 2002; Friedheim, 2016). Assessment of *B. flavicollis* revealed high levels of genetic distinction between populations. Those populations which exhibited genetic similarity demonstrated low genetic divergence estimates and statistically significant polymorphism data to support the likelihood of shared ancestry or gene exchange. Due to the low vagility and dispersal rate of millipedes, it would be beneficial and informative to revise morphology–based descriptions of taxa as speciation may have occurred due to environmental barriers and adaptations. Sampling of taxon and individuals within a population must be increased to better perceive the genetic structure of the species being studied. The analysis conducted on *B. flavicollis* can be implemented on other genera of millipede and answer questions of cryptic speciation.

5.3 References

Bond, J. and Sierwald, P. (2002). Cryptic Speciation In The *Anadenobolus excisus* Millipede Species Complex On The Island Of Jamaica. *Evolution*, 56(6): 1123–1135.

Friedheim, S. (2016). Comparison of Species Identification Methods: DNA Barcoding versus Morphological Taxonomy. *Manoa Horizons*, 1: 74 – 86.

Mwabvu, T., Hamer, M. and Slotow, R. (2007). A taxonomic review of the southern African millipede genus, *Bicoxidens* Attems, 1928 (Diplopoda: Spirostreptida: Spirostreptidae), with the description of three new species and a tentative phylogeny. *Zootaxa*, (1452): 1– 23.

Mwabvu, T., Lamb, J., Slotow, R., Hamer, M. and Barraclough, D. (2013). Is Millipede Taxonomy Based on Gonopod Morphology Too Inclusive? Observations on Genetic Variation and Cryptic Speciation in *Bicoxidens flavicollis* (Diplopoda: Spirostreptida: Spirostreptidae). *African Invertebrates*, 54(2): 349– 356.

Mwabvu, T., Lamb, J., Slotow, R., Hamer, M. and Barraclough, D. (2015). Do cytochrome c oxidase 1 Gene Sequences Differentiate Species of Spirostreptid Millipedes (Diplopoda: Spirostreptida: Spirostreptidae)?. *African Invertebrates*, 56(3): 651– 661.

Pires, A. and Marinoni, L. (2010). DNA barcoding and traditional taxonomy unified through Integrative Taxonomy: a view that challenges the debate questioning both methodologies. *Biota Neotropica*, 10(2): 339– 346.

Tinago, T., Mwabvu, T. and MacDonald, A. (2017). Evidence of multiple divergent mitochondrial lineages within the southern African diplopod genus *Bicoxidens* Attems, 1928 (Spirostreptida). *African Zoology*, 52(4): 229– 235.