Evolution of tangle-veined flies: systematics, biogeography and functional traits in southern African Nemestrininae (Nemestrinidae)

by Genevieve Lee Theron

Thesis presented in fulfilment of the academic requirements for the degree of Doctor of Philosophy at the University of KwaZulu-Natal

School of Life Sciences,

University of KwaZulu-Natal

Pietermaritzburg

August 2021

Supervisor: Dr. Timotheus van der Niet

Co-supervisors: Prof. Bruce C Anderson, Prof. Allan G Ellis and Prof. Steven D Johnson

PREFACE

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

The work contained in this thesis was carried out in the Republic of South Africa at the University of KwaZulu-Natal under the supervision of Dr. Timotheus van der Niet, Prof. Allan G. Ellis, Prof. Bruce C. Anderson and Prof. Steven D. Johnson.

[redacted]

Genevieve L. Theron

29 July 2021

I certify that the above statement is correct

[redacted]

Dr. Timotheus van der Niet (supervisor)

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

[redacted]

Dr. Timotheus van der Niet (supervisor)

[redacted]

Prof. Allan G. Ellis (co-supervisor)

[redacted]

Prof. Bruce C. Anderson (co-supervisor)

[redacted]

Prof. Steven D. Johnson (co-supervisor)

DECLARATION 1 - PLAGIARISM

I, Genevieve Lee Theron, 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.

Genevieve L. Theron

DECLARATION 2 - PUBLICATIONS

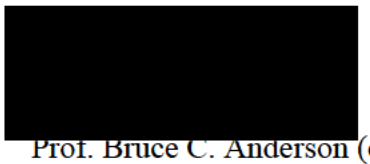
Chapter 2 of this thesis has been published in the Biological Journal of the Linnean Society in 2019. Chapters 3-5 have been prepared as manuscripts that will be submitted to relevant journals for publication.

Author contributions to all manuscripts are as follows: GLT, AGE, BCA, SDJ and TVD contributed to field collections. GLT analysed data and wrote the manuscripts with the assistance of AGE, BCA, SDJ and TVD.

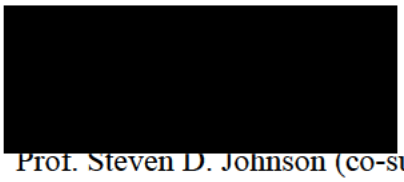
Genevieve L. Theron

Dr. Timotheus van der Niet (supervisor)

Prof. Allan G. Ellis (co-supervisor)



Prof. Bruce C. Anderson (co-supervisor)



Prof. Steven D. Johnson (co-supervisor)

Abstract

The evolution of traits and biogeography of the three southern African endemic genera of the Nemestrinidae: *Moegistorhynchus*, *Prosoeca* and *Stenobasipteron.* These genera are of particular interest due to the exaggerated mouth parts of some species and their role as important pollinators of numerous plants, including rare and endangered species. Most taxonomic studies on southern African nemestrinids date back 50 or more years ago, and the group lacks a phylogenetic framework, thus hindering comprehensive study of their systematics, trait evolution and biogeography. In this thesis, I evaluate the boundaries of a species complex in *Prosoeca* and reconstruct a phylogenetic framework for the southern African Nemestrininae. Furthermore, I use the phylogenetic framework to reconstruct the evolution of proboscis length and biogeographic patterns. To delimit species in the *Prosoeca peringueyi* complex, I quantified morphological variation and established whether this was associated with genetic variation within and between gene regions. Phylogenetic analysis of the complex using the mitochondrial COI gene revealed two well-supported clades, that are supported by morphological traits, one of which is described as a new species. Four gene regions were also used to reconstruct a phylogenetic tree of the three southern African Nemestrininae genera, including 58 morphospecies. The topology suggests that a monophyletic *Moegistorhynchus* is sister to a paraphyletic *Prosoeca*, with *Stenobasipteron* nested within *Prosoeca*. Half of the morphospecies in this phylogeny did not correspond to described species, thus highlighting a substantial taxonomic impediment in this group. The phylogenetic tree was used to reconstruct the evolution of proboscis length in the southern African Nemestrininae. Stochastic character mapping showed transitions between all states (short, long and very long), but shifts occurred more frequently from shorter to longer lengths. The ancestral proboscis state was estimated to be longer than the median proboscis length of the clade. Lastly, I reconstructed the biogeographical patterns of the southern

African Nemestrininae. A Fynbos origin during the Miocene was estimated for this clade, with multiple shifts between biomes along the tree. Together, these results illustrate the need for further systematic and taxonomic work in this clade, as well as in the Nemestrinidae more broadly to gain a firmer understanding of their phylogenetic relationships and diversity. The evolution of proboscis length and biome occupancy appear to be labile within this clade. This work provides a phylogenetic framework for the southern African clade of Nemestrininae and contributes to our understanding of the patterns of evolution, diversification and migration of these ecologically important pollinators.

Acknowledgments

A special thanks to Dr. Timo van der Niet who has been my supervisor for the duration of my PhD. I thank him for having much patience and guidance. I also thank Profs Allan Ellis, Bruce Anderson and Steve Johnson for bringing their extensive knowledge and experience to this project. To Kurt Jordaens, Ashley Kirk-Spriggs and John Midgley from the PINDIP project and additionally Kirstin Williams from the KZN museum I owe a great deal of thanks for their patience in teaching me all things Diptera. I thank my partner Johan for always encouraging me to peruse new heights, always putting a smile on my face, for much love and support. My precious lab mates, field buddies and neighbours, especially, Ruth, Saskia and Carolina, I thank you whole heartedly for keeping me sane, always being positive and sharing an enthusiasm for life, good food, hiking, adventure and much more. A warm thanks to Lynx and Beani for being my constant companions, for always wanting cuddles and never saying no to a relaxing walk. The inhabitants of both labs I was privileged enough to be a part of deserve some special thanks for welcoming me with open arms, for distractions, field trips, cake, enlightening conversations and arguments as well as for much academic guidance. I especially thank my parents and my extended family for their unfailing love, support, curiosity, encouragement and enthusiasm. To those who joined me in the field I offer a particular thanks. Thank you to Ruth Cozien, Saskia Klumpers, Terrence Bellingham, Burgert Muller, John Midgley, Miguel Castañeda-Zárate, Nina Hobbhahn, Mathew Rule and Lawrence Harder who put in concerted effort to catch specimens for this project, even though they were not directly involved. I also acknowledge Mountainlands Nature Reserve, Cape Nature, KZN wildlife, Limpopo Province, Mpumalanga Province, Eastern Cape Tourism, Northern Cape and the South African National Parks (SANParks) who provided permits and

access to land for sampling. I acknowledge the funding provided by the National Research Foundation.

This work has been presented in part at two conferences: International Congress of Dipterology 2018 and the Entomological Society of Southern Africa 2019.

Table of Contents

Preface ……...ii

Declaration iii

Abstract v

Acknowledgments..... vii

Table of Contents ix

List of Tables x

List of Figures xii

Chapter 1: Introduction 1

Chapter 2: Key long proboscid fly pollinator overlooked: morphological and molecular analyses reveal a new *Prosoeca* (Nemestrinidae) species 14

Chapter 3: You don't know the half of it: molecular phylogenetics reveals dramatic underestimation of diversity in a key pollinator group (Nemestrinidae) 47

Chapter 4: Labile proboscis length evolution in a lineage of flower-visiting flies (Nemestrinidae) 82

Chapter 5: Repeated biome shifts following a Miocene Fynbos origin for southern African tangle-vein flies in the Nemestrininae subfamily (Nemestrinidae) 114

Chapter 6: Conclusions ……..... 140

References ……150

List of tables

Supplementary tables are indicated with an S prefixed to the numbering

Table 2.1 Summary of consistent morphological traits to distinguish *Prosoeca torquata* sp. nov from *P. peringueyi* and *P. marinusi*.

Table S2.1 Table of localities and specimen details of accessions used for the phylogenetic analyses.

Table S2.2 Primer sequences used to amplify the gene regions included in this study.

Table S3.1 Table of localities and specimen details of accessions used for the phylogenetic analyses.

Table S3.2 Primer sequences used to amplify the gene regions included in this study.

Table 4.1 Morphological measurements and linear regression statistics of morphospecies used to examine static allometry.

Table S4.1 Localities, specimen details and morphological measurements of accessions used for the phylogenetic analyses.

Table S4.2 Number of transitions between states of proboscis length using a discrete character state evolution analysis on 100 phylogenetic trees.

Table 5.1 Average number of transitions for each biome as a source or destination. Bold numbers indicate within-biome speciation without a change in biome. Non-integer numbers result from averaging transitions for 10 phylogenetic trees to account for phylogenetic uncertainty.

Table S5.1 Table of localities and specimen details of accessions used for the dating and biogeography analyses. Number of localities included refers to mapping species into biomes in Fig. 5.1 for further biogeographic analyses.

Table S5.2 Weighted AIC values for the six biogeographic models tested in RASP using BioGeoBEARS.

List of figures

Supplementary figures are indicated with an S prefixed to the numbering

Figure 1.1 Images of *Moegistorhynchus* and *Prosoeca* species *in situ*. (A) *Prosoeca sp. 5* visiting *Satyrium longicauda* (B) Mating pair of *P. ignita* (C) *P. sp. 6* visiting a Lamiaceae sp. (D) *P. robusta* visiting *Protea punctata* (E) *Moegistorhynchus longirostris* visiting *Lapeirousia anceps* (F) *P. sp. 17* visiting *Watsonia watsoniodes.* Photo credits: (A) Miguel Castañeda-Zárate, (B) Jan-Hendrik Keet (C) Ruth Cozien (D & F) Steven Johnson (E) Bruce Anderson

Figure 1.2 Edited map of long-tongued fly (Nemestrinidae) pollination guilds of southern Africa from Johnson (2006b).

Figure 2.1 The frequency of unextended proboscis lengths (in millimetres) measured in the Kosiesberg population (A) and the Naries population (B). Colour-coded bars refer to the operational taxonomic units designated in the genetic analysis: *Prosoeca torquata* (green) and *Prosoeca peringueyi* (blue).

Figure 2.2 Phylogenetic relationships among accessions of the *Prosoeca peringueyi* complex sampled across its range. (A) Bayesian inference majority rule consensus tree of COI sequences for the *P. peringueyi* complex. Posterior probabilities are shown for major nodes. Species names associated with clades represent the new taxonomy proposed in the present study. (B) distribution of allopatric populations of *P. peringueyi* (blue) and *Prosoeca torquata* (green) and of sympatric populations (gradient of blue and green) that were sampled for this study (circles). Tip labels refer to population codes provided in the Supporting Information (Table S2.1).

Figure 2.3 Relationship between body size (in millimetres) and proboscis length (in millimetres) for *Prosoeca peringueyi* (blue) and *Prosoeca torquata* (green) ($F_{1,27} = 33.67$, $r^2 = 0.54$, $P < 0.001$ and $F_{1,33} = 44.80$, $r^2 = 0.58$, $P < 0.001$, respectively; overall model, $F_{3,60} = 161$, $r^2 = 0.89$, $P < 0.001$; and interaction, $F_1 = 5.43$, $P = 0.023$). The shaded area around the trend line indicates the 95% confidence interval.

Figure 2.4 Photographs of the head, dorsal view of the body and the wing of *Prosoeca torquata* (A–C) and *Prosoeca peringueyi* (D–F). Arrow in A indicates the band of white pile after which P. torquata is named. Orange particles visible in D are pollen grains. Abbreviations: *C*, costal vein; *CuA*, anterior branch of cubital vein; *CuP*, posterior branch of cubital vein; *M*, medial vein; M_1, first branch of media; M_2, second branch of media; M_4, fourth branch of media; R_1, anterior branch of radius; R_{2+3}, second branch of radius; R_4, upper branch of third branch of radius; R_5, lower branch of third branch of radius; *Rs*, radial sector; *Sc*, subcostal vein.

Figure 2.5 Photographs of the male genitalia of *Prosoeca peringueyi* (A, B) and *Prosoeca torquata* (C, D) in dorsal (A, C) and ventral view (B, D). Abbreviations: gc, gonocoxite; gs, gonostylus; hyp, hypandrium; ph, phallus.

Figure 2.6 Photographs of *in situ* adult *Prosoeca peringueyi* visiting *Zaluzianskya* sp. (A) and visiting *Lapeirousia silenoides* (B), and *Prosoeca torquata* hovering over *Lapeirousia dolomitica* (C) and mating (D). Photo credits: (A-B) Steven Johnson, (C-D) Florent Grenier.

Figure S2.1 28S haplotype network of ten individuals from the *Prosoeca peringueyi* complex and four Nemestrinidae species. Numbers in parentheses correspond to the hypothetical number of haplotypes that exist between sampled haplotypes. The size of circles and numbers within them indicate the number of individuals sampled for a particular haplotype.

Figure 3.1 Distribution of sampled Nemestrinidae in South Africa, eSwatini and Lesotho. White circles indicate the localities of accessions used for the phylogenetic analysis, blue circles indicate additional sites that were sampled for this study and grey circles are historical records from the Global Biodiversity Information Facility.

Figure 3.2 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the concatenated dataset of COI, 16S, 28S and CAD. Numbers alongside branches refer to posterior probability values and branches are proportional to the number of substitutional changes per site. Letters indicate nodes referred to in the text. Colour blocks indicate the three major clades of *Prosoeca*. Vertical bars show morphospecies delineation for morphology, Automatic Barcode Gap Discovery and Bayesian Poisson Tree Process respectively. Colour sections of the vertical bars indicate clades that contain type specimens of defined genera. Asterisks (*) refer to sequences obtained from GenBank. The branch lengths of *Nemestrinus* and *Terphis* have been artificially shortened for display purposes.

Figure 3.3 The distribution of all Nemestrinidae sampled for the phylogenetic analysis in this study in South Africa, Lesotho and eSwatini. (A) Global biodiversity hotspots are marked on the map in greyscale with the proportion of described (black), cf. (grey) and undescribed (white) morphospecies occuring i) outside the biodiversity hotspots and ii) within the biodiversity hotspots. The relative size of the pie charts is representative for differences in the total number of morphospecies. (B) Biomes of South Africa with bar charts indicating the proportion of described (black), cf. (grey) and undescribed (white) morphospecies found in each biome. The total number of morphospecies found in each biome is indicated in brackets. The total number of mophospecies per biome includes cases where morphospecies occur in multiple biomes.

Figure S3.1 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the separate gene trees for COI. Numbers alongside branches refer to posterior probability support values and branches are proportional to the number of substitutional changes per site.

Figure S3.2 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the separate gene trees for 16S. Numbers alongside branches refer to posterior probability support values and branches are proportional to the number of substitutional changes per site.

Figure S3.3 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the separate gene trees for CAD. Numbers alongside branches refer to posterior probability support values and branches are proportional to substitutional changes.

Figure S3.4 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the separate gene trees for 28S. Numbers alongside branches refer to posterior probability support values and branches are proportional to substitutional changes.

Figure S3.5 (A) 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the concatenated dataset with the selected single accession per species overlayed with black branches. Colour of sequences ID corresponds to designation as described (red), cf. (pink) and undescribed (blue). (B) Frequency distribution of terminal branch lengths of the whole tree with one accession per species in light grey and a tree with only described and cf. morphospecies included in dark grey.

Figure 4.1 Photographs of *in situ* adult *Prosoeca* species visiting flowers. (A) *Prosoeca ganglbaueri* visiting an *Agapanthus sp.* (B) *Prosoeca sp. 6* visiting a Lamiaceae sp. (C) *Prosoeca umbrosa* visiting *Nerine angustifolia* (D) *Prosoeca peringueyi* visiting a *Zaluzianskya sp.* Photo credits: (A) Michael Whitehead, (B) Ruth Cozien, (C) Genevieve Theron, (D) Steven Johnson.

Figure 4.2 Frequency distribution of morphospecies (A) mean proboscis length (mm) and (B) relative proboscis length.

Figure 4.3 Regression of mean body length and mean proboscis length. Grey shading indicates SE. Fly illustrations are to scale relative to each other and are connected to the morphospecies that they represent on the graph by a dotted line.

Figure 4.4 Regressions of log10 body length and log10 proboscis length for 14 morphospecies. Grey shading indicates SE. Asterisks on the figure legend indicate a significate relationship between body and proboscis length for the relevant morphospecies.

Figure 4.5 Continuous character evolution mapping of relative proboscis length with mean absolute proboscis length indicated in grey bars.

Figure 4.6 Evolutionary proboscis state transitions within the nemestrinid clade containing *Moegistorhynchus*, *Prosoeca* and *Stenobasipteron*. The size of the circles is proportional to total number of species sampled within each state: short, 39; long, 13; very long, 6. Arrow thickness is proportional to the number of transitions in each direction averaged over 100 phylogenetic trees.

Figure S4.1 Stochastic character mapping of absolute proboscis length with three character states.

Figure S4.2 Variation in morphological traits measured for the 14 morphospecies for which allometry was assessed. (A) Scatterplot of proboscis length and body length of all measured individuals. Median and standard deviation indicated with dark vertical and horizontal bars. (B) Raincloud plot indicating the distribution of specimen data for log10 of relative proboscis length.

Figure 5.1 The distribution of all Nemestrinidae sampled for this study in the biomes of South Africa and eSwatini as designated by Mucina & Rutherford, 2006.

Figure 5.2 BAYAREALIKE + J reconstruction of ancestral biome reconstruction on one of the ten phylogenetic trees. Single letters corrospond to biomes: (A) Grassland, (B) Fynbos, (C) Forest, (D) Albany Thicket, (E) Savanna, (F) Indian Ocean Coastal Belt, (G) Succulent Karoo. Letter combinations corrospond to polymorphic species that can be found in both biomes.

Figure 5.3 Evolutionary biome transitions within the nemestrinid clade containing *Moegistorhynchus*, *Prosoeca* and *Stenobasipteron*. The size of the circles is proportional to the total number of species sampled within each biome (polymorphic species are included in both biome totals in which they occur): Grassland, 22; Fynbos, 23; Forest, 12; Albany thicket, 1; savanna, 3; Indian Ocean Coastal Belt, 2; Succulent Karoo, 5. The size of the bi-coloured circles is proportional to the number of polymorphic species in each set of biomes. Arrow thickness is proportional to the number of transitions in each direction averaged over 10 phylogenetic trees, and absence of an arrow indicates the lack of any events.

Figure S5.1 Dated maximum clade credibility tree with the estimated crown age next to relevant nodes and 95% confidence interval indicated with blue bars.

Chapter 1

Introduction

The field of molecular systematics has proliferated into an ever-expanding discipline, in part due to technological innovations that make large amounts of genotypic data relatively easy to acquire (Caterino *et al.* 2000). As molecular data became more readily available, the number of fields in which they are used is constantly expanding. In the study of biodiversity, many questions and hypotheses in the fields of molecular systematics, population genetics, ecology and biogeography have been addressed with the use of molecular data. For example, molecular phylogenetic studies have frequently shown that traditional classifications do not always provide a good indication of the phylogenetic relationships of taxa (Soltis *et al.* 1998).

The identification and delimitation of species is vital to most fields of biology for comparison and compilation of knowledge, as well as for assigning conservation priorities. New species are still being described at a high rate worldwide, even in relatively well-studied groups like plants (Christenhusz & Byng 2016) and vertebrates (Tapley *et al.* 2018). It is, however, estimated that for hyper-diverse orders of insects, like Coleoptera, Diptera, Hymenoptera, and Lepidoptera, only 20-28% of species have been described thus far (García-Robledo *et al.* 2020). With innumerable undescribed species, especially in hyper-diverse orders, it is likely that the number of extinctions, as a result of the global increase in disturbance of natural ecosystems, may be larger than current estimates (Brooks *et al.* 2006; Tedesco *et al.* 2014). Thus, efforts to study and catalogue biodiversity need to be given precedence (Sangster 2009). The effects of the taxonomic impediment (ie. large gaps in taxonomic knowledge and a limited number of specialists to address the deficit) on the study of biodiversity is of particular concern for invertebrates which often have high proportions of undescribed or cryptic diversity (New 1999). While species identification and discovery has historically relied on a small number of taxonomists using morphological traits, molecular data has been increasingly used to aid in this endeavour (Hebert *et al.* 2003a, 2003b; Jinbo *et al.* 2011).

Molecular tools such as DNA barcodes have been employed ever-more for species identification and species discovery (Jinbo *et al.* 2011). Many animal species descriptions published since 2003 contain DNA barcodes from holotypes and/or paratypes to assist with species identification (Brown *et al.* 2003; Yoshitake *et al.* 2008; Adamski *et al.* 2009). Species discovery (ie. identifying clusters of individuals as a species) using DNA barcodes has accelerated the identification of cryptic species (Hebert *et al.* 2004). Cryptic species are defined as species that are overlooked because they are superficially morphologically indistinguishable from other species or only differ in ecology or molecular data from these (Bickford *et al.* 2007). The discovery of cryptic species has, for example, caused the re-evaluation of some assumed specialized ecological interactions (Molbo *et al.* 2003; Smith *et al.* 2006). The dependence on a single gene region for species identification has limitations that necessitate the use of other traits and/or genes for recognition of cryptic species, especially in newly evolved systems (Hebert *et al.* 2003a).

A lack of robust identification tools for adequate species delineation can have severe consequences for practical management and ecological or evolutionary studies (Ravaoarimanana *et al.* 2004; Johnson *et al.* 2005; Hayes 2021). For instance, when geographically broad species complexes are subsequently split into species that may be vulnerable due to occupying a limited range, conservation priorities may be affected (Johnson & Linder 1995; Ravaoarimanana *et al.* 2004; Stuart *et al.* 2006). Species delimitation using an integrative taxonomic approach, which aims at considering multiple and complementary independent lines of evidence to delimit diversity, is ideal to ensure the accuracy of the conclusions that are drawn from it (Mishler & Donoghue 1982; Funk & Omland 2003; Dayrat 2005). Conclusions drawn from an integrative taxonomic approach, which includes the use of molecular, morphological, ecological, behavioural, chemical and/ or eco-geographic data, can provide confidence in species boundary delimitation and diversity

counts and can lead to formal taxonomic changes (Puillandre *et al.* 2012b; Carstens *et al.* 2013).

Beyond species relationships and boundaries, phylogenies can also provide insight into the biogeography and evolution of traits within a lineage (Harvey & Pagel 1992). By using a sufficiently realistic model to infer evolutionary changes along all internal nodes of a phylogeny from the measured characteristics of tree tips, ancestral states of a measurable trait, and the ancestral geographic range of population or species can be reconstructed (Ronquist 2004; Joy *et al.* 2016). Ancestral character state reconstruction is an important tool for understanding the origins and patterns of dispersal or the evolution of key features of extant organisms (Maddison 1995; Liberles 2007). The colonisation of a new area or the evolution of a new character may generate new ecological opportunities by providing access to previously unavailable habitats or resources (Simpson 1953). These new ecological opportunities provided by the generation of new traits or the colonisation of new areas may furthermore lead to increased diversification, possibly through adaptive radiation (Schluter 1996; Glor 2010; Yoder *et al.* 2010). On the other hand, physiological limitations, such as climatic tolerance, and/or interactions with biotic partners, such as larval hosts or food plants, may constrain the evolution of traits or biome colonisation (Wiens & Graham 2005; Waterman *et al.* 2011; Duffy & Johnson 2017). The scaling relationships between various traits and body size within species may act as a constraint on trait evolution (Huxley 1924; Gould 1971). Therefore, assessing allometry in combination with trait reconstructions may be useful to identify patterns indicative of evolutionary constraint or character lability as well as directional trends in character evolution.

Studying the distribution of taxa among biomes (ecologically distinct geographic units based on community composition or structure and climatic variables as defined by Rutherford *et al.* (2006)), specifically in conjunction with phylogenetic inference, provides information on

how species may have evolved through space and time (Ronquist & Sanmartín 2011). The use of fossil evidence provides additional data to potentially aid in the accurate dating of evolutionary changes and the history of colonisation and vicariance (Forest 2009). Biogeographic studies examining patterns of diversification and biome shifts through time and space are often examined in plants, but less so for insects. Studies on plant lineages suggest that speciation driven by ecological shifts within a particular biome appear to be much more frequent than speciation driven by ecological shifts among biomes, suggesting that biome conservatism (ie. the tendency of closely related species to inhabit the same biome) is common (Crisp *et al.* 2009). In some clades, however, diversification appears to be linked to repeated biome transitions rather than within-biome radiations (Mitchell *et al.* 2014; Linder & Verboom 2015; Cardillo *et al.* 2017). Thus, with biogeography we can attempt to understand the spatial and temporal components of the origin and diversification of a lineage as well as the historical assembly of biomes (Linder 2005).

This thesis explores the taxonomy, systematics, and evolution of the southern African Nemestrininae. I primarily focus on the southern African endemic genera of *Moegistorhynchus*, *Stenobasipteron* and *Prosoeca*, three of six genera of Nemestrinidae occurring in southern Africa. These three genera belong to the Nemestrininae subfamily and are all considered ecologically important genera as pollinators of a plethora of plant species in southern Africa.

Study area

South Africa is an incredibly biodiverse country, hosting three biodiversity hotspots (Hrdina & Romportl 2017) comprising nine biomes (Mucina & Rutherford 2006), distributed over two regions which predominantly receive rain in either winter or summer (Tyson & Preston-

Whyte 2000). This substantial landscape heterogeneity likely contributes to the region's high biodiversity (Mazijk *et al.* 2021). South Africa contains over 20 000 plant species (Manning & Goldblatt 2012) occupying biomes ranging from Forests with complete canopy cover to xeric Succulent Karoo with sparse low growing vegetation (Mucina & Rutherford 2006). Other groups, like birds and mammals, are also considered relatively diverse in South Africa. Insects have been relatively poorly studied with the best estimates currently suggesting *ca*. 43 000 species (Scholtz & Chown 1995).

Study system

Nemestrinidae is a small, early diverging dipteran family that has *ca*. 277 species worldwide (Barraclough 2017). Nemestrinids are easily recognizable due to the unique diagonal wing veins that give the vein structure a "tangled" appearance, hence the common name "tangle-vein flies". Although their appearance is remarkably variable, species belonging to the family are often relatively large with well-developed proboscides and profusely pubescent bodies (Barraclough 2017).

Extant Nemestrinidae are thought to be obligate nectar feeders (Karolyi *et al.* 2012) and were likely acting as pollinators as early as the late Triassic (Labandeira *et al.* 2007). Species in many nemestrinid genera are thought to be important flower visitors, but are usually considered generalist flower visitors (Devoto & Medan 2006; Potgieter *et al.* 2009). In southern Africa, however, many of the flower species visited form tight associations with their nemestrinid pollinators (Manning & Goldblatt 2000). In some species, reciprocal selection from corolla tube length and from fly proboscis length appears to have led to co-evolutionary arms races between plants and nemestrinid flies (Anderson & Johnson 2008; Pauw *et al.* 2009). Many well-studied nemestrinid species are collectively considered to be

important pollinators of 100-200 species (Fig. 1.1 & 1.2) (Manning & Goldblatt 2000). They are also known to be the sole pollinators for many rare and endangered plant species (Johnson 2006b), making them of particular conservation importance.

Although the feeding habits of several adult southern African Nemestrinidae are well studied, the immature stages of these flies are virtually unstudied (Barraclough 2017). All members of the Nemestrinidae family are assumed to have parasitic larvae. Known larval hosts include a variety of insect orders, ranging from preying mantids (Mantodea: Tarachodidae) (Haenni & Borer 2007) to katydids (Orthoptera: Tettigoniidae) (Kanmiya 1987). The most common hosts are thought to be grasshoppers (Orthoptera: Acrididae) with very high parasite loads found in certain grasshopper populations (York & Prescott 1952; Prescott 1960). In southern Africa the economically important pest, the brown locust, has been reported as a host of *Trichopsidea costata* (Potgieter 1929) but very little is known about the larval habits of the other southern African species (Barraclough 2006).

The Nemestrinidae's closest relative remains tenuous with a recent study suggesting that the they may be sister to the Xylophagidae (Wiegmann *et al.* 2011) and another rather suggesting that the Nemestrinidae are sister to the Acroceridae (Shin *et al.* 2018). Southern Africa is one of the main hotspots for nemestrinid diversity, with representatives of six genera. Of these *Prosoeca* Schiner 1867, *Stenobasipteron* Lichtwardt 1910 and *Moegistorhynchus* Macquart 1840, have received the most attention for their role as important pollinators of countless plant species (Fig. 1.1) (Barraclough 2006). The other three genera; *Nycterimyia* Lichtwardt 1909, *Trichopsidea* Westwood 1839 and *Atriadops* Wandolleck 1897, collectively have just four described species found in southern African and belong to the subfamily Trichopsideinae (Barraclough 2006). In contrast to the debate about the phylogenetic placement of the family, there has been little debate concerning the placement of *Prosoeca*, *Stenobasipteron* and *Moegistorhynchus* within the subfamily Nemestrininae (Bernardi 1973). All of the genera

within the Nemestrininae have a well-developed proboscis, while those in the Trichopsideinae subfamily have proboscides that are absent or non-functional (Bernardi 1973). Within the subfamily Nemestrininae there are six genera, four of which have species found in Africa, whereas the other two genera can be found in the Palearctic, Australian and Neotropical regions (Bernardi 1973). Species of *Nemestrinus* Latreille 1802 can be found in North Africa, while *Stenobasipteron*, *Moegistorhynchus* and *Prosoeca* are endemic to southern Africa. *Stenobasipteron*, with three described species, is restricted to the eastern part of the country (summer-rainfall), *Moegistorhynchus* with four species is limited to the western region (winter-rainfall area) of southern Africa and *Prosoeca*, with a distribution covering the whole of southern Africa, currently has 37 described species (Barraclough 2017). Previous authors have suggested that *Stenobasipteron* is very closely related to, and possibly a junior synonym of, *Prosoeca* (Bernardi 1973; Barraclough 2006).

As many lineages of plants and animals tend to occupy relatively few biomes, the diversity of biome occupation in southern African Nemestrininae is rather unique for a relatively small group (Crisp *et al.* 2009). Together, the South African Nemestrininae likely contain species in all of South Africa's nine biomes (Barraclough 2006), with the possible exception of the true Desert Biome. *Prosoeca*, the largest of the three genera can be found in eight biomes, whereas *Moegistorhynchus* is predominantly a Fynbos genus. *Stenobasipteron weidemanni* is the only species of the southern African Nemestrininae to be found feeding deep within closed-canopy Forests (although various *Prosoeca* species can be found visiting flowers in light patches within Forests or on Forest ecotones). Several undescribed *Stenobasipteron* species have, however, been found in Savanna and Grassland potentially broadening the habitat preferences of this genus (Barraclough 2017; see results).

In addition to their extensive biome occupation, the southern African Nemestrininae exhibit substantial variation in the length of their poboscides and are particularly unique in having

various species with extremely long proboscides. These exaggerated mouthparts allow them to feed as generalists on specialized long-tubed plant species, using proboscides ranging from 4mm to 100mm among different taxa (Barraclough 2006). Taxa with long proboscides may be able to acquire large quantities of nectar from relatively few flower visits and have access to a broader range of feeding niches than species with short proboscides (Haber & Frankie 1989; Martins & Johnson 2013; Klumpers *et al.* 2019). However, long proboscides may also be associated with functional costs such as reduced nectar uptake rate and increased flower handling time (Kunte 2007; Karolyi *et al.* 2013). The uptake of nectar relies on a complex system that makes use of pharyngeal pumping organs, the contraction and dilation of a cibarium and a straw like proboscis (Bauder & Karolyi 2019). The longer the proboscis, the larger the pumping organs required for efficient nectar uptake (Kunte 2007; Karolyi *et al.* 2013).

Although the Nemestrininae of southern Africa are known to be ecologically and evolutionarily important as pollinators, they are poorly studied in terms of their taxonomy, systematics, and evolution. Additionally, each of the three southern African Nemestrininae genera are thought to contain many undescribed species (Barraclough 2006, 2017). The last revision of the genus *Prosoeca* was more than 90 years ago (Bezzi 1924) with only two species described in the last 20 years (Barraclough *et al.* 2018; chapter 2: Theron *et al.* 2020). The lack of attention to this group of ecologically important flies has resulted in many studies referring to undescribed species, making comparison between study systems difficult. Furthermore, no modern phylogeny exists, hindering comparative biology studies in the group.

The importance of southern African nemestrinids in both basic and applied research, means it is essential to have a clear understanding of their diversity, phylogenetic relationships, biogeography, evolution, and diversification. The unique combination of diverse biome

inhabitance and large variation in proboscis lengths makes this an ideal system for the reconstruction of phylogenetic relationships and investigation of trait evolution.

Thesis structure

In this thesis I use an integrated approach to delineating a species complex and reconstruct the phylogenetic relationships of *Moegistorhynchus*, *Stenobasipteron* and *Prosoeca*. Furthermore, I use the phylogeny to identify patterns of diversity and reconstruct proboscis length evolution and biogeography. In chapter two I identify a new *Prosoeca* species after examining the *Prosoeca peringueyi* complex using two gene regions and morphology. In the third chapter, I investigated species level phylogenetic relationships in three southern African endemic nemestrinid genera; *Moegistorhynchus*, *Stenobasipteron* and *Prosoeca*. Additionally, I used independent species delimitation methods to compare the number of operational taxonomic units estimated with my diversity count from integrated morphology. I also use a measure of phylogenetic diversity (PD) to illustrate the contribution of PD by undescribed morphospecies compared to described morphospecies to the clade. In the fourth chapter I make use of the phylogeny, reconstructed in chapter three, to examine the evolution of the extraordinary variation in proboscis length and its association with diversification rate. To do this, I implement both continuous and discrete ancestral state reconstruction methods, together with a trait-dependent Bayesian analysis of macroevolutionary mixtures (BAMM) analysis. Additionally, I investigate patterns of allometry in 14 morphospecies to examine scaling relationships between proboscis length and body size. In the fifth chapter I investigate the ancestral biome state of the clade comprising *Moegistorhynchus*, *Stenobasipteron* and *Prosoeca*. In this chapter I made use of the phylogenetic topology from chapter three and conducted a molecular dating analysis using COI substitution rates to estimate the timing of

divergence. Phylogenetic trees were then used to estimate the ancestral biome of the clade and the frequency of dispersal and speciation within and between biomes. In the final chapter I briefly summarize the findings of the four data chapters and discuss the broader implications of these results as a whole and discuss potential directions for future study.

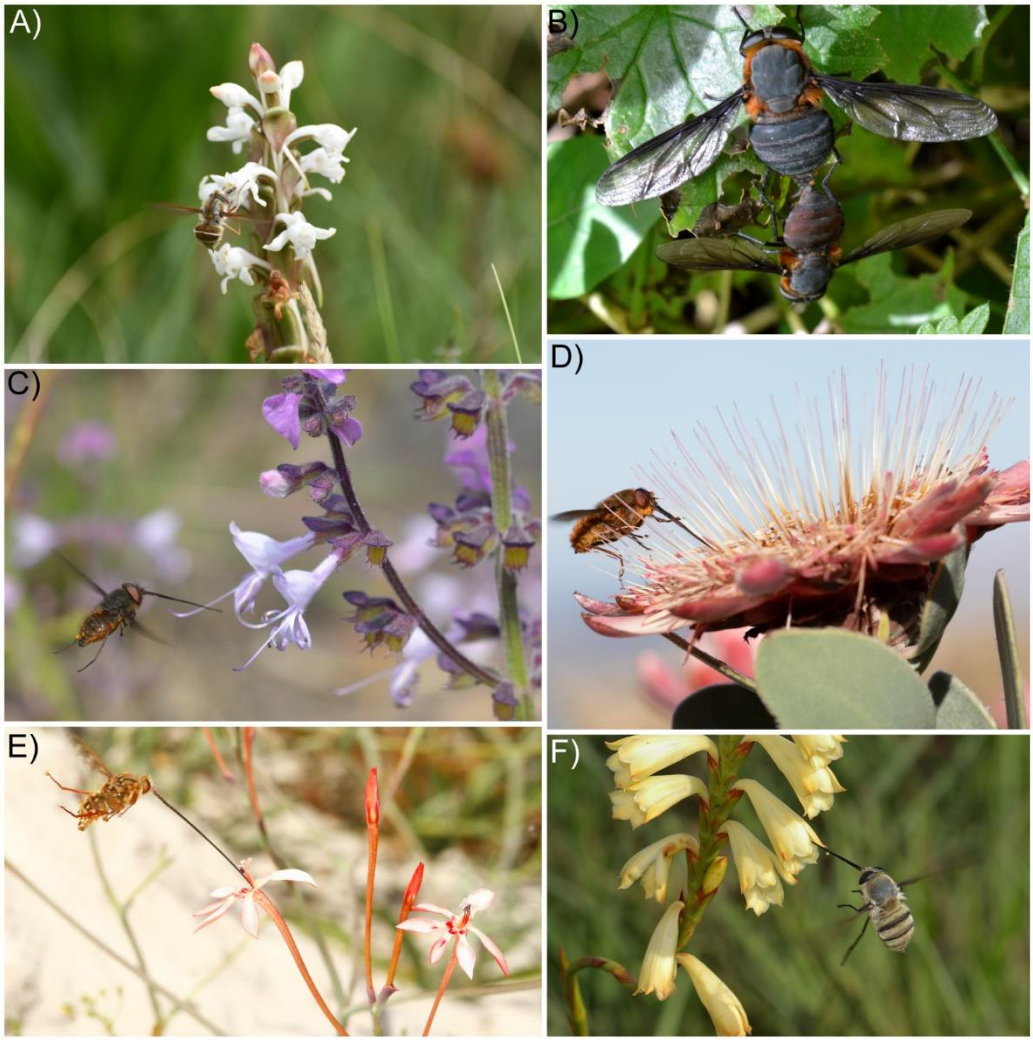


Figure 1.1 Images of *Moegistorhynchus* and *Prosoeca* species *in situ*. (A) *Prosoeca sp. 5* visiting *Satyrium longicauda* (B) Mating pair of *P. ignita* (C) *P. sp. 6* visiting a Lamiaceae sp. (D) *P. robusta* visiting *Protea punctata* (E) *Moegistorhynchus longirostris* visiting *Lapeirousia anceps* (F) *P. sp. 17* visiting *Watsonia watsoniodes.* Photo credits: (A) Miguel Castañeda-Zárate, (B) Jan-Hendrik Keet (C) Ruth Cozien (D & F) Steven Johnson (E) Bruce Anderson

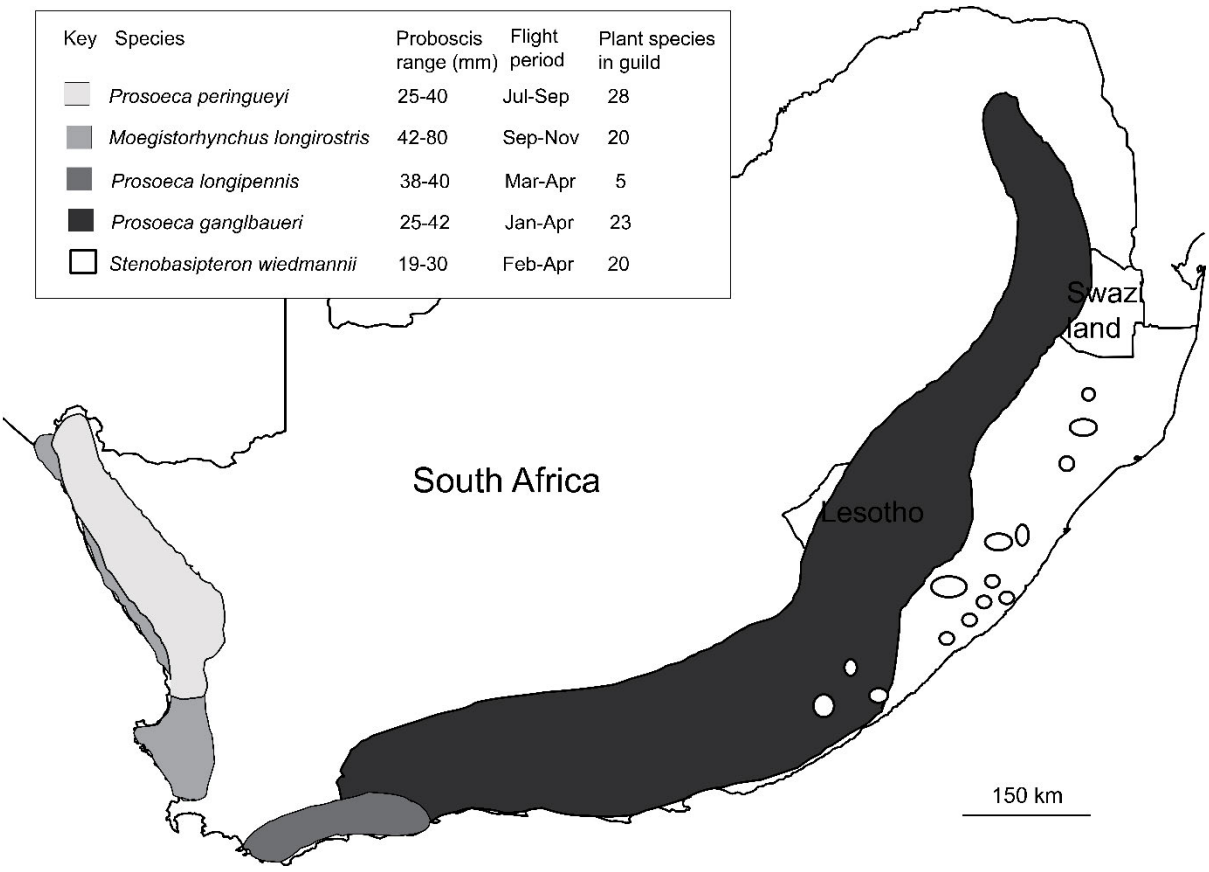


Figure 1.2 Edited map of long-tongued fly (Nemestrinidae) pollination guilds of southern Africa from Johnson (2006b).

Chapter 2

Key long-proboscid fly pollinator overlooked: morphological and molecular analyses reveal a new *Prosoeca* (Nemestrinidae) species

Genevieve L. Theron, Florent Grenier, Bruce C. Anderson, Allan G. Ellis, Steven D. Johnson, John M. Midgley, & Timotheüs van der Niet

This paper has been published in the Biological Journal of the Linnean Society (2020)

http://zoobank.org/urn:lsid:zoobank.org:pub:09F7102B-CB91-4105-908D-6B0A8643145B

ABSTRACT

Long proboscid nemestrinid flies are keystone pollinators of dozens of Southern African plants and consequently their taxonomic status may have important consequences for insect and plant conservation. We focus on *Prosoeca peringueyi*, considered to be a single, morphologically variable species upon which a guild of *ca.* 28 plants in the winter rainfall region depends for pollination. We quantified morphological variation and established whether it is associated with genetic variation within and among sites. Phylogenetic analyses of the mitochondrial COI gene revealed two well-supported clades. One clade contains long proboscid individuals that conform morphologically to the holotype of *P. peringueyi*. The sister clade contains individuals which frequently occur sympatrically with *P. peringueyi* and have shorter proboscides, as well as additional diagnostic characters which set it apart from *P. peringueyi*. A haplotype analysis based on nuclear ribosomal 28S DNA sequences of a subset of individuals corroborated these results. Based on our results we propose to recognize two species: *P. peringueyi* and *Prosoeca torquata* sp. nov., which is described here. Future research is required to quantify the interaction networks of these two fly species and the plant guilds they interact with, to facilitate conservation in the global biodiversity hotspot where they occur.

Keywords: cryptic species - DNA barcoding - keystone pollinator - morphological variation - Proboscis - Succulent Karoo - tangle-veined flies - taxonomy.

INTRODUCTION

Species delineation is a necessary starting point for asking and answering many evolutionary and ecological questions (Knowlton & Jackson 1994; Herre 2006). Certain species complexes can be difficult to accurately delimit morphologically, potentially resulting in erroneous classification as a single species (Bickford *et al.* 2007). When morphological distinctions are not evident, species may be discriminated by differences in other traits (Bickford *et al.* 2007) such as mating pheromones (Kozlov *et al.* 1996), calls (Henry 1994) and other ecological or behavioural differences. To establish species boundaries, it is often best to use an integrative taxonomic approach, considering multiple independent lines of evidence including DNA, morphology, ecological, geographic, and behavioural components (Mishler & Donoghue 1982).

Revision of species boundaries may affect our understanding of mutualisms and have associated consequences for conservation. For example, the presence of cryptic species (distinct species' classified under a single species name, see Bickford *et al*., 2007) affects the level of specialization of biotic interactions and may make interactions more generalised (eg. Molbo *et al.*, 2003) or more specialized (Eastwood *et al.* 2006; Smith *et al.* 2006) than previously thought. Specialization is thought to increase the risks of extinction (Bond 1994). Thus, differences in the levels of specialization or generalization also have important implications for the way we think about the conservation of mutualisms. For example, generalization may buffer plant species against the loss of particular pollinator species (Pauw, 2007).

In the hyper-diverse southern African flora, long proboscid flies are considered to be important pollinators of 100-200 species (Manning & Goldblatt, 2000; Newman *et al.*, 2014*;* Barraclough, 2017*)*. Allopatric populations of long proboscid fly pollinators often differ in

mouthpart morphology and these differences have frequently been implicated in the floral tube length divergence of plant populations (e.g. Anderson *et al.*, 2014; Pauw *et al*., 2009). Floral tube length divergence can in turn facilitate reproductive isolation among plant populations (Minnaar *et al.* 2019). Therefore, long proboscid flies are likely to be important drivers of ecological speciation in southern African plants (*sensu* Ellis *et al.* 2014; van der Niet *et al.* 2014; Minnaar *et al.* 2019). Most pollinating long proboscid fly species are members of the Nemestrinidae (Manning & Goldblatt 2000), and often a single fly species interacts with several plant species at a site (Anderson & Johnson, 2009; Pauw *et al.*, 2009; Newman *et al*., 2014). In contrast, many floral guild members are thought to be highly dependent on just a single nemestrinid species for pollination (Manning & Goldblatt 1996; Anderson & Johnson 2008; Newman *et al.* 2014), making these plant populations potentially vulnerable to local pollinator extinctions (Pauw 2007). Although ecologically and evolutionarily important as pollinators and agents of diversification, the Nemestrinidae of South Africa are poorly studied in terms of taxonomy and biology, and the group is thought to contain many undescribed species (Barraclough 2017). This opens up the possibility that updated fly taxonomy may affect our understanding of the specialization in plant-pollinator interactions.

In the Greater Cape Floristic Region (GCFR), and particularly the Succulent Karoo biodiversity hotspot, *Prosoeca peringueyi* Lichtwardt, 1920 is thought to be the sole pollinator of a guild of at least 28 plant species throughout the majority of the guild members' range (Manning & Goldblatt 1996). In addition to *P. peringueyi*, Manning and Goldblatt (1996) also recognised the presence of an undescribed nemestrinid species (now *P. marinusi* Barraclough, 2018) in one particular geographically restricted area, which is associated with the same floral guild. Within *P. peringueyi*, Manning and Goldblatt (1996) recorded substantial morphological variability (20-45mm) in the proboscis length, but this

morphological variation has yet to be studied in a taxonomic context. Our own preliminary field observations revealed that proboscis length variation did not only occur among populations, but also varied within populations. This, in combination with the need for taxonomic revision of the Nemestrinidae in South Africa (Barraclough 2017), suggests the potential existence of cryptic species. To evaluate the possible existence of cryptic species, we sampled individuals from across the range and first quantified proboscis length variation within and between populations. We then used DNA sequences from the mitochondrial and nuclear genomes to identify discrete genetic lineages within *P. peringueyi s.l.* and to test the utility of proboscis length, body size, and other morphological features for the separation of Operational Taxonomic Units (OTU's). Based on our results, we describe a new species in the *P. peringueyi* species complex.

METHODS

Taxon sampling and modality analysis

To quantify the variation in proboscis length, and to inform the selection of specimens for DNA sampling, individuals of both *Prosoeca peringueyi s. l.* sexes were collected and measured between July and October 2016 from across the known South African range. A single museum specimen from outside of this sampled range (from Namibia) was seen after the completion of field sampling. Proboscis length was measured with a pair of digital callipers from the junction of the proboscis and the face to the tip of the proboscis, without extending the proboscis. Measurements were analysed from two different data sets. To determine if proboscis length has a multimodal distribution across the species range, we used measurements taken across the whole range (range-wide, n = 553). To determine if proboscis length has a multimodal distribution within populations, we used measurements from two

extensively sampled populations (within population analysis: Kosiesberg, n = 196; Naries, n = 58). To test for multimodality, we implemented Hartigans' dip test using the diptest package (Maechler 2004). All statistics were run in R unless stated otherwise (R Core Team 2017).

DNA extraction

To infer whether samples with distinct proboscis lengths, both within and among populations, could be distinguished genetically we implemented a phylogenetic approach. Sixty-three individuals, for which proboscis length measurements were available and which represent the morphological variation across the range and within sites, were selected from 13 sites (Table S2.1). To infer phylogenetic relationships, we used sequences of the mitochondrial Cytochrome Oxidase 1 (COI) gene region, part of which constitutes the universal barcoding region (Hebert *et al.* 2003a). Monophyly of *P. peringueyi* was tested by adding four other *Prosoeca* species to the analysis (Table S2.1). To root the tree, we used the COI sequence of a *Trichophthalma* (Nemestrinidae) species from GenBank (Table S2.1). To corroborate results based on mitochondrial DNA, we also analysed DNA sequences from the 28S nuclear ribosomal gene for a subset of 10 individuals. These individuals were sampled from across the range of the putative species complex. Three additional *Prosoeca* species and a *Trichophthalma* species were also included in the analysis (Table S2.1).

Total DNA was extracted from a single hind leg of each individual using the Qiagen DNeasy blood and tissue kit (Qiagen, supplied by Whitehead Scientific, Cape Town, South Africa). We followed the manufacturer's protocol, but incubated samples for 48 hours to ensure complete chitin breakdown, and used only 50 µl of elution buffer to increase the final DNA concentration.

PCR amplification and DNA sequencing

The entire COI gene region was either amplified as a single amplicon using custom primer NE (Table S2) and TL2-N-3014 (Simon *et al.* 1994), or as two separate amplicons using 1) primers NE and C1-N-2191 (Simon *et al.* 1994) and 2) C1-J-2183 (Simon *et al.* 1994) and TL2-N-3014 (Table S2). PCR conditions were identical, irrespective of whether one or two separate regions were amplified. Each 30 µl reaction contained 3µl DNA Template, 3µl 10× Super-Therm PCR buffer (Super-Therm JMR-801; Separation Scientific SA (Pty) Ltd, Cape Town, South Africa), 3µl $MgCl_2$ (25mM), 0.6µl of Taq (Super-Therm JMR-801; Separation Scientific SA (Pty) Ltd, Cape Town, South Africa), 3µl of each primer (0.05mM), 0.6µl of BSA (10mg/ml), 0.6µl of dNTP (10mM) (AB gene; supplied by LCT Tech, South Africa) and 13.2µl MilliQ H_2O. Amplifications were performed using a Veriti PCR Thermal Cycler (Applied Biosystems, Carlsbad, CA, USA). The cycling protocol consisted of an initial denaturation step at 95ºC for 5 mins, followed by 45 cycles of denaturation at 94ºC for 1 min, annealing at 48ºC for 30 s and extension at 72ºC for 1 min, and a final extension at 72ºC for 30 mins before holding at 10ºC. The 28S gene region was amplified using the 2F and 3DR primers (Reemer & Ståhls 2013) (Table S2). Each 30 µl reaction contained 3µl DNA Template, 3µl 10× Super-Therm PCR buffer (Super-Therm JMR-801; Separation Scientific SA (Pty) Ltd, Cape Town, South Africa), 1.8µl $MgCl_2$ (25mM), 0.6µl of Taq (Super-Therm JMR-801; Separation Scientific SA (Pty) Ltd, Cape Town, South Africa), 3µl of each primer (0.05mM), 0.6µl of BSA (10mg/ml), 0.6µl of dNTP (10mM) (AB gene; supplied by LCT Tech, South Africa) and 14.4µl MilliQ H2O. The cycling protocol consisted of an initial denaturation step at 95ºC for 2 min, followed by 30 cycles of denaturation at 94ºC for 1 min, annealing at 49ºC for 30 s and extension at 72ºC for 2 min, and a final extension at 72ºC for 7 mins and then held at 10ºC. PCR products were sent to the Central Analytical Facility at Stellenbosch University (South Africa) to be purified and sequenced.

Phylogenetic analysis

The retrieved COI and 28S sequences were edited and aligned manually using BioEdit 7.2 (Hall 1999) (see Table S2.1 for genbank accession numbers). The final COI alignment consisted of 1384 base pairs (bp), whereas the 28S alignment consisted of 572 bp. The resulting COI sequence matrix was used to reconstruct a phylogenetic tree using Bayesian Inference in MrBayes 3.2.7 (Ronquist *et al.* 2012). JModelTest 2.1.10 (Darriba *et al.* 2015) was used to identify the best substitution model for each codon position separately. The first two codons could be combined and run with the GTR + G substitution model, whereas the third codon position was partitioned separately with the HKY + G substitution model. Two independent Markov chains were run for 10 million generations, sampling a tree every 10000 generations. The first ten percent of sampled trees was discarded as burn-in before constructing a 50% majority-rule consensus tree. Mean percentage sequence divergence within and between the clades derived from the consensus tree was calculated with MEGA 10.1 (Kumar *et al.* 2018) using p-distances. The 28S sequence matrix was used to reconstruct a haplotype network in POPART (Leigh & Bryant 2015) using median joining.

Morphological analyses

To assess whether the OTUs identified in the genetic analyses differ in morphological features, the 63 individuals of the *P. peringueyi* species complex included in the phylogenetic analysis were examined using a Leica M4 microscope. Body length was measured from the frons to the end of the abdomen, excluding the genitalia. A Students unpaired t-test was used to compare proboscis length and body length respectively, between members of the two main clades retrieved in the phylogenetic analysis. To test for an allometric relationship between body and proboscis length we used linear regression, whereas differences in the allometric

relationship between the OTU's identified were tested using an ANCOVA. Overall morphology of the individuals represented in the COI molecular phylogenetic analysis was also compared across samples and against the type specimen of *P. peringueyi*. Museum name abbreviations follow Evenhuis (2020).

RESULTS

The distribution of proboscis length across all sites was found to be bimodal ($D = 0.039$, $p < 0.001$). At both sites which were sampled extensively, the distribution of proboscis length was not unimodal (Kosiesberg: $D = 0.064$, $p < 0.001$; Naries: $D = 0.010$, $p < 0.001$) (Fig. 2.1).

The phylogenetic analysis recovered two well supported clades within *P. peringueyi* (Fig. 2a), each of which consisted of multiple geographically structured subclades (Fig. 2.2a, b). Within-clade sequence divergence in COI was relatively low (clade A: 5% and clade B: 2%) while between clade divergence was higher (11%). The 28S haplotype network similarly illustrates that there are two haplotypes separated by 4 mutations within the *P. peringueyi* species complex, which correspond to the two COI clades (Fig. S2.3).

Proboscis length and body length differed significantly between individuals from the two main clades (proboscis: $t = 14.87$, $p < 0.001$; body: $t = 6.55$, $p < 0.001$) with little overlap in proboscis length but extensive overlap in body length (Fig. 2.3). The mean proboscis length and body length of individuals from clade A was shorter than that of individuals from clade B (proboscis mean ± SE for clade A: 17.9 ± 0.37; clade B: 29.7 ± 0.66 and body mean ± SE for clade A: 16.6 ± 028; clade B: 19.1 ± 0.25). Body size is allometrically related to proboscis length of individuals within clades A and B ($F_{1,33} = 44.80$, $r^2 = 0.58$, $p < 0.001$ and $F_{1,27} =$

33.67, $r^2 = 0.54$, $p < 0.001$, respectively), but these allometric relationships differ between the two clades (overall model: $F_{3,60} = 161$, $r^2 = 0.89$, $p < 0.001$; interaction: $F_1 = 5.43$, $p = 0.023$).

The individuals of clade B correspond morphologically to the originally described *P. peringueyi* whereas individuals of clade A have unique characters that distinguish it from those in clade B (see diagnosis), which leads us to describe a new species below as *Prosoeca torquata* sp. nov. (Fig. 2.4 & Fig. 2.5).

Taxonomy

Genus *Prosoeca* Schiner

Prosoeca Schiner, 1867. Type species: *Nemestrina westermanni* Wiedemann, 1821.

***Prosoeca torquata* Theron sp. nov.**

http://zoobank.org/urn:lsid:zoobank.org:act:D4FA6366-B1FE-4F76-B973-6437E8DF7FAC

Type material: (in NMSA): South Africa: Northern Cape: HOLOTYPE: ♂: "South Africa, N. Cape, Steinkopf, Kosiesberg, -29.126317, 17.556865, 03 August 2019, F. Grenier, NMSA-DIP 166602".

Paratypes: (in NMSA) 1 ♀ "South Africa, N. Cape, Steinkopf, Kosiesberg, -29.126317, 17.556865, 03 August 2019, Timotheüs van der Niet". 1 ♂ 3 ♀ "South Africa, N. Cape, Steinkopf, Kosiesberg, -29.126317, 17.556865, 03 August 2019, F. Grenier NMSA-DIP 166602-166606." 3 ♂ "South Africa, N. Cape, Springbok, Naries, -29.690433, 17.664916, 04 August 2019, F. Grenier, NMSA-DIP 166607-166609". (in RMCA): 1 ♀ 1 ♂ "South Africa, Steinkopf, Kosiesberg, -29.126317, 17.556865, 03 August 2019, Timotheüs van der Niet, RMCA ENT 13300-13301".

Etymology: *torquata* (Latin) = being adorned with a neck collar; referring to the characteristic white band of hairs on the anterior margin of the thorax.

Diagnosis: small to medium sized (length 14 – 21mm), dark body with intricate patterns on the abdomen, dark brown legs, proboscis length 1.09 ± 0.09 times the length of the body (range of unextended proboscis length: 14 – 22mm), smokey brown infuscation on the anterior half of wing. *Prosoeca torquata* and *P. peringueyi* Lichtwardt, 1920 can be distinguished from all other species in the genus by the distinct wing patterning of a smoky brown anterior abruptly becoming hyaline on the posterior section (Fig. 2.4c, f). *Prosoeca torquata* differs most noticeably from *P. peringueyi* and *P. marinusi* Barraclough, 2018, both ecologically similar species distributed in the winter-rainfall region of South Africa, by the presence of a white band of pile on the anterior of the thorax and white pile on the face, which is largely lacking in the latter two species. Additionally, *P. torquata* has a darker thorax than *P. peringueyi*. *Prosoeca torquata* generally has a proboscis which is only slightly longer than the body compared to *P. marinusi* and *P. peringueyi*, which both have a proboscis substantially longer than the body. The frons of *P. torquata* is subtriangular, with the width above the antennal insertions 1.6 times the width below the anterior ocellus, whereas the frons is more rectangular with the width above the antennal insertions 1.2 and 1.3 times the width below the anterior ocellus in *P. peringueyi* and *P. marinusi* respectively (Table 2.1).

Geographical distribution: *Prosoeca torquata* occurs between Lüderitz in Namibia (single NMNW specimen) in the north and the southern border of the Kamiesberg region in the Northern Cape province at Uilklip in the south and is generally abundant when host plants are flowering. The northern most individual found in our own field sampling was at Eksteenfontein in the Northern Cape province and no individuals were found further north than the Richtersveld area despite extensive sampling. *Prosoeca peringueyi* occurs from Khuboes near the Namibian border in the north to Piketberg in the Western Cape province in

the south (per obs. F. Grenier) with a distribution gap across the knersvlakte. *Prosoeca marinusi* occurs in a limited area around Nieuwoudtville in the Northern Cape province.

Description (modelled on Barraclough *et al.* (2018)):

Body length: mean 16.6 mm; range 14–21 mm. Proboscis length: mean 18mm; range 14-22mm. Wing length: mean 15 mm; range 12-18mm.

Head: Colour generally dark brown to black. Face sublaterally irregular yellow-brown; slightly bulbous in profile; pruinescence silver to brown, but largely absent from medial section of face. Sublateral region of face covered by dense, elongate and mostly white pile. Frons subtriangular in shape; width anterior to ocellar tubercle less than maximum width of tubercle, width above antennal insertions 1.6 times the width below anterior ocellus, slightly to moderately swollen between antennal insertions and anterior ocellus, swelling receding strongly towards eye margin. Pruinescence on frons relatively dense, silver to brown. Pile on frons largely absent, only along eye margins if present. Gena with pile forming the beard being elongate, profuse and a mixture of black, white and yellow. Ocellar tubercle somewhat bulbous and well developed; ocellus just evident above upper eye margin in profile; anterior ocellus separated from posterior ocelli by shallow transverse groove. Ocellar triangle with generally long black pile, with longer black pile on the posterior margin; occiput with dense silvery pruinescence. Antenna with scape 1.2–1.5 times the length of pedicel; first flagellomere longer than the length of scape and pedicel combined; third flagellomere 1.5–2 times the length of second flagellomere; style subequal to scape + pedicel + flagellomere 1. Colour generally dark brown to black, ventral surface of scape irregularly yellow-brown, style darker than remainder of antenna. Scape, pedicel and flagellomere 1 with irregular silver to brown pruinescence; scape and pedicel with mixture of elongate and short, black and white pile. Proboscis 0.9–1.2 times the length of the body; black with dorsal part of basal

third brown. Palpus with first segment significantly longer than that of second segment and with much longer pile, second segment much narrower than first segment, colour generally dark brown to black.

Thorax: Scutum generally grey and brown; median and paired sublateral vittae absent or indistinct; vittae wider anteriorly, narrowing towards the scutellum; paired sublateral vittae distinctly black, ending at transverse suture if present. Pile of scutum sparse, of intermediate length and mostly black, longer posteriorly toward scutellum and laterally on scutum; postalar callus with longer, black pile on posterior region; ventral side of postalar callus with a tuft of white pile. Scutellum with distinct central diamond-shaped, or circular mark; posterior margin appearing to form a strikingly dark, black border. Anterior margin covered by silver to brown pruinescence; pile on disc of scutellum sparse, relatively long and black. Pile along posterior margin of scutellum elongate, of intermediate density and a mixture of black and paler white or yellow. Pleuron mostly blackish with silver pruinescence. Pile generally a mixture of black and white to yellow; over much of pleuron pile moderately long and of intermediate density. Pile most dense and elongate in two tufts, ventral and anterior to the base of the wing and between postalar callus and posterior spiracle; tuft of pile ventral to wing base pure white, pile absent over area below anterior spiracle and sometimes on anterior part of anepisternum; pile absent between mid and hind coxae and posterior spiracle. **Legs:** Coxae yellow brown to black with elongate, dense, white pile. Trochanters mostly blackish with some yellow-brown colouring; pile short, very sparse and white. Femora yellow-brown with a black mark ventrally on the hind femora; with a mixture of short dense and long sparse, mostly black pile. Tibia yellow-brown to brown; pile sparse, hind tibia with long sparse pile. Tarsi mostly dark brown to black, hind tarsi tend to be darker. **Wings:** Shape relatively slender; broadest just basal to insertion of *CuP* on posterior margin; costal margin close to straight, without distinct anteriorly curved flexure. *Sc* inserted on *C* approximately

coincident to the insertion of M_4 on the posterior margin of the wing; R_1 inserted closer to R_{2+3} than to *Sc*; insertions of *Sc* and R_1 well separated. Short appendix just beyond fork on R_{4+5} sometimes present; R_4 gently bowing upward; R_5 slightly bowing upward; M_1 and M_2 gently bowing upward; cell *cua* open at margin. Dark marking on R_1 positioned just basal to humeral cross vein; membrane pale smoky brown infuscated, appearing darker on anterior half to one-third of wing; posterior region of wing mostly hyaline; isolated darker patches present in hyaline region; distinction between brown infuscation and hyaline membrane sharply delineated. Tuft of pile at base of wing white. Halteres with light brown to yellow stalk; bulb dark brown.

Abdomen: Colour generally grey, with silvery pruinescence. Pile generally sparse, long and short intermixed and mostly black on tergites. T1 with pile relatively profuse; T2 with pile along anterior margins dense, elongate, white to yellow; pile over anterior corner of lateral margin of T2 dense, elongate, mostly white to yellow and directed anteriorly; pile along posteriolateral margins on T2 to T4 dense, elongate and black; T5 with pile along lateral margins more evenly distributed than that of T2 to T4, elongate and black. Median dark pruinescent vitta distinct, extending from medial section of posterior margin of T1 to terminalia, usually not covering full length of each tergite; paired sublateral markings of indistinct shape on T2 to T4. Membrane between T1 and T2 grey to brown; T2 with posterior margin stout and relatively broad; abdomen tapering abruptly after T3. Sternites typically paler than tergites with reddish brown colouring. Pile on sternites short, sparse and a mixture of black and white; S1 to S3 intermixed with long, white pile; membrane adjacent to lateral margins of S2 to S4 typically with profuse, elongate and white pile.

Male Genitalia: Hypandrium triangular in shape; 1.8 times longer than basal width; short, sparse, vestiture present on the apical 1/3. Gonocoxite rounded apically; apical half parallel sided, slightly broader basally; long, laterally projecting vestiture present on the

lateral half of the apical half. Gonostylus with concave outer edge, more or less parallel to inner edge and globular apical section. Phallus narrowing apically, more or less triangular in shape.

Other material examined

P. peringueyi types: (in SAM): 1 ♀, Namaqualand, Klipfontein, Cape Colony, August 1890. R. M. Lightfoot, SAM-DIP A009009. 1 ♂ O'okiep, Namaqua Div, R. Lightfoot, IX. 90, SAM-DIP-A009013. Non-type material: (in NMSA): 1 Van Rhynsdorp, 1927-7-28, Brauns, H. NMSA-Dip 049943. 1 Clanwilliam dist. Pakhuis Pass, 1964-10-17, Stuckenberg, B.R.; Stuckenberg, P. NMSA-Dip 052845. 1 Pakhuis Mts Pakhuis Farm 2mi. NNE.,1972-9-14, Irwin, M.E.; Irwin, B.J. NMSA-Dip 054873. 1 Gifberge, Van Rhynsdorp, 1911-9-1, C.C. NMSA-Dip 055007. (in NMNW): *P. torquata* 1 ♀ Namuskluft 88, Luderitz, SE 2710 Dd 20-22 Sept 1973, H14602.

Biology

Prosoeca torquata Theron sp. nov. is abundant in the winter-rainfall region of the Northern Cape province of South Africa, particularly at the start of the local flowering season. The peak activity of *Prosoeca torquata* is dependent on rainfall and elevation, but individuals are generally on the wing from early August to mid-October, when the diversity of host plants is most abundant. The peak activity period is slightly later than that of *P. peringueyi* which ranges from mid-July to mid-September. As with most *Prosoeca* species, these two species often feed while hovering over a host flower, probing many flowers in a single foraging bout (Fig. 2.6). Flies are active throughout the day but the peak activity is during the morning, after it has warmed up sufficiently and in the afternoon, before it cools down for the evening.

Based on examination of photographs and data presented in the preprint of Pauw *et al.* (2020), we note that the important *Pelargonium incrassatum* (Geraniaceae) pollinator, referred to informally as *Prosoeca* '*namaquensis*', corresponds to our concept of *P. torquata.*

DISCUSSION

This study combines genetic and morphological data from the *Prosoeca peringueyi* species complex from 13 sites across the range to show that two clearly distinct species exist in this complex. These two species are genetically distinct with COI sequence divergence between clades being 2.34 times larger than the average divergence within clades. The genetic divergence between these two *Prosoeca* taxa is 11%. Genetic divergence between species from other dipteran families ranges between 0-16% (Tabanidae: 0.0-16.2% (Changbunjong *et al.* 2018), Syrphidae: 4.08-15.28% (Mengual *et al.,* 2008), Muscidae: 0.77-11.33% (Renaud *et al.,* 2012)). Consequently, the observed levels of divergence between the two major clades identified here is comparable to the upper levels of divergence among fly species in general. The separation of these two clades into different species is further supported by the nuclear haplotype network as each OTU comprises a single, unique, haplotype. Members of the two clades also have distinct morphological features, consistent with the genetic differentiation found. Proboscis length differed between the two species even within sites where they co-occur. On the other hand, substantial overlap in traits like body length means that this feature is not diagnostic for the two species. The sympatric occurrence of the two species without intermediate individuals suggests the presence of a strong gene flow barrier/s, consistent with their recognition as separate species according to the biological species concept (Yoder *et al.,* 2002; Coyne & Orr, 2004; Blair *et al.,* 2005). There is also anecdotal evidence for assortative mating where they occur sympatrically based on at least 20 observations of mating pairs over multiple years (pers. obs. F. Grenier).

Previous taxonomic accounts of *P. peringueyi* were based on very limited material, which may explain why the obvious presence of two species, even within sites, has been hitherto overlooked. The original description of *P. peringueyi* by Lichtwardt (1920) featured four specimens all from the Northern Cape, whereas Bezzi (1924) redescribed *P. peringueyi* based on one of these specimens. Both emphasised the cross vein between R_4 and R_{2+3} to distinguish it from all other *Prosoeca* species. With more extensive sampling we found that this trait is not stable across the range for *P. peringueyi* but is always absent in *P. torquata*. *Prosoeca torquata* and *P. peringueyi* can be separated by several reliable characters in fresh specimens (Table 2.1). Many of these characters, such as hair densities and body colouration may become less reliable as living individuals age or if preserved specimens are mishandled. Traits such as the white pile on the anterior of the thorax and the overall darker colouring are often easy to use in the field as well as in flight for identification. Further investigation of morphological variation may ultimately form the basis for further description of taxa within the *P. peringueyi* species complex.

The recognition of similar taxa that occur sympatrically begs the question what mechanism underpins this. The theory of competitive exclusion suggests that when two species compete similarly for the same limited resource, one is likely to go locally extinct (Hardin 1960). Alternatively, co-existence may be possible when species partition the available resources. The co-existence of two closely related fly species may be facilitated by differences in proboscis length. Proboscis length affects the species of plants from which flies can forage from (Klumpers *et al*., 2019); therefore, divergence in proboscis length might facilitate adequate partitioning of nectar resources for co-existence to occur (Pauw, 2013; Maglianesi *et al*., 2015). Although long proboscid insects can potentially forage from long and short tubed flowers (Haber & Frankie 1989; Martins & Johnson 2013), they often display preferences for long-tubed flowers which generally contain greater nectar rewards (Martins &

Johnson 2013; Anderson *et al.* 2016; Klumpers *et al.* 2019). In contrast, foragers with shorter proboscides are often unable to access nectar from flowers with deeper tubes (Pauw *et al*., 2009) and are consequently restricted to the subset of available flowers with short tubes (Ranta & Lundberg, 1980; Haber & Frankie, 1989; Geerts & Pauw, 2009; Martins & Johnson, 2013; Klumpers *et al*., 2019 but see Merxem *et al*., 2009). Co-existence may however, also be facilitated by the use of divergent larval hosts (Berlocher & Feder 2002). Although South African Nemestrinidae are assumed to be parasitic as larvae, almost nothing is known about their larval biology (Barraclough 2006, 2017).

This study points to the co-occurrence of two, rather than one long proboscid fly species within the *Prosoeca peringueyi* complex in the arid biodiversity hotspot of the Succulent Karoo. The ability to identify these two species has implications for the conservation and management of the large floral guild which relies exclusively on long proboscid flies for seed production. In particular, the loss of a single pollinator species in even very generalised systems has been shown to have negative effects on plant reproduction (Brosi & Briggs 2013) and this loss is much more severe in highly specialized systems (Pauw 2007). Being able to utilize sympatric fly species that have partial foraging niche overlap is likely to give many guild members more resilience to stochastic pollinator failure (Pauw 2007), a growing concern as climates become more unpredictable. Future study of the functional specialisation of the plants in this guild will provide insights on the potential resilience of the plant species and the mechanism facilitating co-existence of the fly species.

ACKNOWLEDGMENTS

We thank Célian Bimbard, Laure Schneider-Maunoury, Marion Poiré, Matt Jasper, Martin Plancke and Guilhem Duvot for assistance in the field. We thank Dr. David Barraclough, Dr.

Jason Londt and an anonymous reviewer for their helpful comments on this manuscript. Additionally, we thank Cape Nature (CN44-31-2588), Northern Cape Province (FAUNA 1230) for permits. This study was funded by National Research Foundation of South Africa – Foundational Biodiversity Programme (NRF-FBIP grant # 110440 to A.G.E), National Research Foundation of South Africa (SARChI grant # 46372 to S.D.J) and the Erasmus+ international credit mobility grant to F.G.

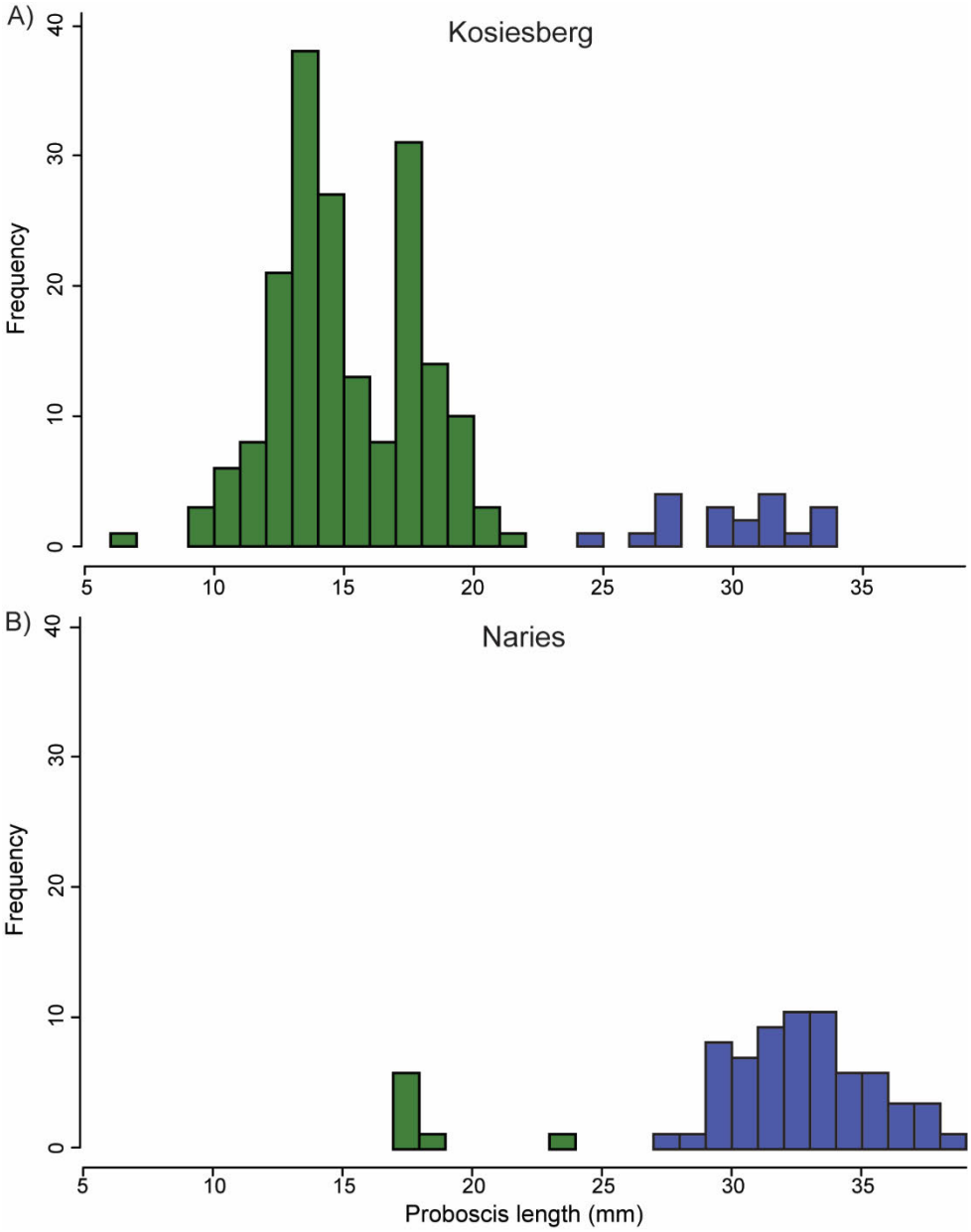


Figure 2.1 The frequency of unextended proboscis lengths (in millimetres) measured in the Kosiesberg population (A) and the Naries population (B). Colour-coded bars refer to the operational taxonomic units designated in the genetic analysis: *Prosoeca torquata* (green) and *Prosoeca peringueyi* (blue).

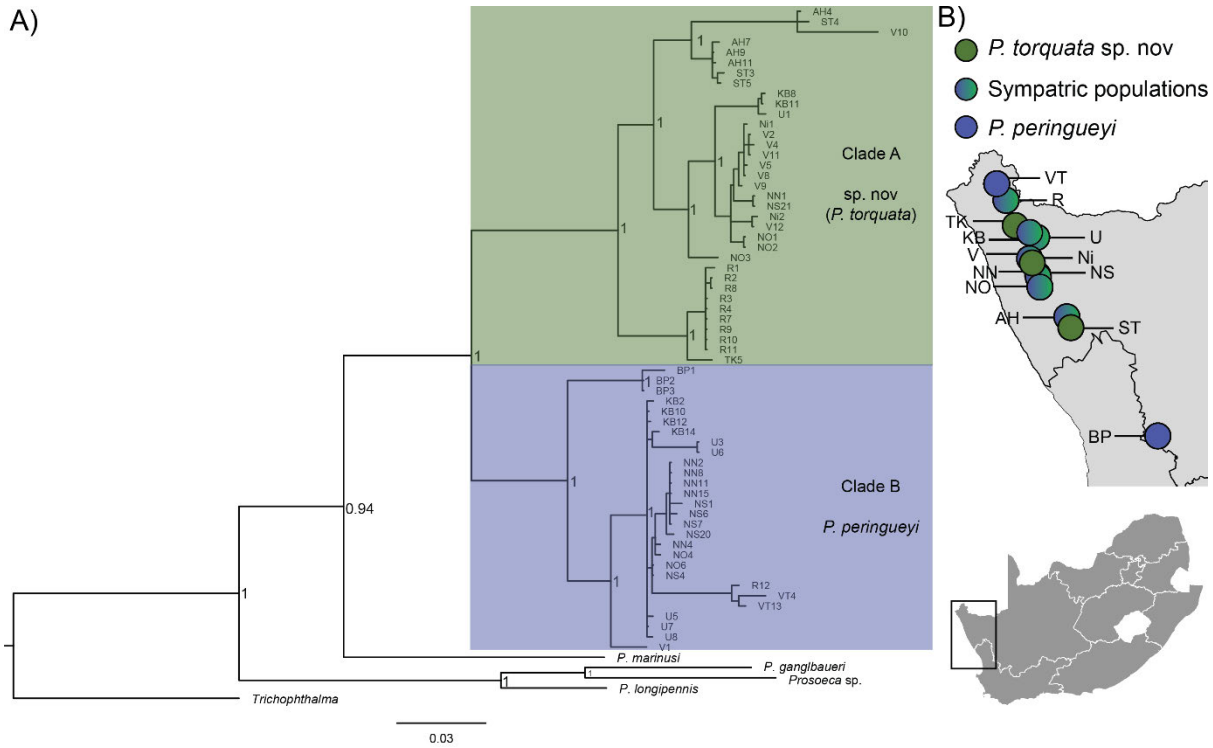


Figure 2.2 Phylogenetic relationships among accessions of the *Prosoeca peringueyi* complex sampled across its range. (A) Bayesian inference majority rule consensus tree of COI sequences for the *P. peringueyi* complex. Posterior probabilities are shown for major nodes. Species names associated with clades represent the new taxonomy proposed in the present study. (B) distribution of allopatric populations of *P. peringueyi* (blue) and *Prosoeca torquata* (green) and of sympatric populations (gradient of blue and green) that were sampled for this study (circles). Details of tip labels and population codes are provided in the Supporting Information (Table S2.1)

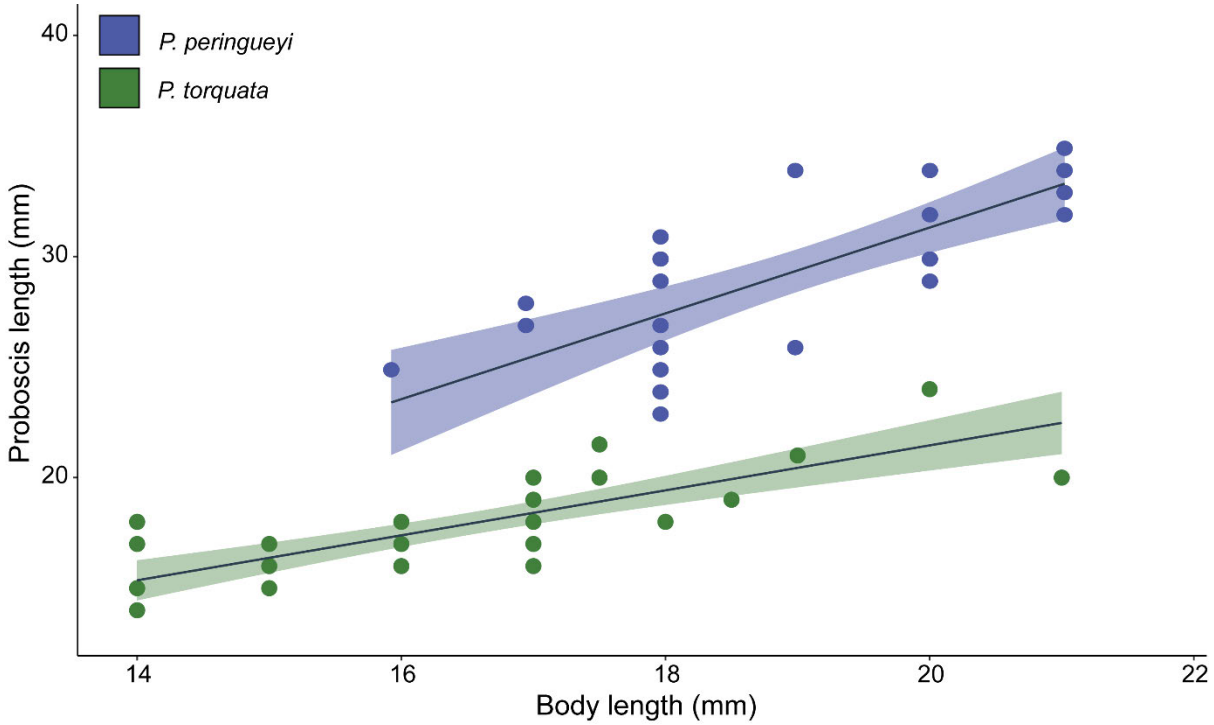


Figure 2.3 Relationship between body size (in millimetres) and proboscis length (in millimetres) for *Prosoeca peringueyi* (blue) and *Prosoeca torquata* (green) ($F_{1,27} = 33.67$, $r^2 = 0.54$, $P < 0.001$ and $F_{1,33} = 44.80$, $r^2 = 0.58$, $P < 0.001$, respectively; overall model, $F_{3,60} = 161$, $r^2 = 0.89$, $P < 0.001$; and interaction, $F_1 = 5.43$, $P = 0.023$). The shaded area around the trend line indicates the 95% confidence interval.

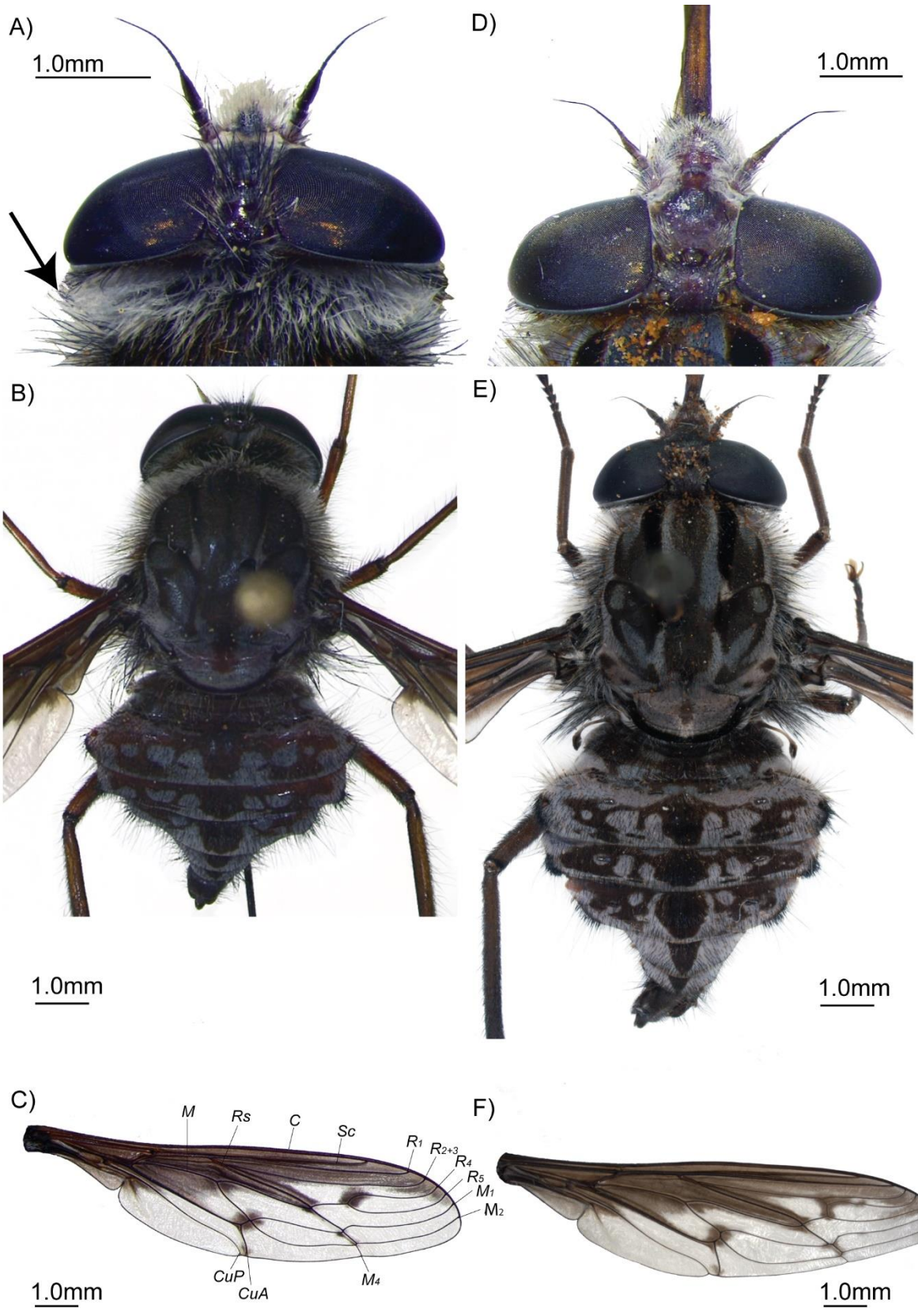

A)
1.0mm
D)
1.0mm
B)
1.0mm
E)
1.0mm
C)
M
Rs
C
Sc
R1
R2+3
R4
R5
M1
M2
CuP
CuA
M4
1.0mm
F)
1.0mm

Figure 2.4 Photographs of the head, dorsal view of the body and the wing of *Prosoeca torquata* (A–C) and *Prosoeca peringueyi* (D–F). Arrow in A indicates the band of white pile after which P. torquata is named. Orange particles visible in D are pollen grains. Abbreviations: *C*, costal vein; *CuA*, anterior branch of cubital vein; *CuP*, posterior branch of cubital vein; *M*, medial vein; M_1, first branch of media; M_2, second branch of media; M_4, fourth branch of media; R_1, anterior branch of radius; R_{2+3}, second branch of radius; R_4, upper branch of third branch of radius; R_5, lower branch of third branch of radius; *Rs*, radial sector; *Sc*, subcostal vein.

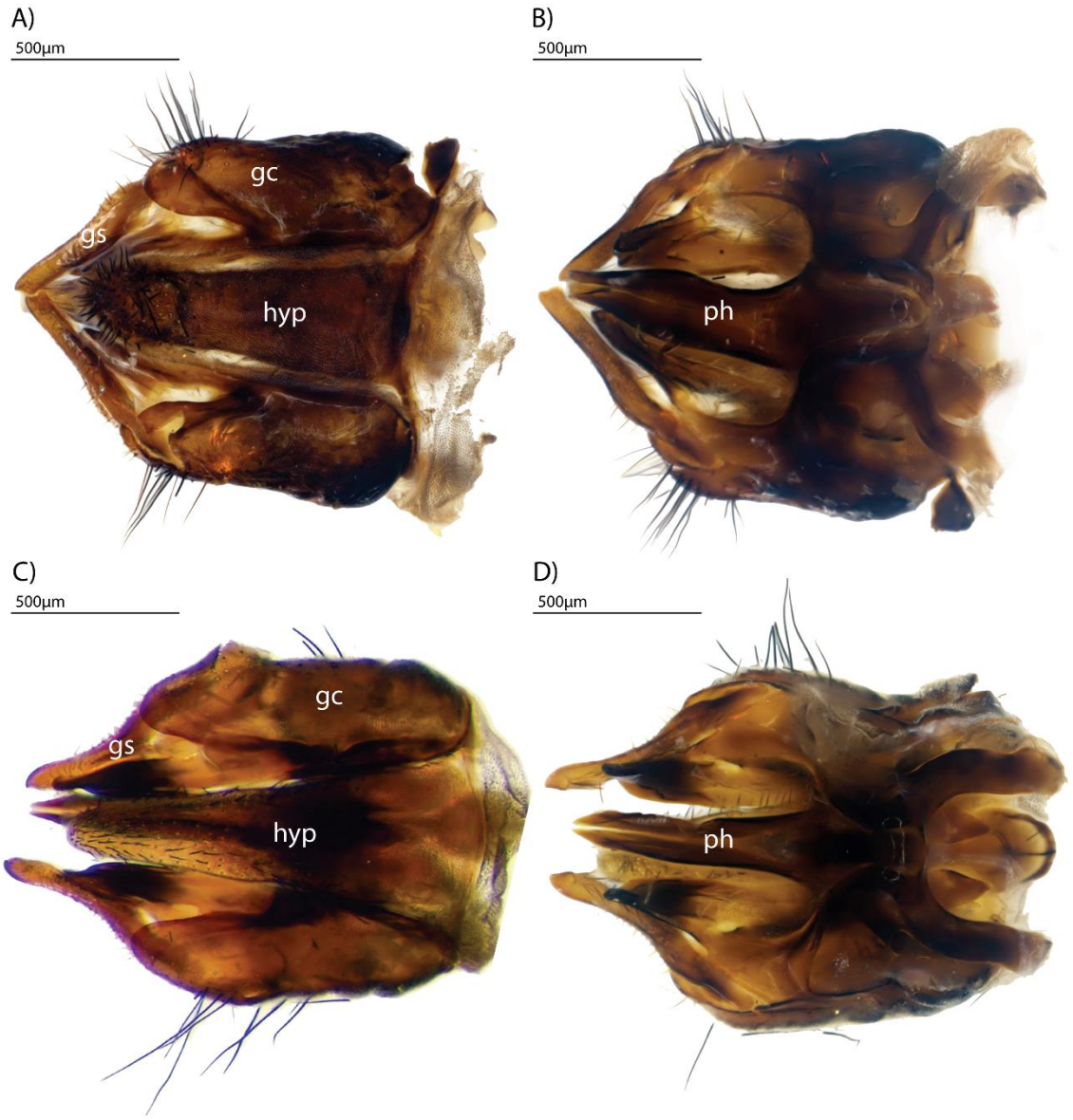


Figure 2.5 Photographs of the male genitalia of *Prosoeca peringueyi* (A, B) and *Prosoeca torquata* (C, D) in dorsal (A, C) and ventral view (B, D). Abbreviations: gc, gonocoxite; gs, gonostylus; hyp, hypandrium; ph, phallus.

Figure 2.6 Photographs of *in situ* adult *Prosoeca peringueyi* visiting *Zaluzianskya* sp. (A) and visiting *Lapeirousia silenoides* (B), and *Prosoeca torquata* hovering over *Lapeirousia dolomitica* (C) and mating (D). Photo credits: (A-B) Steven Johnson, (C-D) Florent Grenier.

Table 2.1 Summary of consistent morphological traits to distinguish *Prosoeca torquata* sp. nov from *P. peringueyi* and *P. marinusi*.

	P. peringueyi	*P. torquata sp. nov*	*P. marinusi*
Proboscis length relative to body length	Long	Subequal	Long
White pile on anterior of thorax	Largely absent	Present	Absent
Paired dark sublateral vittae on thorax	Dark and distinct	Light or absent	Absent
Width above antennal insertions relative to width below anterior ocellus	1.2	1.6	1.3
Pile on hind tibia	No long pile	Sparse long pile	Sparse long pile

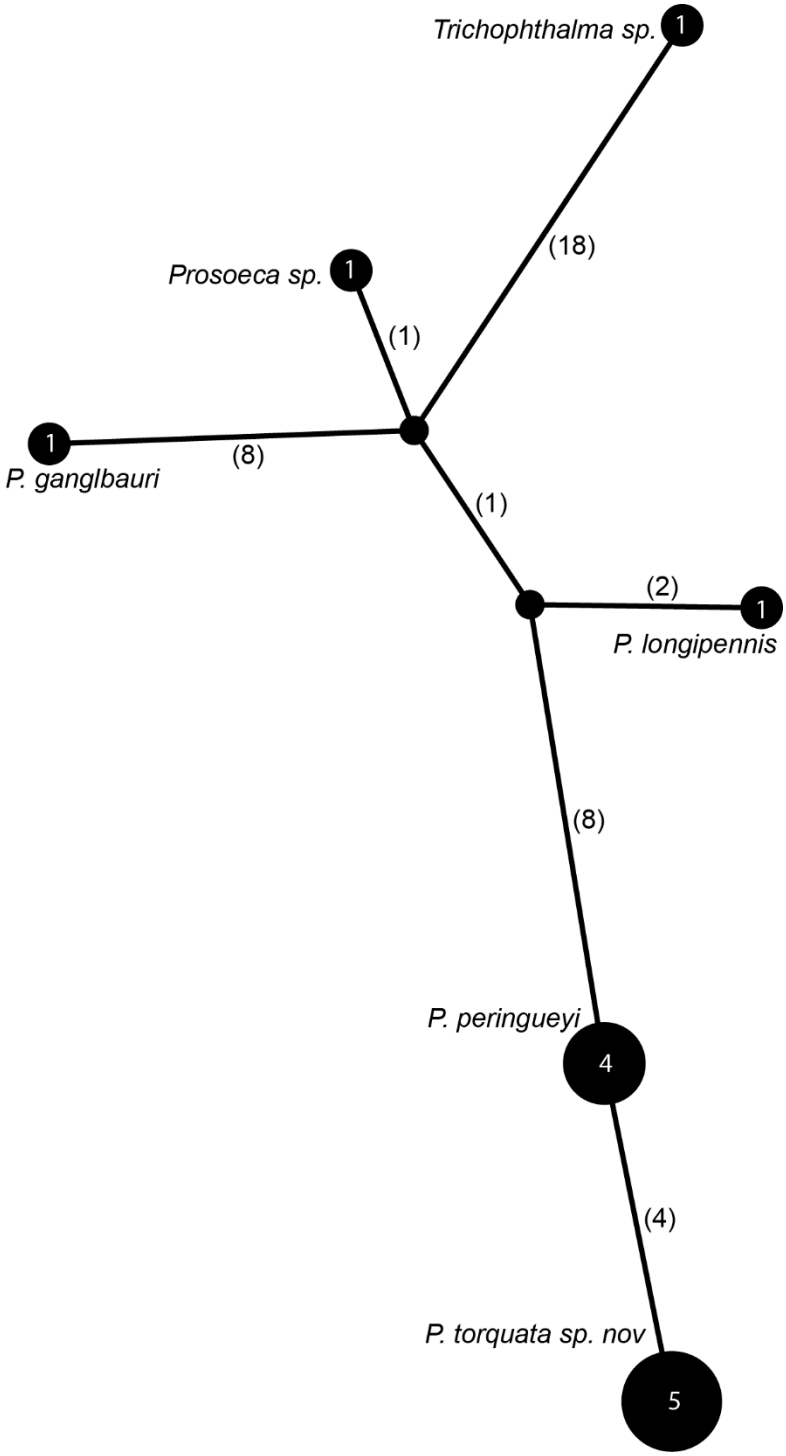


Figure S2.1 28S haplotype network of ten individuals from the *Prosoeca peringueyi* complex and four Nemestrinidae species. Numbers in parentheses correspond to the hypothetical number of haplotypes that exist between sampled haplotypes. The size of circles and numbers within them indicate the number of individuals sampled for a particular haplotype.

Table S2.1 Table of localities and specimen details of accessions used for the phylogenetic analyses.

Locality	Latitude	Longitude	Specimen reference	Species	Body length (mm)	Proboscis length (mm)	GenBank accession numbers	
							CO1	28S
Ariehoek	-30.244758	18.050258	AH4	*P. torquata*	17	16	MT487492	MT436414
			AH7	*P. torquata*	14	18	MT487493	
			AH9	*P. torquata*	16	17	MT487494	
			AH11	*P. torquata*	15	17	MT487495	
Botterkloof	-31.823601	19.256734	BP1	*P. peringueyi*	19	26	MT487496	MT436410
			BP2	*P. peringueyi*	18	26	MT487497	
			BP3	*P. peringueyi*	20	29	MT487498	
Umdaus	-29.189137	17.649852	U1	*P. torquata*	17	16	MT487538	MT436413
			U3	*P. peringueyi*	18	27	MT487539	
			U5	*P. peringueyi*	18	31	MT487540	
			U6	*P. peringueyi*	17	27	MT487541	
			U7	*P. torquata*	16	25	MT487542	
			U8	*P. peringueyi*	21	35	MT487543	
Kosiesberg	-29.126317	17.556865	KB2	*P. peringueyi*	18	19	MT487499	
			KB8	*P. torquata*	18	18	MT487500	
			KB10	*P. peringueyi*	20	30	MT487501	
			KB11	*P. torquata*	18,5	25	MT487502	

			KB12	*P. peringueyi*	21	32	MT487503	
			KB14	*P. peringueyi*	21	34	MT487504	
Naries North	-29.690433	17.664916	NN1	*P. torquata*	16	18	MT487505	
			NN2	*P. peringueyi*	20	34	MT487506	
			NN4	*P. peringueyi*	18	30	MT487507	
			NN8	*P. peringueyi*	20	34	MT487508	MT436411
			NN11	*P. peringueyi*	20	30	MT487509	
			NN15	*P. peringueyi*	17	28	MT487510	
			NN20	*P. peringueyi*	20	34	MT487522	
Naries South	-29.716007	17.671079	NS1	*P. torquata*	18	24	MT487518	
			NS4	*P. peringueyi*	21	34	MT487519	
			NS6	*P. peringueyi*	19	34	MT487520	
			NS7	*P. peringueyi*	21	33	MT487521	
			NS21	*P. torquata*	17	18	MT487523	
Nigramoep	-29.529318	17.591411	Ni1	*P. torquata*	15	15	MT487505	MT436417
			Ni2	*P. torquata*	17	18	MT487506	
Nooitgedacht	-29.849447	17.693732	NO1	*P. torquata*	15	15	MT487513	MT436415
			NO2	*P. torquata*	17	17	MT487514	
			NO3	*P. torquata*	15	15	MT487515	
			NO4	*P. peringueyi*	20	34	MT487516	
			NO6	*P. peringueyi*	20	32	MT487517	
Stinkfonteinberg	-28.698448	17.243396	R1	*P. torquata*	15	16	MT487524	
			R2	*P. torquata*	14	15	MT487525	
			R3	*P. torquata*	14	15	MT487526	

			R4	*P. torquata*	14	14	MT487527	
			R7	*P. torquata*	16	16	MT487528	
			R8	*P. torquata*	17	20	MT487529	
			R9	*P. torquata*	19	21	MT487530	MT436416
			R10	*P. torquata*	17	20	MT487531	
			R11	*P. torquata*	17	18	MT487532	
			R12	*P. peringueyi*	20	29	MT487533	
Studers plateau	-30.386984	18.103343	ST3	*P. torquata*	17	20	MT487534	MT436417
			ST4	*P. torquata*	17,5	20	MT487535	
			ST5	*P. torquata*	19	21	MT487536	
Tierkloof	-29.033988	17.366095	TK5	*P. torquata*	17,5	21,5	MT487537	
Vaalheuwel	-29.473022	17.556174	V1	*P. peringueyi*	18	29	MT487544	MT436412
			V2	*P. torquata*	20	24	MT487545	
			V4	*P. torquata*	21	20	MT487546	
			V5	*P. torquata*	14	17	MT487547	
			V8	*P. torquata*	16	18	MT487548	
			V9	*P. torquata*	16	18	MT487549	
			V10	*P. torquata*	17	18	MT487550	
			V11	*P. torquata*	17	19	MT487551	
			V12	*P. torquata*	17	20	MT487552	
Vandersterberg	-28.480114	17.11973	VT4	*P. peringueyi*	19	26	MT487553	
			VT13	*P. peringueyi*	18	25	MT487554	
Nieuwoudtville	-31.3979	19.1406	NV304	*P. marinusi*			MT487555	
Riversdale	-34.036519	21.455702	AH20	*P. longipennis*			MT487556	MT436421
Mokhotlong	-29.29287	29.03628	AH66	*P. ganglbauri*			MT487557	MT436422

Long Tom Pass	-25.095679	30.562498	PP283	*Prosoeca* sp.		MT487558	MT436420
				Trichophthalma sp.		DQ631994	AY144383

Table S2.2 Primer sequences used to amplify the gene regions included in this study.

Primer name	Gene region	Reference	Primer sequence
NE	COI	this study	ACT TTA TAY TTT ATY TTT GGA GC
C1-N-2191	COI	Simon *et al*. (1994)	CCC GGT AAA ATT AAA ATA TAA ACTTC
C1-J-2183	COI	Simon *et al*. (1994)	CAA CAT TTA TTT TGA TTT TTT GG
TL2-N-3014	COI	Simon *et al*. (1994)	TCC ATT GCA CTA ATC TGC CAT ATT A
2F	28S	Reemer *et al.* (2013)	AGA GAG AGT TCA AGA GTA CGT G
3DR	28S	Reemer *et al.* (2013)	TAG TTC ACC ATC TTT CGG GTC

Chapter 3

You don’t know the half of it: molecular phylogenetics reveals dramatic underestimation of diversity in a key pollinator group (Nemestrinidae)

Genevieve L. Theron, Bruce C. Anderson, Ruth Cozien, Allan G. Ellis, Florent Grenier, Steven D. Johnson, Ethan Newman, Anton Pauw, & Timotheüs van der Niet

To be submitted to: Systematic Entomology

ABSTRACT

Nemestrinidae are important pollinators of numerous highly specialized plant species and have radiated extensively in southern Africa. Despite their known ecological importance in southern Africa, Nemestrinidae have been largely neglected in collections and the systematics of the group is relatively poorly resolved. In this study we aimed to assess the phylogenetic relationships and species diversity among three southern African nemestrinid genera from the Nemestrininae subfamily: *Prosoeca*, *Moegistorhynchus* and *Stenobasipteron*, with a focus on *Prosoeca*. We reconstructed a molecular phylogeny using both mitochondrial and nuclear DNA sequence data. The topology from the combined analysis designates a monophyletic *Moegistorhynchus* sister to a paraphyletic *Prosoeca* with *Stenobasipteron* nested inside it, although there was some topological incongruence in the placement of clades among the four gene regions sampled. Almost half of the sampled species diversity was found to be undescribed in all three genera. An analysis of phylogenetic diversity shows that undescribed species make a large contribution to the overall phylogenetic diversity within the sampled group. We find a higher occurrence of sampled diversity from biodiversity hotspots and the Fynbos and Grassland biomes, although further sampling is required to evaluate whether this represents real disparities in the distribution of diversity. The presence of numerous undescribed species, as well as the paraphyletic nature of *Prosoeca*, clearly illustrates the need for increased taxonomic and sampling effort in this ecologically important group of flies.

Keywords: Diptera - Nemestrinidae - phylogeny - southern Africa - species diversity - tangle-vein flies

INTRODUCTION

Despite centuries of taxonomic effort, new species are still being described at a high rate worldwide, even in plants (Christenhusz & Byng 2016) and vertebrates (Tapley *et al.* 2018) which are relatively well-studied compared to invertebrates. The taxonomic impediment is of particular concern for invertebrates, which often have high proportions of undescribed diversity and few specialists to address the description deficit (New 1999). Recently, mass DNA barcoding studies have greatly contributed towards improved numerical estimates of global insect diversity. For large hyper-diverse orders of insects like Coleoptera, Diptera, Hymenoptera, and Lepidoptera, one estimate suggests that only 20-28% of species have been described worldwide (García-Robledo *et al.* 2020) and it is thought that microfauna are generally responsible for the majority of this description deficit (Hebert *et al.* 2016; Forbes *et al.* 2018).

Species are frequently used as units in downstream analyses. Poorly resolved species boundaries therefore hamper ecological and evolutionary studies (Tobias *et al.* 2010). In addition to unclear species boundaries, excessive taxonomic splitting or lumping of species names may be equally problematic for ecological and evolutionary inferences and conservation decisions (Zachos 2018). A phylogenetic approach provides an opportunity for objective assessments of species relationships, boundaries, and higher-level systematics. To aid species taxonomy, single-locus and multilocus species delimitation methods that implement a phylogenetic perspective have become popular tools (Fujita *et al.* 2012; Fontaneto *et al.* 2015). Similarly, phylogenetic diversity, a measure calculated by summing the of branch lengths of relevant taxa in a phylogenetic tree, has been suggested to represent an objective measure of biodiversity, superior to simple species counts (Faith 1992; Miller *et al.* 2018). A resolved, densely sampled phylogeny can form the foundation for systematic,

diversity and quantitative research as well as practical decision making (Freckleton *et al.* 2002; Johnson *et al.* 2005).

Biogeographic region designations such as biomes and biodiversity hotspots have aided practical decision-making in conservation. Global biodiversity hotspots are areas of high plant endemism, and are under considerable threat of extinction (Myers *et al.* 2000). Biomes, such as those defined by Rutherford *et al.* (2006), on the other hand, are based on distinct floristic community composition and climatic variables. As both hotspots and biomes are generally modelled on plant biodiversity in Africa (Rutherford *et al.* 2006), little is known about the utility of these biogeographical designations for insects. Biodiversity hotspots encompass the distribution of threatened mammals to a reasonable degree (Schipper *et al.* 2008), apart from rodents (Amori *et al.* 2011). Few studies have, however, found substantial links between plant diversity and insect diversity (Hawkins & Porter 2003; Procheş *et al.* 2009; Kemp & Ellis 2017; Bosc *et al.* 2019) and these studies tend to focus almost exclusively on herbivores rather than pollinators.

Nemestrinidae is a small family of flower-foraging Diptera, comprised of *ca*. 277 species worldwide (Barraclough 2017). Nemestrinids are easily recognizable due to the unique diagonal wing veins that give the vein structure a "tangled" appearance. Although their appearance is remarkably variable, species belonging to the family are often relatively large with well-developed proboscides and profusely pubescent bodies (Barraclough 2017). Adult Nemestrinidae are thought to be obligate nectar feeders (Karolyi *et al.* 2012) and were likely one of the first insect families to act as pollinators since at least the late Triassic (Labandeira *et al.* 2007). All Nemestrinidae are thought to be important pollinators, although this has only been established for species in Chile and southern Africa (Manning & Goldblatt 2000; Devoto & Medan 2006). Southern Africa is one of the main hotspots for nemestrinid diversity, with representatives of two subfamilies, each with three genera. In the

Trichopsideinae subfamily *Nycterimyia* Lichtwardt 1909, *Trichopsidea* Westwood 1839 and *Atriadops* Wandolleck 1897, collectively have four described species found in southern Africa (Barraclough 2006). The subfamily Nemestrininae, with *Prosoeca* Schiner 1867, *Stenobasipteron* Lichtwardt 1910 and *Moegistorhynchus* Macquart 1840, has received the most attention as it contains important pollinators for numerous plant species (Barraclough 2006). The three endemic southern African Nemestrininae genera are unique in having some species with extremely long proboscides that allow them to feed on specialized, long-tubed plant species (Barraclough 2006). Nine well-studied long proboscid Nemestrinidae species are the only known pollinators for *ca.* 150 species of plants (Manning & Goldblatt 1996, 1997, 2000; Potgieter & Edwards 2005; Anderson & Johnson 2009; Newman *et al.* 2014) with many more specialized interactions suspected. Nemestrinids are also often the main pollinators of rare and endangered plants in southern Africa (Johnson 2006b).

The systematics of the higher Nemestrinidae has long-been debated, with the placement of the family in the Brachycera remaining unresolved. Wiegmann *et al.* (2011) suggested that the Nemestrinidae are sister to Xylophagidae while a recent study suggests that they are sister to Acroceridae (Shin *et al.* 2018). In contrast there has been little debate concerning the placement of *Prosoeca*, *Stenobasipteron* and *Moegistorhynchus* within the Nemestrininae. Bequaert (1932) and Greathead (1967) agreed on the placement of these three genera in the Nemestrininae based largely on the length of the proboscis, and details of male genitalia added by Greathead (1967). Bernardi (1973), in agreement with Berquart (1932) and Greathead (1967), then proposed the classification of Nemestrininae that is still currently used. Bernardi (1973) suggested that some of the genera within the Nemestrininae are not well separated and that *Prosoeca* is very closely related to *Stenobasipteron* as they are morphologically very similar. Barraclough (2006) agreed with Bernardi (1973) and additionally suggested that *Stenobasipteron* may be considered a subgenus of *Prosoeca*.

Moegistorhynchus has been considered closely related to *Nemestrinus* due to the similarity in wing venation and genitalia traits (Bernardi 1973). Within the southern African genera, the bulk of the described diversity is in *Prosoeca* with 37 species (Barraclough 2017; Barraclough *et al.* 2018; Theron *et al.* 2020). *Prosoeca* is widespread and found across large parts of southern Africa. On the other hand, *Moegistorhynchus* (four species) and *Stenobasipteron* (three species) have restricted ranges in the winter- and summer-rainfall regions of southern Africa, respectively.

In contrast to research effort dedicated to understanding the importance of southern African Nemestrininae for the ecology and evolution of flora, foundational data on their distribution and taxonomy are surprisingly rudimentary. A revision of the known species together with a key to the southern African nemestrinid species was published more than 90 years ago (Bezzi 1924), and only two new species descriptions have appeared in recent years (Barraclough *et al.* 2018; Theron *et al.* 2020). Recently, Barraclough (2006) transferred some species from *Stenobasipteron* to *Prosoeca* and revised the key to the nemestrinid genera for southern Africa. Many tentative new species have been identified as part of species complexes and noted in published literature (Barraclough 2006; Barraclough & Slotow 2010) while additional complexes and undescribed species have been discovered through detailed examination of large collections (pers. obs. Theron). Therefore, it is unsurprising that many collections potentially contain undescribed species (Barraclough 2006, 2017; Theron *et al.* 2020). One consequence of the presence of large numbers of undescribed species is that many ecological studies have only been able to identify relevant pollinators up to the genus level. Comparisons of pollinating fly species between studies have therefore been challenging (Manning & Goldblatt 2000; Goldblatt *et al.* 2004; Goldblatt & Manning 2007; Hansen *et al.* 2012; de Jager *et al.* 2016; Pauw *et al.* 2020). This is particularly serious given that the

taxonomy of *Prosoeca*, *Stenobasipteron* and *Moegistorhynchus* has relevance for their conservation and that of the plants that they pollinate.

The aim of this study is twofold. First, we use a phylogenetic framework to test the monophyly of *Prosoeca*, *Stenobasipteron* and *Moegistorhynchus* and to delineate the relationships among these southern African genera of the Nemestrininae subfamily. Secondly, we use the species-level phylogeny to assess patterns of sampled species diversity with specific reference to their distribution in southern African hotspots of biodiversity and biomes.

METHODS

Taxon sampling and morphospecies designation

Nemestrinidae were widely collected between January 2014 and February 2020 from the known southern African range, as established from historical records (Fig. 2.1), obtained for southern Africa from the Global Biodiversity Information Facility (GBIF; http://www.gbif.org/). Specimens were assigned to genera based on the key to genera of Barraclough (2017) and identified to species where possible using a combination of descriptions and comparison to type specimens. Specimens that did not match existing descriptions were assigned to distinct morphospecies. To add confidence to our morphological species identifications, multiple accessions were included in the phylogeny for morphologically variable or geographically widespread species where possible. After the initial morphospecies designation, a visual inspection of the phylogenetic placement of specimens from our phylogeny (see results) was used to revisit the initial species identification based on morphological characters. In some cases, the initial designation was then revised based on this integrated morphological and phylogenetic approach. Species were

designated with cf. when there was some degree of uncertainty about their identification (eg. collection locality was far removed from the known distribution, or the type material was unavailable for inspection). We use the term morphospecies throughout to indicate that our designations are mainly based on morphological differences. All subsequent analyses of diversity use this conservative integrated morphology delineation of morphospecies.

A total of 136 individuals representing 58 morphospecies of three genera were sampled; *Prosoeca* N = 49, *Moegistorhynchus* N = 4 and *Stenobasipteron* N = 5 (Table S3.1).

To assess whether our sampling is geographically representative of existing collections, we visually compared our sampling with historical collections/records of Nemestrinidae from southern Africa, Lesotho and eSwatini. All available records were extracted from the GBIF (http://www.gbif.org/) on 26 October 2020. Maps of records were generated in QGIS v3.10 (QGIS Development Team 2014) and plotted by biome as delineated by the South African National Biodiversity Institute (http://bgis.sanbi.org/SpatialDataset/Detail/669, 2011) and global biodiversity hotspots as delineated by Hoffman *et al.* (2016). The presence of morphospecies within individual biomes or hotspots was extracted from QGIS to allow for comparisons among areas.

Data for species delineation and reconstructing phylogenetic relationships were obtained from DNA sequences of the mitochondrial genome, including Cytochrome Oxidase I (COI) (which comprises the universal DNA barcoding region), and 16S ribosomal DNA (16S), and the nuclear genome, including 28S ribosomal DNA (28S) and the Carbamoyl-phosphate synthetase (CAD) gene region (Winterton *et al.* 2007). Outgroups were added to assess the phylogenetic placement of the ingroups: *Prosoeca*, *Stenobasipteron* and *Moegistorhynchus*, within the Nemestrininae subfamily. Representatives from two additional Nemestrininae genera, *Trichophthalma* and *Nemestrinus,* were added as single accessions from GenBank. Additionally, DNA sequences from a single *Hirmoneura* species, from the Hirmoneurinae

subfamily, was downloaded and added to the matrix. A *Terphis* species (Acroceridae) was used to root all gene trees apart from 28S. For 28S, the only suitable outgroup with a sequence in GenBank was *Trichophthalma* (Nemestrinidae), thus in the combined analysis accessions with missing data were generated for the other three outgroups for 28S.

DNA extraction

Total DNA was extracted from a single ethanol-preserved hind leg of each adult individual using the Qiagen DNeasy blood and tissue kit (Qiagen, supplied by Whitehead Scientific, Cape Town, South Africa). We followed the manufacturer's protocol, but incubated samples for 48 hours to ensure complete chitin breakdown and used only 50 µl of elution buffer to increase the final DNA concentration.

PCR amplification and DNA sequencing

The COI gene region was either amplified as a single amplicon using custom primer NE (Theron *et al.* 2020) and TL2-N-3014 (Simon *et al.* 1994), or as two separate amplicons using 1) primers NE and C1-N-2191 (Simon *et al.* 1994) and 2) C1-J-2183 (Simon *et al.* 1994) and TL2-N-3014 (Table S3.2). PCR conditions were identical, irrespective of whether one or two separate regions were amplified. Each 30 µl reaction contained 3µl of extracted DNA, 3µl 10× Super-Therm PCR buffer (Super-Therm JMR-801; Separation Scientific SA (Pty) Ltd, Cape Town, South Africa), 3µl of $MgCl_2$ (25mM), 0.6µl of Taq (Super-Therm JMR-801; Separation Scientific SA (Pty) Ltd, Cape Town, South Africa), 3µl of each primer (0.05mM), 0.6µl of BSA (10mg/ml), 0.6µl of dNTP (10mM) (AB gene; supplied by LCT Tech, South Africa) and 13.2µl of MilliQ H_2O. Amplifications were performed using a Veriti PCR Thermal Cycler (Applied Biosystems, Carlsbad, CA, USA). The cycling protocol consisted

of an initial denaturation step at 95ºC for 5 mins, followed by 45 cycles of denaturation at 94ºC for 1 min, annealing at 48ºC for 30 s and extension at 72ºC for 1 min, and a final extension at 72ºC for 30 mins before holding at 10ºC. The 16S gene region was amplified using the LR-J-12887 and SR-N-13398b primers (Simon *et al.* 1994) (Table S3.2) and the same reaction mixture as above. The cycling protocol consisted of an initial denaturation step at 94ºC for 4 min, followed by 35 cycles of denaturation at 94ºC for 1 min, annealing at 52ºC for 30 s and extension at 72ºC for 1 min, and a final extension at 72ºC for 4 mins and then held at 10ºC. The 28S gene region was amplified using the 2F and 3DR primers (Reemer & Ståhls 2013) (Table S3.2). Each 30 µl reaction contained 3µl DNA Template, 3µl 10× Super-Therm PCR buffer (Super-Therm JMR-801; Separation Scientific SA (Pty) Ltd, Cape Town, South Africa), 1.8µl $MgCl_2$ (25mM), 0.6µl of Taq (Super-Therm JMR-801; Separation Scientific SA (Pty) Ltd, Cape Town, South Africa), 3µl of each primer (0.05mM), 0.6µl of BSA (10mg/ml), 0.6µl of dNTP (10mM) (AB gene; supplied by LCT Tech, South Africa) and 14.4µl MilliQ H2O. The cycling protocol consisted of an initial denaturation step at 95ºC for 2 mins, followed by 30 cycles of denaturation at 94ºC for 1 min, annealing at 49ºC for 30 s and extension at 72ºC for 2 mins, and a final extension at 72ºC for 7 mins and then held at 10ºC. The CAD gene region was amplified using primers 54F and 405R (Moulton & Wiegmann 2004) or as two separate amplicons using primers 1) 54F and CAD-R611-631 (Cisneros 2015) and 2) CAD-F202-222 (Cisneros 2015) and 405R (Table S3.2). The same reaction mixture that was used for 28S was also used for both PCR of both CAD amplicons. The cycling protocol consisted of an initial denaturation step at 94ºC for 3 mins, followed by a touchdown protocol, starting with 5 cycles of denaturation at 94ºC for 30 secs, annealing at 57ºC for 30 secs and extension at 72ºC for 90 secs, a further 5 cycles of denaturation at 94ºC for 30 secs, annealing at 52ºC for 30 secs and extension at 72ºC for 90 secs, followed by a

final 35 cycles of denaturation at 94ºC for 30 secs, annealing at 45ºC for 30 secs and extension at 72ºC for 90 secs and then held at 10ºC.

PCR products were sent to the Central Analytical Facility at Stellenbosch University (Stellenbosch, South Africa) to be purified and sequenced in both directions.

Phylogenetic analyses

Forward and reverse sequences were edited using BioEdit 7.2 (Hall 1999) and aligned using ClustalW (Thompson *et al.* 1994). Minor manual edits were made to the ClustalW alignment. The final alignments consisted of 140 accessions of 1401 base pairs (bp) for COI, 140 accessions of 549 bp for 16S, 138 accessions of 590 bp for 28S and 95 accessions of 918 bp for CAD. The CAD matrix had the fewest accessions but included representatives from all the genera. Where CAD sequences were not available, accessions with missing data were generated for a combined analysis with all four gene regions. PartitionFinder (Lanfear *et al.* 2017) was used on the CIPRES Science Gateway v3.3 (Miller *et al.* 2010) to identify the best partitioning scheme and substitution model for each partition. Separate analyses were conducted for each gene region based on the optimal substitution model using Bayesian Inference in MrBayes 3.2.7 (Ronquist *et al.* 2012) on the CIPRES Science Gateway v3.3 (Miller *et al.* 2010). Two independent Markov chains were run for 10 million generations, sampling a tree and parameters every 10,000 generations. The first 25% percent of sampled trees and parameters were discarded as burn-in before constructing a 50% majority-rule consensus tree. Convergence and effective sample size (ESS) values of >100 were confirmed for all parameters using Tracer 1.7.1 (Rambaut *et al.* 2018). Topologies were visually compared for incongruence. All cases of incongruence were due to the topology of a single gene region deviating from the topologies of the remaining three regions, rather than

replicated patterns of incongruence between mitochondrial and nuclear partitions. Furthermore, incongruence was generally poorly supported for 28S and 16S. We therefore considered this incongruence to be an artifact of low resolution in individual gene trees and proceeded to concatenate all four matrices for a combined analysis. The concatenated matrix was used to reconstruct a phylogenetic tree with the same parameters as specified above. As suggested by the PartitionFinder analysis, each codon position for COI and all other gene regions were partitioned separately.

Diversity analyses

To objectively estimate species richness, independent of potential human bias due to taxonomic splitting or lumping, we implemented several methods of species delimitation based on the molecular dataset. For species delimitation we implemented two single-locus species delimitation methods using the COI dataset. The Automatic Barcode Gap Discovery (ABGD) (Puillandre *et al.* 2012a) method is a distance-based method relying on pairwise sequence distances between specimens to estimate the number of OTUs, whereas the Bayesian Poisson Tree Process (bPTP) (Zhang *et al.* 2013) delimitation method takes evolutionary history into account. To estimate the number of species present in the COI dataset, we used ABGD implementing the JC69 model of DNA sequence evolution and a relative gap =1, and bPTP using the COI genetree with default settings on the web versions (https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html and http://species.h-its.org/, respectively). A high prior intraspecific divergence value (P= 0.012) as suggested by (Puillandre *et al.* 2012a) was used to delineate the number of OTU's from the initial partition for ABGD.

As a large portion of the diversity sampled comprised putative undescribed morphospecies, we conducted phylogenetic diversity analyses to test whether undescribed diversity merely reflects the splitting of known taxa, or whether our sample comprises a set of unknown species distributed randomly across the phylogeny. To quantify the degree of phylogenetic diversity that is contributed by undescribed species, we estimated Faith's phylogenetic diversity (PD) for three different categories of morphospecies classification (all morphospecies, described + cf. morphospecies, only described morphospecies). If the addition of undescribed morphospecies constitutes splitting of described species, we predict that a relatively small addition of PD would occur. However, if undescribed morphospecies represent a random sample of species diversity, then a larger addition of PD can be expected when including them in the analysis. The drop.tip function from the Ape package (Paradis *et al.* 2004) in R (R Core Team 2017) was used to create a tree with a single accession per morphospecies. In most cases a species only had a single accession with data from all four genes present, which was then retained in the tree. However, if multiple accessions with all four genes were available for a species, included accessions were chosen arbitrarily. Phylogenetic diversity was then calculated for three datasets; one with all morphospecies present (whole tree), one with only the described species included (described only), and lastly one with the described species and those designated as cf. (described + cf.), using the pd function in the *picante* package (Kembel *et al.* 2010). The latter analysis accommodates the fact that not all collected specimens could be unambiguously identified as a described or undescribed species. We treated cf. species in two different ways. We either included them as described species (representing a conservative estimate for the number of undescribed species) or as undescribed morphospecies (representing a liberal estimate for the number of undescribed species). To illustrate that an addition of PD by adding undescribed species does not represent the addition of a single long branch (which would be the case if undescribed

species represent a phylogenetically biased sample), we also plotted the frequency distributions of the terminal branch lengths from the three trees for which PD was calculated (a single accession of all the morphospecies, described + cf. and described only). A median terminal branch length was calculated for each of the three trees for comparative purposes.

RESULTS

The combined topology renders *Prosoeca* paraphyletic with respect to *Stenobasipteron*, while *Moegistorhynchus* (clade A) was recovered as monophyletic and sister to *Prosoeca* + *Stenobasipteron* (Fig. 3.2). The paraphyletic *Prosoeca* lineage is resolved into three well supported major clades (clade B, C, D) (Fig. 3.2). The type species for *Prosoeca* (*P. westermanni*) is placed in clade B (*Prosoeca s.s.* clade), the smallest of three main *Prosoeca* clades (PP = 1). The node subtending clade C and *Prosoeca* sp.15 has low support but clade C itself (PP = 1) and the two subclades are well-supported (PP = 1). *Prosoeca* sp.15 is sister to clade D, although this is relatively poorly supported (PP= 0.88). Clade D and its three subclades are all well supported (PP = 1) with the type specimen for *Stenobasipteron* in clade D_1 (PP = 1). *Stenobasipteron weidemanni* along with the multiple accessions of undescribed *Stenobasipteron* morphospecies form a monophyletic clade. In the combined tree *Trichophthalma* and *Hirmoneura* were placed as sister to the southern African Nemestrininae with *Nemestrinus* as sister to them. Thus, based on single representatives of the non-southern African genera, *Hirmoneura* appears to render Nemestrininae paraphyletic.

Individual gene trees (Fig. S3.1-S3.4) varied in the degree of support and topology. The gene trees based on ribosomal loci were characterized by particularly low support and resolution, but nonetheless individually supported many of the same higher-level relationships as the combined tree (see below). On the other hand, COI and CAD showed some incongruence in terms of the monophyly of the different genera. The *Stenobasipteron* clade (clade E1) is

monophyletic for three gene regions but the placement of *P. marinusi* and *P. sp.1* as sister to *S. weidemanni* creates a paraphyletic clade in the COI tree (Fig. S3.1). The placement of *Stenobasipteron* within *Prosoeca* is consistent among all four gene regions, whereas the monophyly of clade B is supported by three gene regions (COI, 16S and 28S). The placement of *Moegistorhynchus* as monophyletic clade sister to *Prosoeca* is supported by 16S and 28S but not all of the *Moegistorhynchus* species are sister to *Prosoeca* in COI and CAD.

We distinguished 58 distinct morphospecies with an integrated morphology approach, twenty-three (40%) of which are currently recognised species. An additional six morphospecies (10%) were assigned cf. status, whereas the remaining 29 (50%) represent undescribed species. Using single-locus delineation, the total number of OTUs varied, depending on the model used to partition the COI data (bPTP or ABGD). ABGD recovered the same number of OTUs identified by integrated morphology (58 species), with 67% of groupings of specimens being consistent with those designated by morphology (Fig. 3.2). bPTP recognised 76 OTUs with 65% of groupings of specimens overlapping with those designated by morphology (Fig. 3.2).

The tree with only a single accession of each described morphospecies had the lowest phylogenetic diversity (total PD:1.71 units). The addition of cf. and of undescribed morphospecies increased total PD considerably (total PD: 2.00 units and 3.43 units, respectively). The distribution of terminal branch lengths illustrates that the addition of PD due to adding undescribed morphospecies is not a result of splitting or the addition of a single long branch (Fig. S3.5b). The median terminal branch length of the whole tree (0.041) was only slightly shorter than for the tree with described + cf morphospecies (0.056) and the tree containing only described morphospecies (0.056). The sampling of described and undescribed morphospecies is more or less evenly distributed across the tree with undescribed morphospecies present in all clades (Fig. S3.5a).

The tree shows some geographic structure, with many clades or subclades being restricted to either the western (winter-rainfall) or eastern (summer-rainfall) region. Clade A contains two described and two undescribed *Moegistorhynchus* morphospecies that are all restricted to the western part of southern Africa. Similarly, clade B contains one described and two undescribed morphospecies from the west of the country. The 14 described and six undescribed morphospecies in clade C, and its two subclades, are a mixture of morphospecies that occur in either the winter-rainfall or summer-rainfall part of the country. Clade D_1 contains the only South African *Stenobasipteron* species as well as four undescribed morphospecies which conform to the genus description. This clade is restricted to the eastern part of South Africa with all the undescribed morphospecies being collected in the north eastern part of southern Africa. Clade D_2 contains a mixture of morphospecies that have either an eastern or western distribution with 11 undescribed and eight described morphospecies. Lastly, clade D_3 is a clade that occurs within the Succulent Karoo hotspot of biodiversity with an even split of three described and three undescribed morphospecies.

Samples of nemestrinid species were obtained from seven of the nine biomes of southern Africa, with only the Desert and Nama Karoo biomes having no species sampled for this study. The Fynbos and Grassland biomes had the most morphospecies sampled while Albany Thicket only had a single morphospecies sampled (Fig. 3.3b). The Fynbos (39%) and Grassland (68%) biomes both had a high percentage of their diversity represented by undescribed morphospecies while all three morphospecies sampled in the Savanna biome were undescribed (Fig. 3.3b).

Overall, there were far more morphospecies sampled in the three hotspots than outside these areas (Fig. 3.3a). Undescribed morphospecies, however, represented 76% of diversity sampled outside of biodiversity hotspots compared to 42% within hotspots.

DISCUSSION

The topology of the combined analysis shows 1) that *Prosoeca* is not monophyletic because accessions representing *Stenobasipteron* form a monophyletic clade nested inside *Prosoeca,* and 2) that *Moegistorhynchus* is sister to the combination of *Prosoeca* + *Stenobasipteron*. Our results suggest that we sampled somewhere between 58-76 species based on morphospecies and OTUs. Even if we follow our most conservative estimate of diversity (i.e. 58 species), this represents a substantial increase in diversity from the 44 species currently described for these three genera. Importantly, of the diversity sampled in this study, only 23 of the morphospecies could be confirmed as being previously described, with another six designated as cf. implying that we were unable to confirm whether they had been previously described. This suggests that, conservatively, only ~50% of the diversity found within these three genera has been described and that total nemistrinid diversity in southern Africa may eventually add up to about 90 species. The paraphyletic nature of *Prosoeca* in addition to the abundance of undescribed morphospecies confirms that the taxonomy of Nemestrinidae as it currently stands is inadequate and requires thorough attention.

The independent species delimitation methods based on molecular data (i.e. AGBD and bPTP) suggested diversity estimates ranging from 58-76 species. This indicates that our estimate of 58 species based on morphology and phylogeny is at the conservative end of the spectrum and is unlikely to be a gross overestimate of diversity (as might be a result of a taxon splitting approach). Additionally, known complexes, such as the *P. ganglbaueri* species complex (Barraclough 2017) which contains multiple genetic lineages from the COI species delimitation analyses (Fig. 3.2), were left as a single species with multiple accessions until further morphological examination can be used to delimit this group. In addition to high richness of undescribed morphospecies in the tree, the undescribed diversity represents a substantial portion of the PD examined. Thus, even given some uncertainty in identification,

there can be no doubt that a substantial portion of the Nemestrinidae diversity of southern Africa remains undescribed. The current state of taxonomy and the number of undescribed species in *Prosoeca*, *Stenobasipteron* and *Moegistorhynchus* means that it is currently difficult to know with certainty how many species were not sampled for the present analyses.

High degrees of undescribed diversity in invertebrates are most often found in groups with small, indistinct or rare species (Jones *et al.* 2009; Hebert *et al.* 2016; Brown *et al.* 2018). However many nemestrinids are charismatic, large-bodied flies, with distinct features and are also ecologically important as pollinators for a diverse array of plant species (Bernardi 1973). Consequently, it is surprising that so much of their diversity has been taxonomically neglected. For most of the twentieth century the group did not attract the attention of specialist museum-based entomological taxonomists (perhaps because of the lack of direct economic/veterinary importance of the group in comparison with other dipteran groups, such as the Tabanidae), and it is only recently that some entomologists have begun to pay close attention to the group (Barraclough 2006; Barraclough *et al.* 2018; Theron *et al.* 2020). Despite their ecological importance, many southern African Nemestrininae remain challenging to identify because of a lack of up-to-date identification resources and keys (Barraclough 2017).

A high percentage of the total nemestrinid diversity was sampled in biodiversity hotspots, especially in the Grassland and Fynbos biomes. On the other hand, far fewer morphospecies were found in the Indian Ocean Coastal Belt, Albany Thicket and Savanna biomes. The distribution of South African Tabanidae shows a similar pattern with high Grassland and Fynbos biome diversity, but Nemestrinidae are far less diverse in Savannas than tabanids are (Snyman *et al.* 2020). The high nemestrinid diversity found within biodiversity hotspots may represent an association with high plant diversity in these regions, as a result of their dependence on floral nectar. Further sampling is required to exclude the possibility that

sampling bias is responsible for the present pattern of diversity in hotspots compared to other areas (Reddy & Dávalos 2003). This similarly applies to a potential sampling bias in the Grassland and Fynbos biomes.

This study makes a contribution to the estimation of species diversity, placement and delineation of nemestrinid genera and species. Furthermore, it clearly illustrates the need for a major revision of all three genera of southern African Nemestrininae. It is, however, clear that this analysis represents a starting point for molecular systematic and taxonomic work to follow, based on the need for broader species sampling of this ecologically important group. In particular, an integrated study of morphological and molecular characters is required to definitively resolve the higher taxonomic placement and delineation of species and genera. Nevertheless, the phylogenetic framework established for this clade of nemestrinids provides a useful template to reconstruct the evolution of their diverse and unique morphology, and historical biogeography and to aid species delineation.

ACKNOWLEDGMENTS

We thank Nina Hobbhahn, Saskia Klumpers, Miguel Castañeda-Zárate, Matthew Rule, Terrence Bellingham and Lawrence Harder for assistance in collecting specimens and in the field. We would also like to thank Dr. David Barraclough for showing us many of the type specimens. Additionally, we thank Cape Nature (CN44-31-2588), Northern Cape Province (FAUNA 1230), Limpopo Department of Economic Development, Environmental & Tourism (ZA/LP/86341), Mpumalanga Tourism and Parks Agency (MPB 5612), KZN wildlife (OP 1614/2018 & OP 492/2020), Eastern Cape Department of Economic Development, Environmental & Tourism (CRO37/19CR) and South African National PARKS (THEG1553) for permits. This study was funded by National Research Foundation

of South Africa – Foundational Biodiversity Programme (NRF-FBIP grant # 110440 to A.G.E), National Research Foundation of South Africa (SARChI grant # 46372 to S.D.J) and the Erasmus+ international credit mobility grant to F.G.

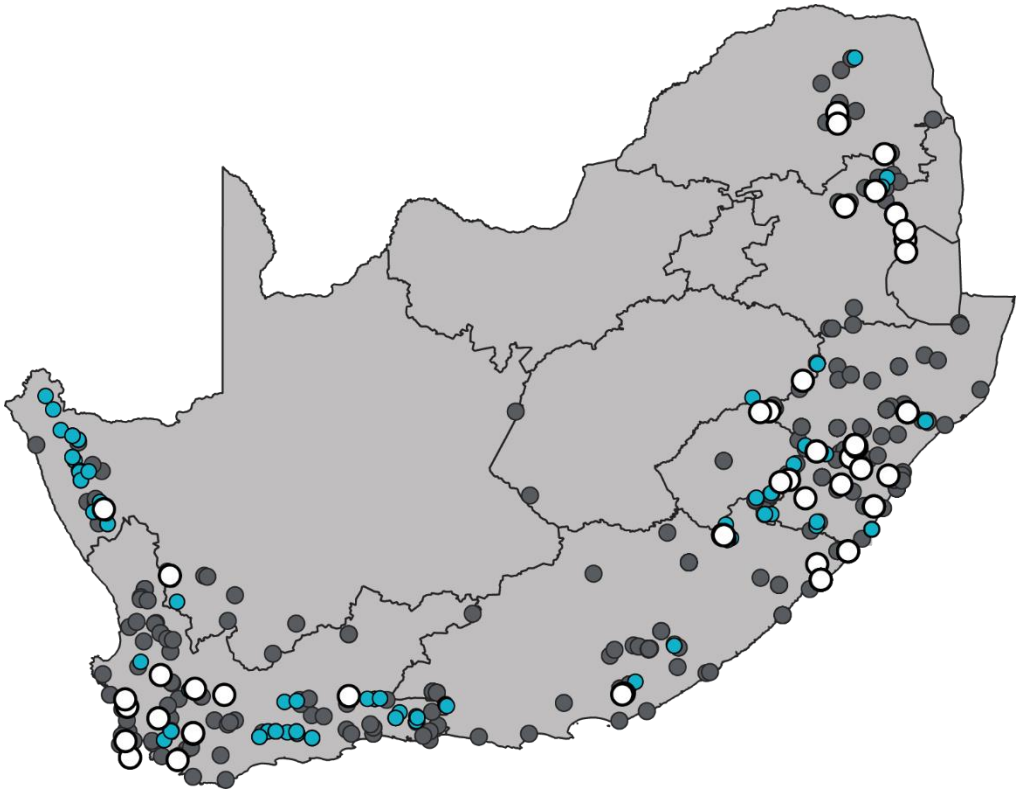


Figure 3.1 Distribution of sampled Nemestrinidae in South Africa, eSwatini and Lesotho. White circles indicate the localities of accessions used for the phylogenetic analysis, blue circles indicate additional sites that were sampled for this study and grey circles are historical records from Global Biodiversity Information Facility.

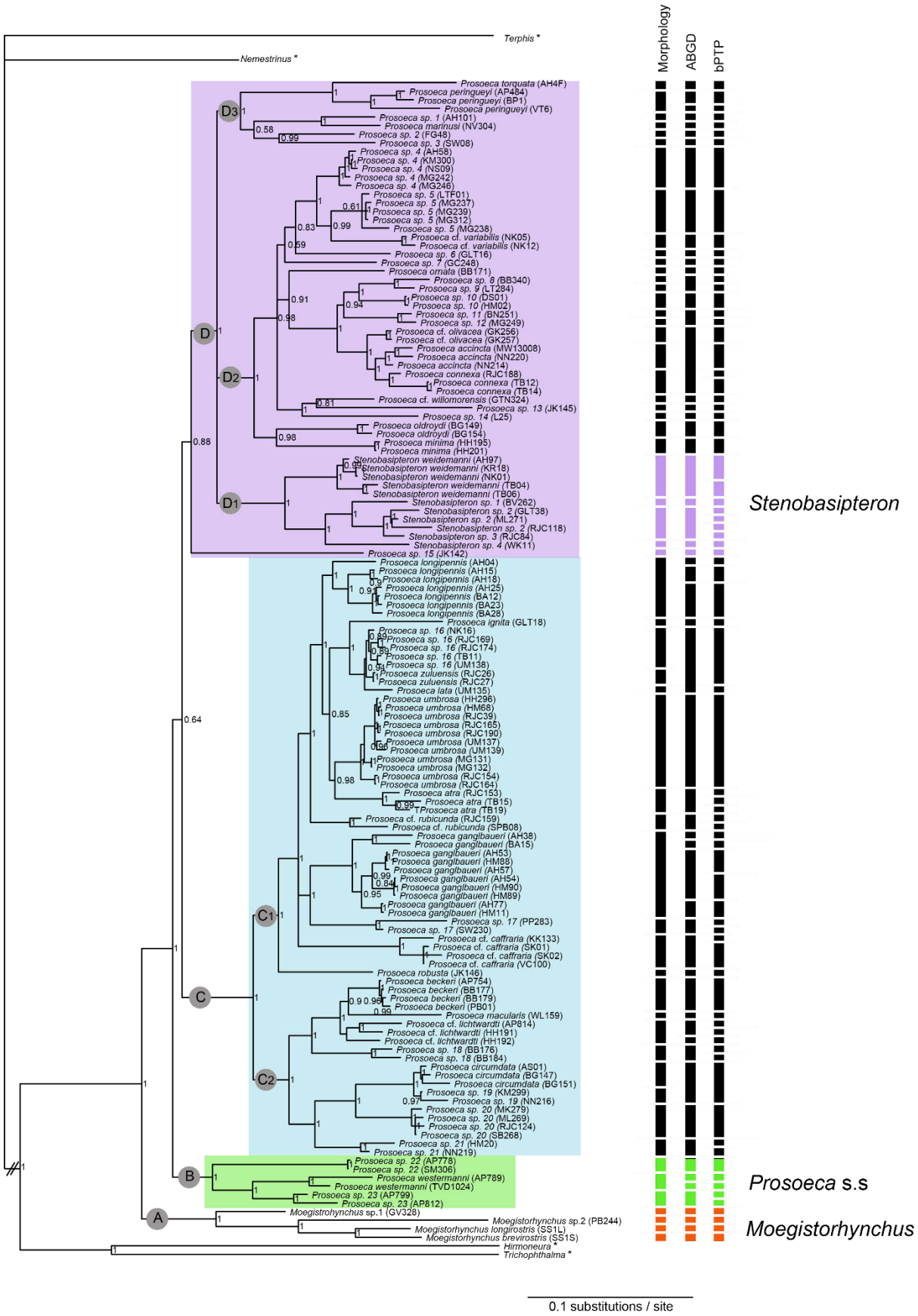

Morphology
ABGD
bPTP
Terphis *
Nemestrinus *
Prosoeca torquata (AH4F)
Prosoeca peringueyi (AP484)
Prosoeca peringueyi (BP1)
Prosoeca peringueyi (VT6)
Prosoeca sp. 1 (AH101)
Prosoeca marinusi (NV304)
Prosoeca sp. 2 (FG48)
Prosoeca sp. 3 (SW08)
Prosoeca sp. 4 (AH58)
Prosoeca sp. 4 (KM300)
Prosoeca sp. 4 (NS09)
Prosoeca sp. 4 (MG242)
Prosoeca sp. 4 (MG246)
Prosoeca sp. 5 (LTF01)
Prosoeca sp. 5 (MG237)
Prosoeca sp. 5 (MG239)
Prosoeca sp. 5 (MG312)
Prosoeca sp. 5 (MG238)
Prosoeca cf. variabilis (NK05)
Prosoeca cf. variabilis (NK12)
Prosoeca sp. 6 (GLT16)
Prosoeca sp. 7 (GC248)
Prosoeca ornata (BB171)
Prosoeca sp. 8 (BB340)
Prosoeca sp. 9 (LT284)
Prosoeca sp. 10 (DS01)
Prosoeca sp. 10 (HM02)
Prosoeca sp. 11 (BN251)
Prosoeca sp. 12 (MG249)
Prosoeca cf. olivacea (GK256)
Prosoeca cf. olivacea (GK257)
Prosoeca accincta (MW13008)
Prosoeca accincta (NN220)
Prosoeca accincta (NN214)
Prosoeca connexa (RJC188)
Prosoeca connexa (TB12)
Prosoeca connexa (TB14)
Prosoeca cf. willomorensis (GTN324)
Prosoeca sp. 13 (JK145)
Prosoeca sp. 14 (L25)
Prosoeca oldroydi (BG149)
Prosoeca oldroydi (BG154)
Prosoeca minima (HH195)
Prosoeca minima (HH201)
Stenobasipteron weidemanni (AH97)
Stenobasipteron weidemanni (KR18)
Stenobasipteron weidemanni (NK01)
Stenobasipteron weidemanni (TB04)
Stenobasipteron weidemanni (TB06)
Stenobasipteron sp. 1 (BV262)
Stenobasipteron sp. 2 (GLT38)
Stenobasipteron sp. 2 (ML271)
Stenobasipteron sp. 2 (RJC118)
Stenobasipteron sp. 3 (RJC84)
Stenobasipteron sp. 4 (WK11)
Prosoeca sp. 15 (JK142)
Prosoeca longipennis (AH04)
Prosoeca longipennis (AH15)
Prosoeca longipennis (AH18)
Prosoeca longipennis (AH25)
Prosoeca longipennis (BA12)
Prosoeca longipennis (BA23)
Prosoeca longipennis (BA28)
Prosoeca ignita (GLT18)
Prosoeca sp. 16 (NK16)
Prosoeca sp. 16 (RJC169)
Prosoeca sp. 16 (RJC174)
Prosoeca sp. 16 (TB11)
Prosoeca sp. 16 (UM138)
Prosoeca zuluensis (RJC26)
Prosoeca zuluensis (RJC27)
Prosoeca lata (UM135)
Prosoeca umbrosa (HH296)
Prosoeca umbrosa (HM68)
Prosoeca umbrosa (RJC39)
Prosoeca umbrosa (RJC165)
Prosoeca umbrosa (RJC190)
Prosoeca umbrosa (UM137)
Prosoeca umbrosa (UM139)
Prosoeca umbrosa (MG131)
Prosoeca umbrosa (MG132)
Prosoeca umbrosa (RJC154)
Prosoeca umbrosa (RJC164)
Prosoeca atra (RJC153)
Prosoeca atra (TB15)
Prosoeca atra (TB19)
Prosoeca cf. rubicunda (RJC159)
Prosoeca cf. rubicunda (SPB08)
Prosoeca ganglbaueri (AH38)
Prosoeca ganglbaueri (BA15)
Prosoeca ganglbaueri (AH53)
Prosoeca ganglbaueri (HM88)
Prosoeca ganglbaueri (AH57)
Prosoeca ganglbaueri (AH54)
Prosoeca ganglbaueri (HM90)
Prosoeca ganglbaueri (HM89)
Prosoeca ganglbaueri (AH77)
Prosoeca ganglbaueri (HM11)
Prosoeca sp. 17 (PP283)
Prosoeca sp. 17 (SW230)
Prosoeca cf. caffraria (KK133)
Prosoeca cf. caffraria (SK01)
Prosoeca cf. caffraria (SK02)
Prosoeca cf. caffraria (VC100)
Prosoeca robusta (JK146)
Prosoeca beckeri (AP754)
Prosoeca beckeri (BB177)
Prosoeca beckeri (BB179)
Prosoeca beckeri (PB01)
Prosoeca macularis (WL159)
Prosoeca cf. lichtwardti (AP814)
Prosoeca cf. lichtwardti (HH191)
Prosoeca cf. lichtwardti (HH192)
Prosoeca sp. 18 (BB176)
Prosoeca sp. 18 (BB184)
Prosoeca circumdata (AS01)
Prosoeca circumdata (BG147)
Prosoeca circumdata (BG151)
Prosoeca sp. 19 (KM299)
Prosoeca sp. 19 (NN216)
Prosoeca sp. 20 (MK279)
Prosoeca sp. 20 (ML269)
Prosoeca sp. 20 (RJC124)
Prosoeca sp. 20 (SB268)
Prosoeca sp. 21 (HM20)
Prosoeca sp. 21 (NN219)
Prosoeca sp. 22 (AP778)
Prosoeca sp. 22 (SM306)
Prosoeca westermanni (AP789)
Prosoeca westermanni (TVD1024)
Prosoeca sp. 23 (AP799)
Prosoeca sp. 23 (AP812)
Moegistrohynchus sp.1 (GV328)
Moegistorhynchus sp.2 (PB244)
Moegistorhynchus longirostris (SS1L)
Moegistorhynchus brevirostris (SS1S)
Himoneura *
Trichophthalma *
D3
D
D2
D1
C1
C
C2
B
A
Stenobasipteron
Prosoeca s.s
Moegistorhynchus
0.1 substitutions / site

Figure 3.2 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the concatenated dataset of COI, 16S, 28S and CAD. Numbers alongside branches refer to posterior probability values and branches are proportional to the number of substitutional changes per site. Letters indicate nodes referred to in the text. Colour blocks indicate the three major clades of *Prosoeca*. Vertical bars show morphospecies delineation for morphology, Automatic Barcode Gap Discovery and Bayesian Poisson Tree Process respectively. Colour sections of the vertical bars indicate clades that contain type specimens of defined genera. Asterisks (*) refer to sequences obtained from GenBank. The branch lengths of *Nemestrinus* and *Terphis* have been artificially shortened for display purposes.

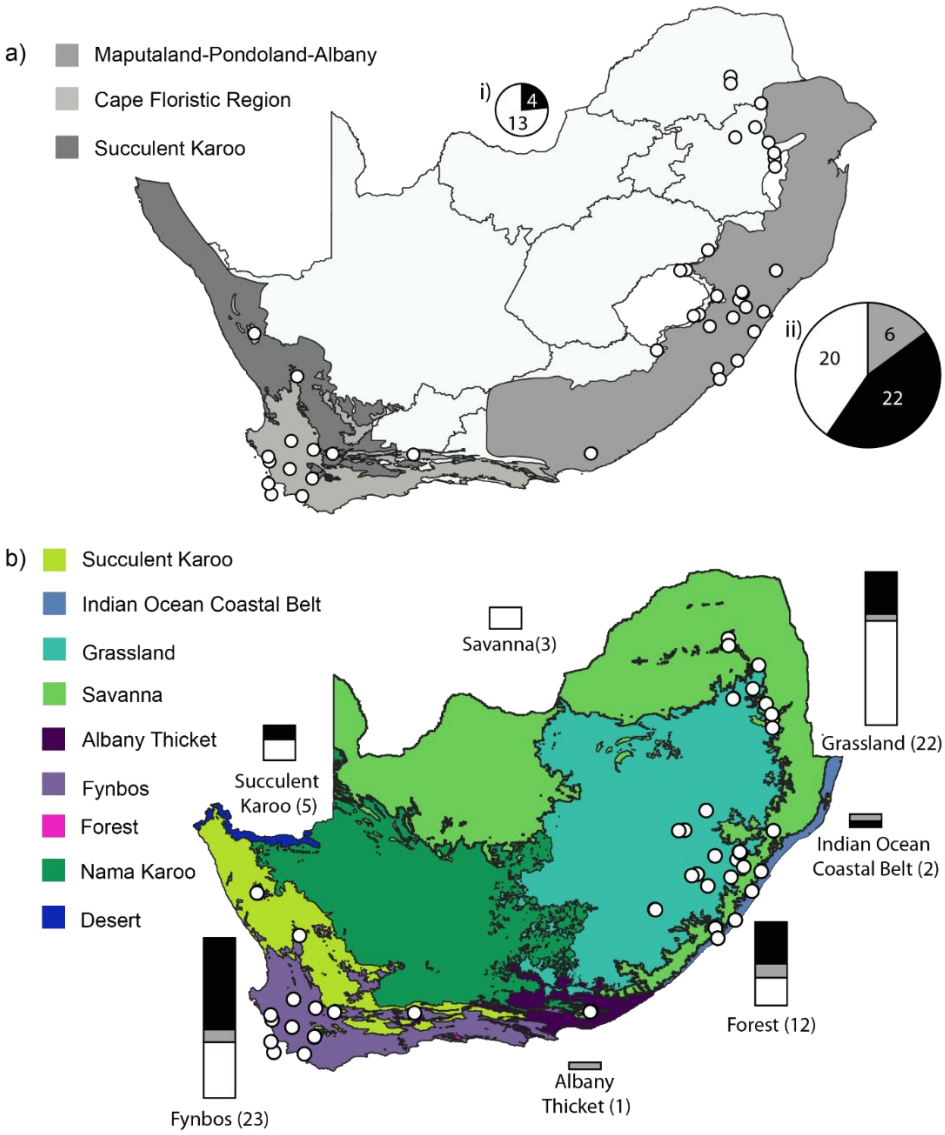


Figure 3.3 The distribution of all Nemestrinidae sampled for the phylogenetic analysis in this study in South Africa, Lesotho and eSwatini. (A) Global biodiversity hotspots are marked on the map in greyscale with the proportion of described (black), cf. (grey) and undescribed (white) morphospecies occuring i) outside the biodiversity hotspots and ii) within the biodiversity hotspots. The relative size of the pie charts is representative for differences in the total number of morphospecies. (B) Biomes of South Africa with bar charts indicating the proportion of described (black), cf. (grey) and undescribed (white) morphospecies found in each biome. The total number of morphospecies found in each biome is indicated in brackets. The total number of mophospecies per biome includes cases where morphospecies occur in multiple biomes.

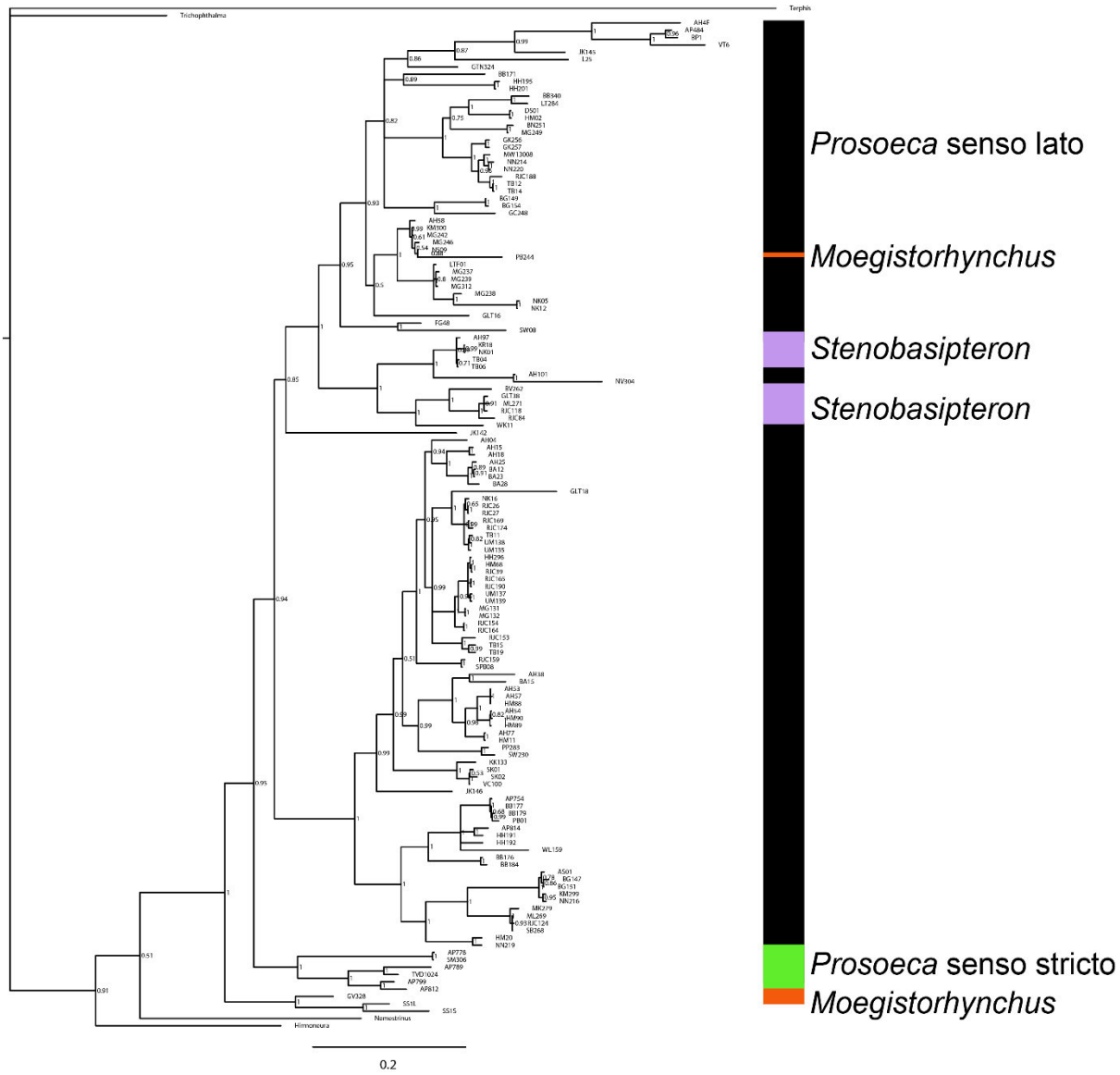


Figure S3.1 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the separate gene trees for COI. Numbers alongside branches refer to posterior probability support values and branches are proportional to the number of substitutional changes per site.

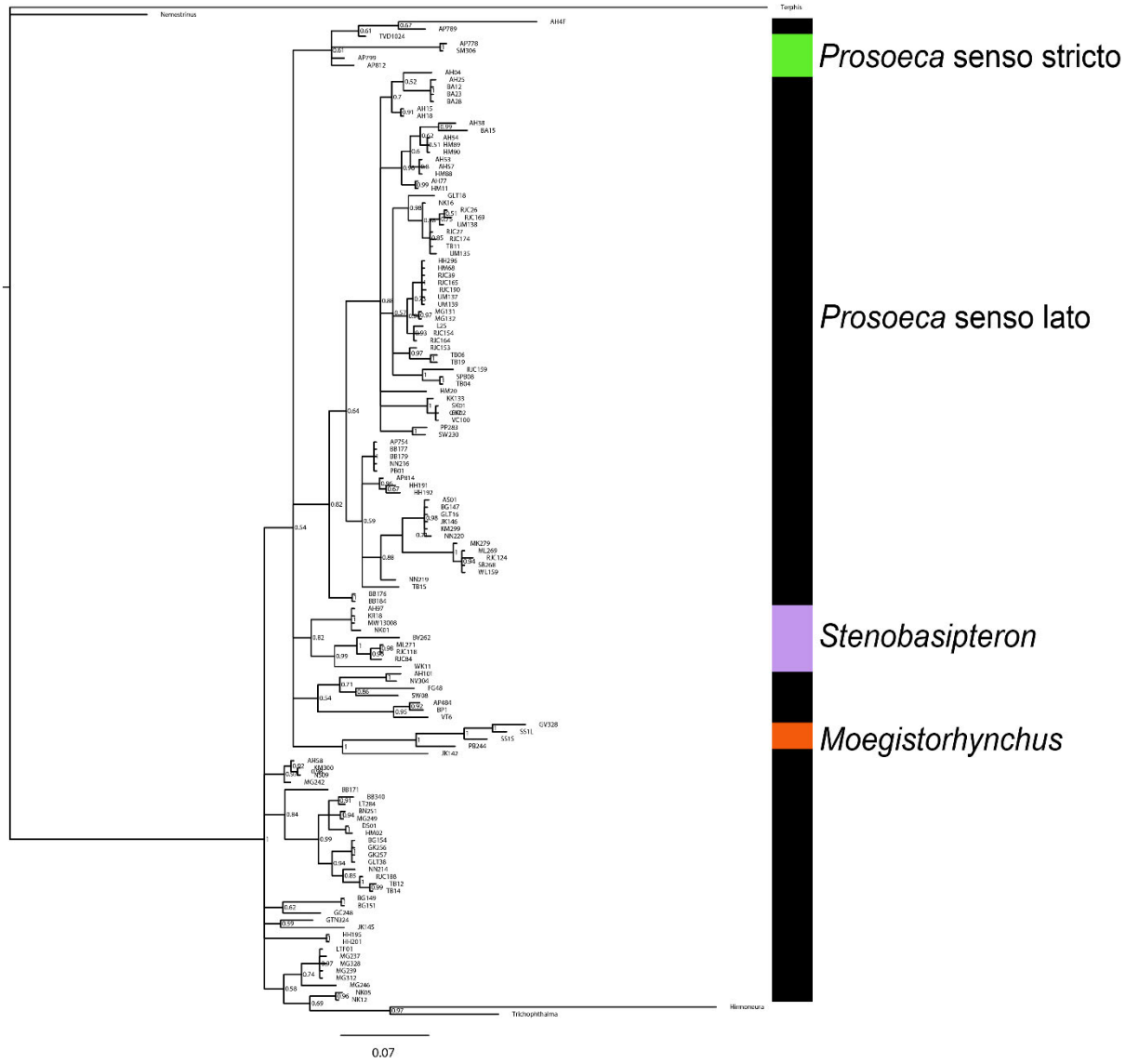


Figure S3.2 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the separate gene trees for 16S. Numbers alongside branches refer to posterior probability support values and branches are proportional to the number of substitutional changes per site.

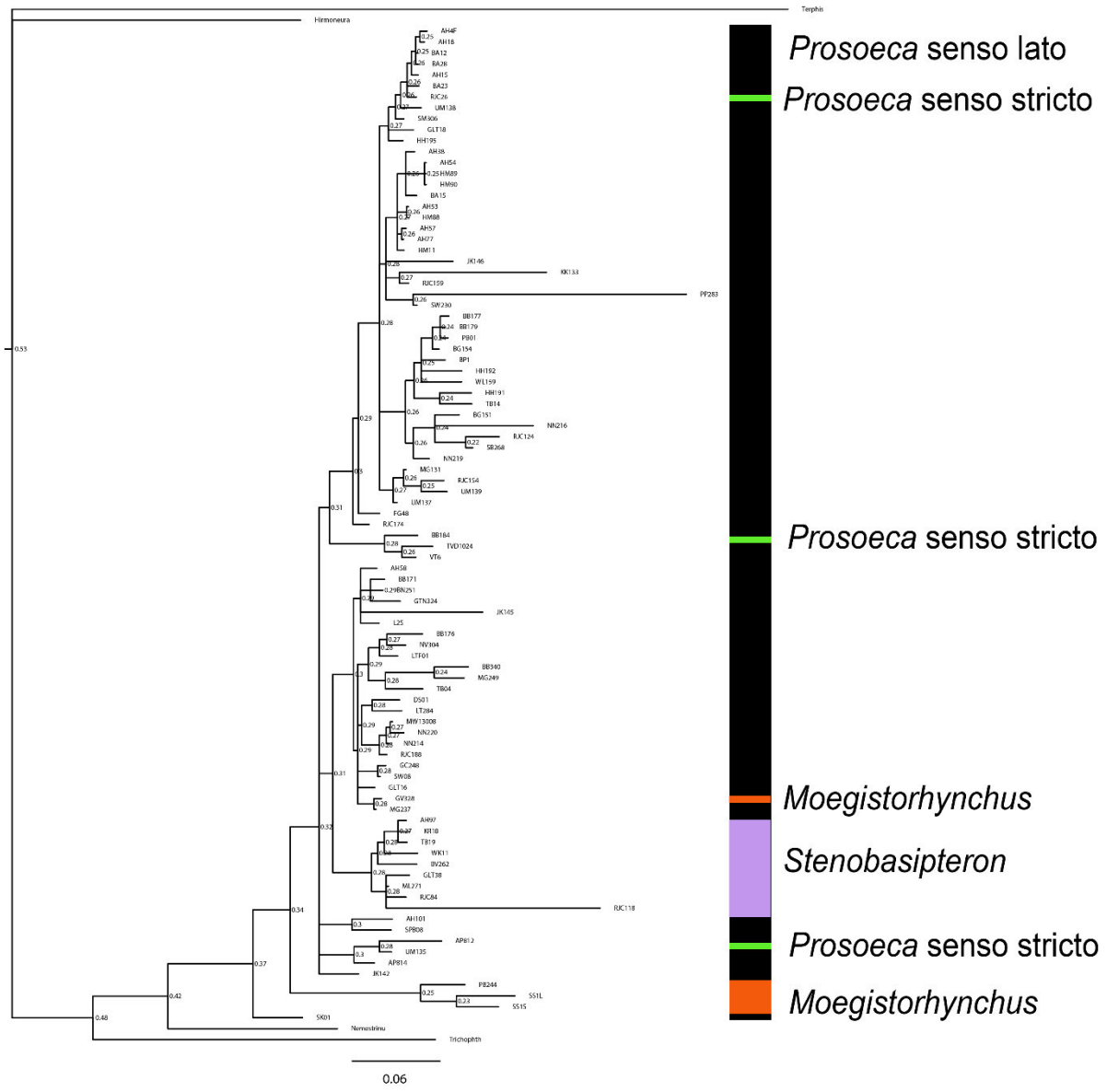


Figure S3.3 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the separate gene trees for CAD. Numbers alongside branches refer to posterior probability support values and branches are proportional to the number of substitutional changes per site.

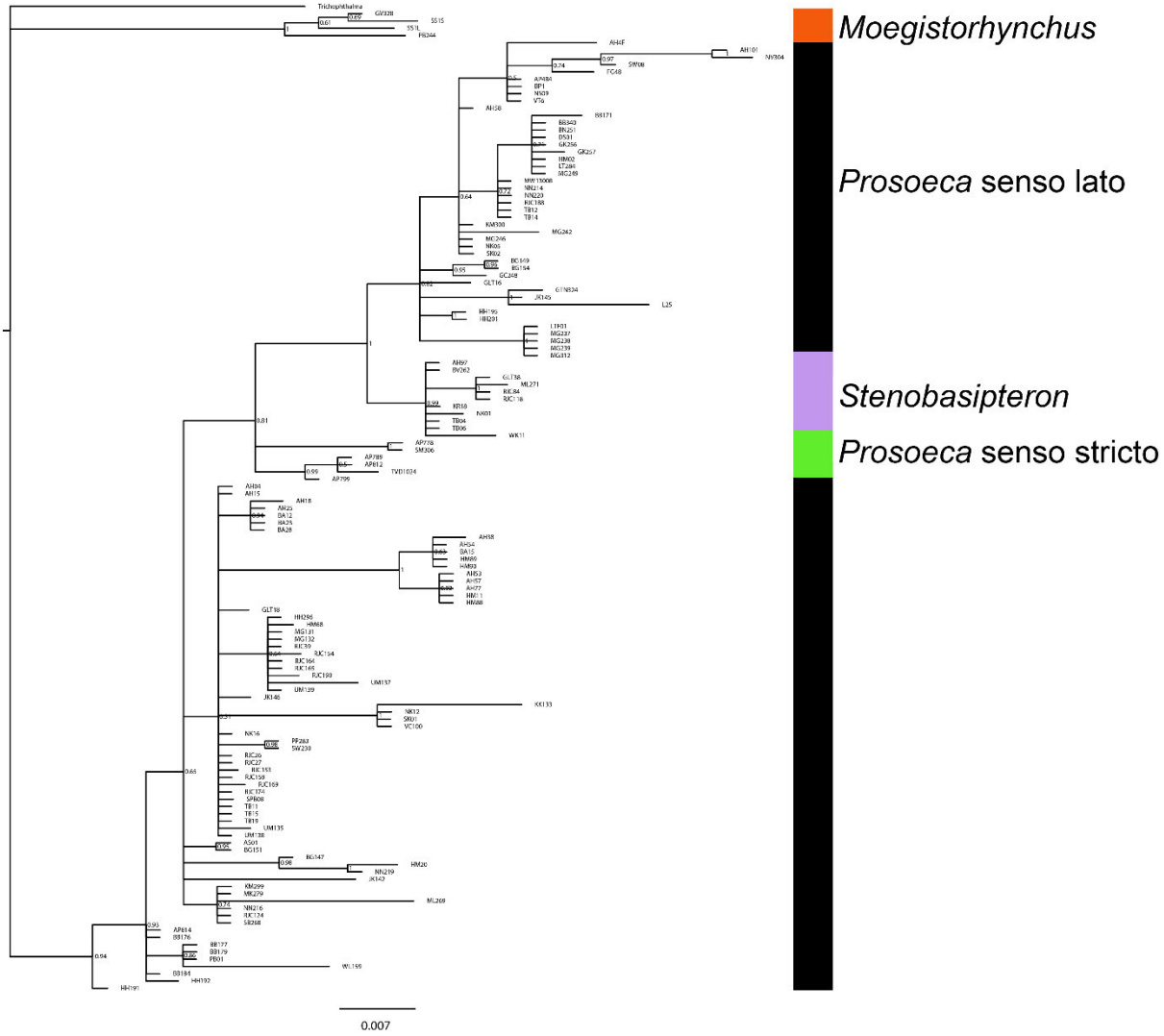


Figure S3.4 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the separate gene trees for 28S. Numbers alongside branches refer to posterior probability support values and branches are proportional to the number of substitutional changes per site.

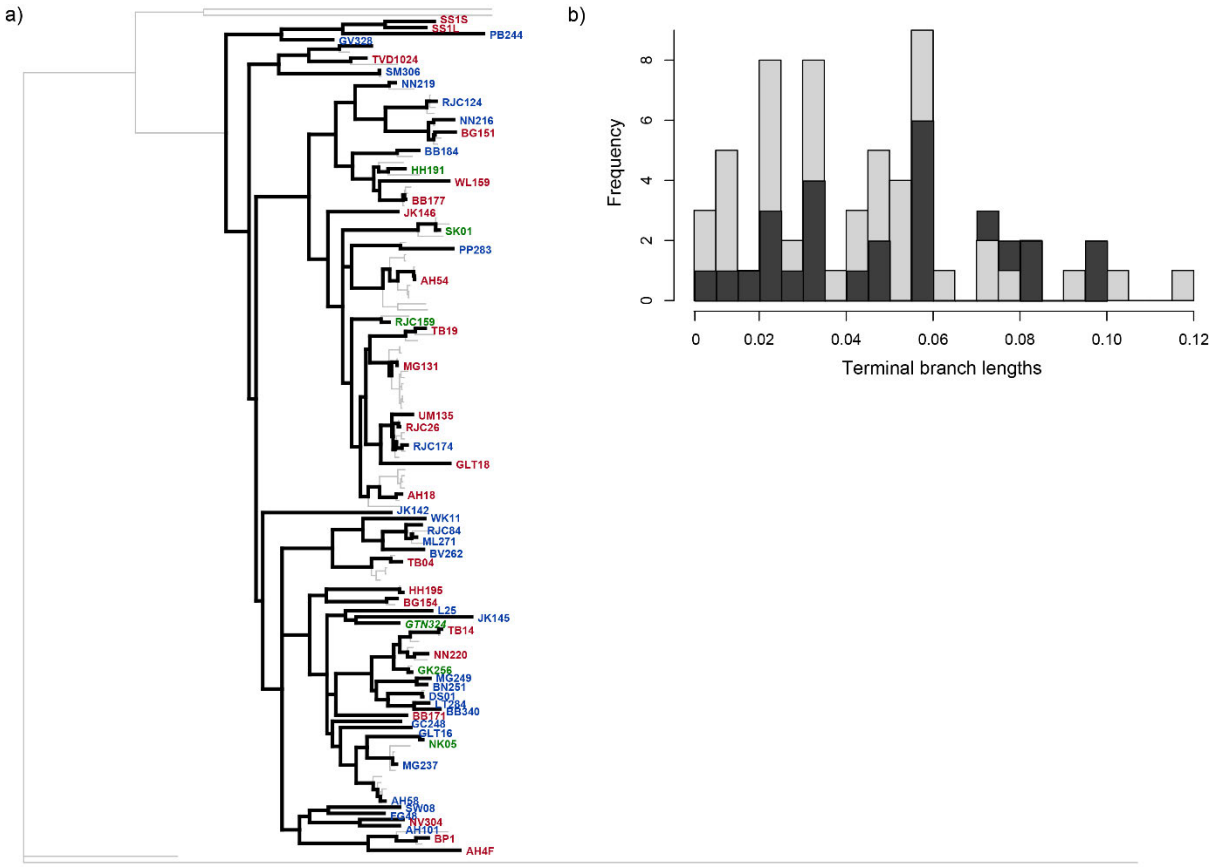


Figure S3.5 (A) 50% majority-rule consensus tree resulting from a Bayesian Inference analysis of the concatenated dataset with the selected single accession per species overlayed with black branches. Colour of sequences ID corresponds to designation as described (red), cf. (pink) and undescribed (blue). (B) Frequency distribution of terminal branch lengths of the whole tree with one accession per species in light grey and a tree with only described and cf. morphospecies included in dark grey.

Table S3.1 Table of localities and specimen details of accessions used for the phylogenetic analyses.

Sequence ID	*Species*	Locality	Latitude	Longitude	Date collected	Collected by
SS1S	*Moegistorhynchus brevirostris*	Silverstroomstrand	-33.528550	18.446101	20 October 2015	B. Anderson
SS1L	*Moegistorhynchus longirostris*	Silverstroomstrand	-33.528550	18.446101	20 October 2015	B. Anderson
GV328	*Moegistorhynchus sp. 1*	Groenvlei	-30.319344	18.067708	10 August 2015	A. Ellis
PB244	*Moegistorhynchus sp.2*	Picketberg	-32.79855	18.665922	01 December 2015	F. Grenier
MW13008	*Prosoeca accincta*	Bushmans Nek Pass	-29.879977	29.075102	22 February 2013	M. Whitehead
NN214	*Prosoeca accincta*	Naudes Nek Pass	-30.732387	28.138476	12 February 2019	S. Klumpers
NN220	*Prosoeca accincta*	Naudes Nek Pass	-30.732387	28.138476	12 February 2019	S. Klumpers
RJC153	*Prosoeca atra*	Kologha Forest Reserve	-32.53742	27.346	12 April 2018	R. Cozien
TB15	*Prosoeca atra*	Mbotyi	-31.465392	29.730392	12 May 2019	T. Bellingham
TB19	*Prosoeca atra*	Mbotyi	-31.465393	29.730393	12 May 2019	T. Bellingham
AP754	*Prosoeca beckeri*	Kleinrivierberge	-34.395444	19.265669	10 October 2017	A. Pauw
BB177	*Prosoeca beckeri*	Boosmansbos	-33.929018	20.86353	13 January 2019	B. Anderson
BB179	*Prosoeca beckeri*	Boosmansbos	-33.929018	20.86353	13 January 2019	B. Anderson
PB01	*Prosoeca beckeri*	Boosmansbos	-33.929018	20.86353	22 February 2015	P. Botha
KK133	*Prosoeca cf. caffraria*	Kranzkloof	-29.772535	30.83043	04 April 2018	G. Theron
SK01	*Prosoeca cf. caffraria*	Vernon Crooke Nature Reserve	-30.274244	30.594822	08 April 2019	S. Klumpers
SK02	*Prosoeca cf. caffraria*	Vernon Crooke Nature Reserve	-30.274244	30.594822	08 April 2019	S. Klumpers
VC100	*Prosoeca cf. caffraria*	Vernon Crooke Nature Reserve	-30.274244	30.594822	30 March 2018	G. Theron
AP814	*Prosoeca cf. lichtwardti*	Porterville	-33.014188	18.991504	September 2018	A. Pauw
HH191	*Prosoeca cf. lichtwardti*	Hottentots Holland Nature Reserve	-34.069298	19.047765	25 January 2019	G. Theron
HH192	*Prosoeca cf. lichtwardti*	Hottentots Holland Nature Reserve	-34.069298	19.047765	25 January 2019	B. Anderson
GK256	*Prosoeca cf. olivacea*	Graskop	-24.932787	30.808752	25 March 2019	G. Theron
GK257	*Prosoeca cf. olivacea*	Graskop	-24.932787	30.808752	25 March 2019	T. van der Niet
RJC159	*Prosoeca cf. rubicunda*	Swartberg	-33.400	22.355	25 March 2019	R. Cozien
SBP08	*Prosoeca cf. rubicunda*	Swartberg Pass	-33.355044	22.052139	12 April 2015	E. Newman

NK05	*Prosoeca cf. variabilis*	Nkandla Forest Reserve	-28.743981	31.139830	13 April 2018	T. van der Niet
NK12	*Prosoeca cf. variabilis*	Nkandla Forest Reserve	-28.743981	31.139830	13 April 2019	T. van der Niet
GTN324	*Prosoeca cf. willomoresis*	Grahamstown	-33.311184	26.526693	20 April 2017	
AS01	*Prosoeca circumdata*	Howick	-29.474308	30.216809	13 April 2019	A. Shuttleworth
BG147	*Prosoeca circumdata*	Pietermaritzburg Botanical Gardens	-29.607538	30.34779	24 April 2018	G. Theron
BG151	*Prosoeca circumdata*	Pietermaritzburg Botanical Gardens	-29.607538	30.34779	24 April 2018	G. Theron
RJC188	*Prosoeca connexa*	Ingeli	-30.520	29.684	09 May 2018	R. Cozien
TB12	*Prosoeca connexa*	Mbotyi	-31.465390	29.730390	12 May 2019	T. Bellingham
TB14	*Prosoeca connexa*	Mbotyi	-31.465391	29.730391	12 May 2019	T. Bellingham
AH38	*Prosoeca ganglbaueri*	Uniondale	-33.684495	23.14911	18 March 2011	M. Whitehead
AH53	*Prosoeca ganglbaueri*	Naudes Nek Pass	-30.73242	28.14157	08 February 2013	M. Whitehead
AH54	*Prosoeca ganglbaueri*	Naudes Nek Pass	-30.77598	28.22365	07 February 2013	M. Whitehead
AH57	*Prosoeca ganglbaueri*	Naudes Nek Pass	-30.73242	28.14157	08 February 2013	M. Whitehead
AH77	*Prosoeca ganglbaueri*	Witsieshoek Resort	-28.68521	28.9	27 January 2013	B. Anderson
BA15	*Prosoeca ganglbaueri*	Meiringspoort	-33.392784	22.559937	14 April 2013	E. Newman
HM11	*Prosoeca ganglbaueri*	Witsieshoek Resort	-28.728861	28.891777	28 January 2014	H. de Moor
HM88	*Prosoeca ganglbaueri*	Naudes Nek Pass	-30.73677	28.1395277	27 February 2014	H. de Moor
HM89	*Prosoeca ganglbaueri*	Naudes Nek Pass	-30.73677	28.138722	27 February 2014	H. de Moor
HM90	*Prosoeca ganglbaueri*	Naudes Nek Pass	-30.743166	28.154055	27 February 2014	H. de Moor
GLT18	*Prosoeca ignita*	Haenertsberg	-23.880965	29.99877	19 March 2018	R. Cozien
UM135	*Prosoeca lata*	Umtamvuna Nature Reserve	-31.00363	30.17356	09 April 2018	G. Theron
UM138	*Prosoeca lata*	Umtamvuna Nature Reserve	-31.00363	30.17356	09 April 2018	L. Harder
AH04	*Prosoeca longipennis*	Qacha's Nek	-30.13991	28.67692	Missing	B. Anderson
AH15	*Prosoeca longipennis*	Suurbraak	-34.017633	20.59802	19 March 2010	E. Newman
AH18	*Prosoeca longipennis*	Riversdale	-34.036519	21.455702	26 March 2010	E. Newman
AH25	*Prosoeca longipennis*	Riversdale	-33.966434	21.211	31 March 2010	E. Newman
BA12	*Prosoeca longipennis*	Baviaanskloof	-33.489712	23.625214	14 April 2013	E. Newman
BA23	*Prosoeca longipennis*	Baviaanskloof	-33.508497	23.63649	09 March 2013	E. Newman
BA28	*Prosoeca longipennis*	Uniondale	-33.711015	22.813879	07 March 2014	E. Newman

WL159	*Prosoeca macularis*	Waylands Nature Reserve	-33.408801	18.413050	22 September 2018	S. Johnson
NV304	*Prosoeca marinusi*	Niewoutville Botanical Garderns	-31.398152	19.141070	25 August 2019	S. Johnson
HH195	*Prosoeca minima*	Hottentots Holland Nature Reserve	-34.069298	19.047765	25 January 2019	G. Theron
HH201	*Prosoeca minima*	Hottentots Holland Nature Reserve	-34.069298	19.047765	25 January 2019	B. Anderson
BG149	*Prosoeca oldroydi*	Pietermaritzburg Botanical Gardens	-29.607538	30.34779	09 May 2018	G. Theron
BG154	*Prosoeca oldroydi*	Pietermaritzburg Botanical Gardens	-29.607538	30.34779	09 May 2018	G. Theron
BB171	*Prosoeca ornata*	Boosmansbos	-33.929018	20.86353	13 January 2019	B. Anderson
AP484	*Prosoeca peringueyi*	Towsrivier	-33.336127	20.025910	08 August 2009	A. Pauw
BP1	*Prosoeca peringueyi*	Botterkloof	-31.823601	19.256734	August 2017	F. Grenier
VT6	*Prosoeca peringueyi*	Vandersterberg	-28.480114	17.11973	01 September 2017	F. Grenier
JK146	*Prosoeca robusta*	Jonaskop	-33.957261	19.520465	17 April 2018	G. Theron
AH4F	*Prosoeca torquata*	Ariehoek	-30.244758	18.050258	August 2017	F. Grenier
HH296	*Prosoeca umbrosa*	Hella Hella	-29.919002	30.063956	19 May 2019	S. Klumpers
HM68	*Prosoeca umbrosa*	Nelson's Kop	-28.232194	29.437277	22 February 2014	H. de Moor
MG131	*Prosoeca umbrosa*	Mount Gilboa	-29.308598	30.309024	01 April 2018	M. Castañeda-Zárate
MG132	*Prosoeca umbrosa*	Mount Gilboa	-29.308598	30.309024	01 April 2018	M. Castañeda-Zárate
RJC39	*Prosoeca umbrosa*	Royal Natal	-28.742383	28.742383	06 March 2016	R. Cozien
RJC154	*Prosoeca umbrosa*	Kologha Forest Reserve	-32.53742	27.346	12 April 2018	R. Cozien
RJC164	*Prosoeca umbrosa*	Hogsback	-32.601	26.960	22 April 2018	R. Cozien
RJC165	*Prosoeca umbrosa*	Ingeli	-30.520	29.684	26 April 2018	R. Cozien
RJC190	*Prosoeca umbrosa*	Ingeli	-30.520	29.684	11 May 2018	R. Cozien
UM137	*Prosoeca umbrosa*	Umtamvuna Nature Reserve	-31.00363	30.17356	09 April 2018	L. Harder
UM139	*Prosoeca umbrosa*	Umtamvuna Nature Reserve	-31.00363	30.17356	09 April 2018	L. Harder
AP789	*Prosoeca westermanni*	Boland	-33.717485	18.954361	02 October 2017	A. Pauw
TVD1024	*Prosoeca westermanni*	Prince Alfred Hamlet	-33.236666	19.55	14 September 2016	T. van der Niet
RJC26	*Prosoeca zuluensis*	Bisley Nature Reserve	-29.658909	30.387038	20 February 2014	R. Cozien
RJC27	*Prosoeca zuluensis*	Bisley Nature Reserve	-29.658909	30.387038	20 February 2014	R. Cozien
AH101	*Prosoeca sp. 1*	Nieuwoudtville	-31.387089	19.176518	01 August 2012	J. coville

FG48	*Prosoeca sp. 2*	Namaqualand	-30.371906	18.092013	September 2018	F. Grenier
SW08	*Prosoeca sp. 3*	Soetwater			September 2018	F. Grenier
AH58	*Prosoeca sp. 4*	Naudes Nek Pass	-30.73242	28.14157	13 February 2013	M. Whitehead
KM300	*Prosoeca sp. 4*	Kamberg	-29.381052	29.660820	15 March 2019	M. Castañeda-Zárate
MG242	*Prosoeca sp. 4*	Mount Gilboa	-29.843930	29.210246	17 January 2019	M. Castañeda-Zárate
MG246	*Prosoeca sp. 4*	Mount Gilboa	-29.843930	29.210246	21 February 2019	S. Klumpers
NS09	*Prosoeca sp. 4*	Ntsikeni Nature Reserve	-30.141562	29.478780	8 February 2019	T. van der Niet
LTF01	*Prosoeca sp. 5*	Mount Gilboa	-29.276283	30.281003	03 January 2014	M. Castañeda-Zárate
MG237	*Prosoeca sp. 5*	Mount Gilboa	-29.276283	30.281003	14 January 2019	M. Castañeda-Zárate
MG238	*Prosoeca sp. 5*	Mount Gilboa	-29.276283	30.281003	06 January 2019	M. Castañeda-Zárate
MG239	*Prosoeca sp. 5*	Mount Gilboa	-29.276283	30.281003	06 January 2019	M. Castañeda-Zárate
MG312	*Prosoeca sp. 5*	Mount Gilboa	-29.843930	29.210246	15 January 2020	T. van der Niet
GLT16	*Prosoeca sp. 6*	Haenertsberg	-23.880965	29.99877	19 March 2018	G. Theron
GC248	*Prosoeca sp. 7*	Giants Castle Nature Reserve	-29.276286	30.281006	22 February 2019	S. Klumpers
BB340	*Prosoeca sp. 8*	Bulembo border post	-25.943825	31.106867	15 January 2019	S. Johnson
LT284	*Prosoeca sp. 9*	Long Tom Pass	-25.149548	30.619543	31 March 2019	G. Theron
DS01	*Prosoeca sp. 10*	Dullstroom	-25.396739	30.126015	19 January 2018	S. Johnson
HM02	*Prosoeca sp. 11*	Verloren vlei	-25.402837	30.119673	24 January 2014	H. de Moor
BN251	*Prosoeca sp. 12*	Bushmans Nek Pass	-29.843928	29.210244	01 March 2019	S. Klumpers
MG249	*Prosoeca sp. 12*	Mount Gilboa	-29.278267	30.287912	24 February 2019	S. Klumpers
JK145	*Prosoeca sp. 13*	Jonaskop	-33.968685	19.518869	17 April 2018	G. Theron
L25	*Prosoeca sp. 14*	Langkloof	-30.55937	18.128272	01 September 2017	F. Grenier
JK142	*Prosoeca sp. 15*	Jonaskop	-33.957259	19.520465	17 April 2018	B. Anderson
NK16	*Prosoeca sp. 16*	Nkandla Forest Reserve	-28.743981	31.139830	13 April 2019	T. van der Niet
RJC169	*Prosoeca sp. 16*	Eshowe	-28.891083	31.4349	02 May 2018	R. Cozien
RJC174	*Prosoeca sp. 16*	Eshowe	-28.891083	31.4349	02 May 2018	R. Cozien

TB11	*Prosoeca sp. 16*	Mbotyi	-31.465389	29.730389	12 May 2019	T. Bellingham
PP283	*Prosoeca sp. 17*	Paardeplaats Nature Reserve	-25.095679	30.562498	30 March 2019	M. Rule
SW230	*Prosoeca sp. 17*	Malololtja National Park	-26.140859	31.117373	February 2019	S. Johnson
BB176	*Prosoeca sp. 18*	Boosmansbos	-33.929018	20.86353	13 January 2019	B. Anderson
BB184	*Prosoeca sp. 18*	Boosmansbos	-33.929018	20.86353	13 January 2019	B. Anderson
KM299	*Prosoeca sp. 19*	Kamberg	-29.381052	29.660820	15 March 2019	M. Castañeda-Zárate
NN216	*Prosoeca sp. 19*	Naudes Nek Pass	-30.736268	28.13939	12 February 2019	S. Klumpers
MK279	*Prosoeca sp. 20*	Mariepskop Nature Reserve	-25.78014	31.09808	29 March 2019	M. Rule
ML269	*Prosoeca sp. 20*	Mountainlands Nature Reserve	-25.7735	31.08543	27 March 2019	M. Rule
RJC124	*Prosoeca sp. 20*	Saddleback Pass	-25.79536	31.094741	25 March 2018	R. Cozien
SB268	*Prosoeca sp. 20*	Saddleback Pass	-25.79787	31.104046	27 March 2019	M. Rule
HM20	*Prosoeca sp. 21*	Sentinel	-28.728861	28.891777	28 January 2014	H. de Moor
NN219	*Prosoeca sp. 21*	Naudes Nek Pass	-30.736268	28.13939	12 February 2019	S. Klumpers
AP778	*Prosoeca sp. 22*	Cape Peninsula	-34.354020	18.493001	23 October 2017	A. Pauw
SM306	*Prosoeca sp. 22*	Silvermine Nature Reserve	-34.085932	18.418983	05 October 2019	B. Anderson
AP799	*Prosoeca sp. 23*	Boland	-33.717485	18.954361	23 October 2017	A. Pauw
AP812	*Prosoeca sp. 23*	Cape Peninsula	-34.354020	18.493001	September 2018	A. Pauw
BV262	*Stenobasipteron sp. 1*	Bridal Veil Falls	-25.082029	30.723459	25 March 2019	T. van der Niet
GLT38	*Stenobasipteron sp. 2*	Nelspruit	-25.531727	30.953037	24 March 2018	G. Theron
ML271	*Stenobasipteron sp. 2*	Mountainlands Nature Reserve	-25.7735	31.08543	27 March 2019	G. Theron
RJC118	*Stenobasipteron sp. 2*	Nelspruit	-25.531727	30.953037	24 March 2018	R. Cozien
RJC84	*Stenobasipteron sp. 3*	Abel Erasmus Pass	-24.55128	30.76479	22 March 2018	R. Cozien
WK11	*Stenobasipteron sp. 4*	Serala Forest Reserve	-24.055322	30.005245	17 March 2018	R. Cozien
NK1	*Stenobasipteron weidemanni*	Nkandla Forest Reserve	-28.743981	31.139830	13 April 2019	S. Johnson
AH97	*Stenobasipteron weidermanni*	Msikaba Campsite	-31.212619	29.670470	February 2014	J. coville
KR18	*Stenobasipteron weidermanni*	Karkloof	-29.308598	30.309024	30 March 2018	G. Theron
TB04	*Stenobasipteron weidermanni*	Grahamstown	-33.328778	26.500194	30 April 2019	T. Bellingham
TB06	*Stenobasipteron weidermanni*	Grahamstown	-33.328778	26.500194	30 April 2019	T. Bellingham

Table S3.2 Primer sequences used to amplify the gene regions included in this study.

Primer name	Gene region	Reference	Primer sequence
NE	CO1	Theron *et al.* (2020)	ACT TTA TAY TTT ATY TTT GGA GC
C1-N-2191	CO1	Simon *et al.* (1994)	CCC GGT AAA ATT AAA ATA TAA ACTTC
C1-J-2183	CO1	Simon *et al.* (1994)	CAA CAT TTA TTT TGA TTT TTT GG
TL2-N-3014	CO1	Simon *et al.* (1994)	TCC ATT GCA CTA ATC TGC CAT ATT A
2F	28S	Reemer & Ståhls (2013)	AGA GAG AGT TCA AGA GTA CGT G
3DR	28S	Reemer & Ståhls (2013)	TAG TTC ACC ATC TTT CGG GTC
LR-J-12887	16S	Simon *et al.* (1994)	CCG GTT TGA ACT CAG ATC ATG T
SR-N-13398b	16S	Simon *et al.* (1994)	CRC YTG TTT AWC AAA AAC AT
54F	CAD	Moulton & Wiegmann (2004)	GTN GTN TTY CAR ACN GGN ATG GT
405R	CAD	Moulton & Wiegmann (2004)	GCN GTR TGY TCN GGR TGR AAY TG
CAD-F202-222	CAD	Cisneros (2015)	AAT AAG TGG AAT TGA TAC TAG
CAD-R611-631	CAD	Cisneros (2015)	TGA GGA CTT GGA AGT GAA TGT

Chapter 4

Labile proboscis length evolution in a lineage of flower-visiting flies (Nemestrinidae)

Genevieve L. Theron, Bruce C. Anderson, Allan G. Ellis, Steven D. Johnson & Timotheüs van der Niet

To be submitted to: Evolutionary biology

ABSTRACT

Modification of mouthparts into a proboscis allows for acquisition of nectar from floral tubes. Proboscides evolved independently in most orders of flower-visiting insects and through coevolution with flowers they can even greatly exceed body length. However, extremely exaggerated traits likely incur developmental and functional costs. These costs may constrain, and under some ecological conditions even reverse, evolutionary transitions to longer proboscides in lineages. Using species-level phylogenetic comparative methods we reconstructed the evolution of proboscis length in a clade of long proboscid nemestrinid flies from southern Africa that includes the genera *Moegistorhynchus*, *Prosoeca* and *Stenobasipteron*. Using a pruned phylogenetic tree, we found a positive interspecific allometric relationship between body size and proboscis length after correcting for phylogenetic relationship, but there was no clear phylogenetic signal in either of these traits. A similar positive intraspecific relationship between proboscis length and body size is evident across morphospecies with long proboscides. Using continuous character state reconstruction, we inferred that the ancestor had a proboscis longer than the median of the current ingroup. Discrete stochastic character state mapping showed that there were transitions between all proboscis length categories (short, long and very long) and that transitions away from short proboscides were most common. An analysis of structured rate permutations on phylogenies did not detect any association between proboscis length and the rate of diversification. Our results indicate that absolute proboscis length is a highly labile trait that may reflect adaptations to the local environment for optimal foraging on flowers.

Keywords: allometry, ancestral state reconstruction, Nemestrinidae, proboscis length, trait evolution

INTRODUCTION

Proboscides have evolved convergently in various flower-visiting insects to access energy-rich nectar resources (Krenn *et al.* 2005). A small proportion of flower-visiting insects possess extraordinary proboscides that are longer than their bodies. Such extreme relative lengths have evolved several times independently in Hymenoptera (Euglossini), Lepidoptera (Papilionoidea, Sphingidae) and Diptera (Nemestrinidae, Tabanidae, Acroceridae) (Bauder & Karolyi 2019), but are arguably most impressive in flies because their proboscides cannot be rolled up. Proboscides that are particularly long may allow taxa to exploit a broader range of feeding niches by increasing the variety of resources that are accessible (Haber & Frankie 1989; Martins & Johnson 2013; Klumpers *et al.* 2019). They may additionally facilitate the acquisition of large quantities of nectar from relatively few flower visits and should result in a high net-energy reward for visitors (Haverkamp *et al.* 2016). A longer proboscis may also allow insects to access a larger portion of a nectar resource (ie. to reach the bottom of the nectar tube) as well as the larger quantities of nectar that are associated with longer-tubed flowers (Haber & Frankie 1989; Martins & Johnson 2013; Klumpers *et al.* 2019). Reciprocal directional selection for longer traits between flower corolla tubes and pollinator proboscides may lead to a co-evolutionary race (Anderson & Johnson 2008; Pauw *et al.* 2009), potentially leading to unidirectional trait evolution (Klimov *et al.* 2017; Cardoso-Gustavson *et al.* 2018).

Proboscis lengths are likely under diverse selective pressures to balance optimal foraging performance with potential developmental constraints and functional costs (Krenn *et al.* 2005; Anderson *et al.* 2010). Functional costs of a long proboscis may include reduced nectar uptake rate and increased flower handling time (Kunte 2007; Karolyi *et al.* 2013). The larger pumping organs required for the adequate functioning of the suction action of long proboscides may also represent an additional material cost (Karolyi *et al.* 2012). Increased costs associated with the evolution of particularly long proboscides may therefore restrict the

evolution of this trait (Bauder & Karolyi 2019). Developmental constraints may limit changes in the mechanism by which many exaggerated traits evolve, such as changes in the slope and/or shape of allometric relationships (scaling pattern relating trait size to overall body size) (Wilkinson & Reillo 1994; Klingenberg 2005). These constraints are thought to be the cause of the low variability in static allometric slopes generally found between closely related species, thus potentially restricting phenotype evolution (Gould 1966). Studying the changes in scaling relationships within and between species may provide insights into the evolution of trait morphology.

The evolution of traits, such as wings and metamorphosis, have been suggested to be key innovations (ie. traits associated with an increased net diversification rate within a lineage), leading to the immense diversity in insects (Mayhew 2007; Nicholson *et al.* 2014). In plants it has been argued that the evolution of nectar spurs in flowers was a key innovation that allowed radiation into new pollination niches for angiosperms, driving increased biodiversity in the associated clades (Hodges & Arnold 1995; Fernández-Mazuecos *et al.* 2019). It is not known whether the evolution of proboscides had similar effects on the diversification of insect groups. Long proboscides and long nectar spurs differ in terms of how they affect levels of specialization; longer proboscides often broaden the feeding niche while deeper flowers narrow the available pollinator niche because of morphological filtering (Borrell 2005). While the costs associated with long proboscides may necessitate the use of exclusive high-energy nectar resources, they also facilitate generalization on a wider array of floral species (Borrell 2005; Martins & Johnson 2013; Klumpers *et al.* 2019).

Southern African Nemestrininae are extremely diverse in terms of morphology and ecology with proboscides ranging from 4mm to 100mm in different taxa (Barraclough 2006) (Fig. 1). The nemestrinid proboscis cannot be rolled up like in Lepidoptera but rather swivels through 100 degrees from feeding to resting where it is carried at full length while tucked back

between the legs (Karolyi *et al.* 2012). Nemestrinid proboscides are able to extend further for up to about a third of the resting length while feeding but the degree of extension differs between species (Morita 2011). A handful of southern African nemestrinids with particularly long proboscides are thought to be the sole pollinators for *ca.* 150 plant species, many of which are rare and endangered (Manning & Goldblatt 2000; Johnson 2006b). Species with shorter proboscides are generally thought to act as visitors to generalist flowers, alongside other pollinators (Devoto & Medan 2006; Potgieter *et al.* 2009). In southern Africa, various species in three genera have been studied as key pollinators of the native flora. The paraphyletic *Prosoeca* currently has 37 described species with *Stenobasipteron* embedded within *Prosoeca* (chapter three). *Moegistorhynchus*, sister to the abovementioned clade, currently has four described species (Barraclough 2017).

Exaggerated proboscis lengths have only been reported in a relatively small proportion of southern African species, but can be found in *Prosoeca*, *Moegistorhynchus* and *Stenobasipteron*. Thus, these exaggerated traits are spread across the phylogenetic tree (chapter 3) and likely have multiple origins. Co-evolution has been suggested to be a major force in shaping exaggerated lengths of the proboscis in the flies, as well as the corolla tubes of the flowers that they visit (Anderson & Johnson 2008; Pauw *et al.* 2009). Selection for longer proboscides may cause the evolution of traits to be directional, evolving from one state to another but never or rarely in the opposite direction (Whittall & Hodges 2007; Klimov *et al.* 2017). The limited number of species with a long proboscis may suggest constraints (eg. developmental) on the origin of this trait state, or that functional costs or trade-offs between resource acquisition and foraging efficiency select against it. The diversity of proboscis lengths of morphospecies in southern African Nemestrininae represents an ideal study system to examine scaling relationships as well as the pattern of state transition in proboscis length and its effects on the rate of diversification.

We aimed to reconstruct the evolution of proboscis length, for the clade comprising *Moegistorhynchus, Prosoeca* and *Stenobasipteron,* using comparative phylogenetic methods and focusing on four aspects. (i) We predict that proboscis and body length evolution would be labile (ie. not phylogenetically conserved) because exaggerated traits are distributed across the phylogeny. (ii) We consequently expect long and short proboscides to have multiple independent origins in the sampled group and (iii) we also predict an overall bias in the direction of transitions from shorter to longer proboscides. (iv) Finally, we expect long proboscides to be associated with increased diversification rates.

METHODS

To reconstruct the evolution of proboscis length in a clade of southern African nemestrinids which includes *Prosoeca, Moegistorhynchus* and *Stenobasipteron* we utilized a pruned version of the majority-rule phylogenetic tree from the Bayesian Inference reconstruction reconstructed in chapter three. The phylogenetic tree was trimmed to include a single accession per morphospecies using the drop.tip function from the APE (Paradis *et al.* 2004) package in R (R Core Team 2017). Additionally, the outgroups were trimmed from the tree. The tree was then forced to bifurcate with the multi2di function and made ultrametric with the chronos function using the phytools package (Revell 2012).

Morphological traits and allometry

Proboscis and body length of all individuals in the phylogenetic tree, as well as other representatives from the same morphospecies, were measured (see Table S1 for sample sizes). Proboscis length was measured with a pair of digital callipers from the junction of the proboscis and the face to the tip of the proboscis, without extending the proboscis. Although

proboscides are able to extend during feeding (Morita 2011), we use the unextended length to allow for the inclusion of museum specimen measurements. Body length was measured from the frons to the end of the abdomen, excluding the genitalia. The mean value for traits of morphospecies was used in all analyses to account for variation from multiple individuals. To normalize the distribution of absolute proboscis and body length, the natural log was used throughout as a continuous trait (LaBarbera 1989). Relative proboscis length was calculated by dividing the mean proboscis length by mean body length for each morphospecies. To assess the allometric relationship between body length and proboscis length across the group (ie. evolutionary allometry), we used the morphospecies means and ran a Phylogenetic Generalised Least Squares (PGLS) regression using the pgls function with the Caper package (Orme 2013). PGLS corrects for phylogenetic relatedness to account for the statistical non-independence of closely related species. An ordinary linear regression (OLS) was also run with body length as the predictor and proboscis length as the response variable. Static allometry of 14 morphospecies which had measurements of proboscis and body length for more than 10 individuals was assessed with an OLS analysis. Differences in the static intercept between species indicates differences in the proportional size of the proboscis length, irrespective of body size, while the relationship between the two traits are maintained (Gould 1966). A difference in the static slope between species, however, indicates a difference in how the proportional size of the proboscis changes with the body size within a species (Gould 1966). Positive allometry with a static slope value > 1 indicates that with a unit increase in body size, there is a proportionally larger increase in the size of the relevant trait. A static slope = 1 indicates isometry and means that a unit increase in body size, is equal to a unit increase in the relevant trait.

Phylogenetic signal

To evaluate whether the phylogenetic relationships can explain the observed trait values, phylogenetic signal of traits was analysed by running 999 simulations of Pagel's lambda (Pagel 1999) and Blombergs K (Blomberg *et al.* 2003). We calculated lambda and K using the species mean data for proboscis length, body length and the relative proboscis length as continuous traits using the phylosig function from phytools. Blombergs K is defined as the ratio between the mean squared error of the observed traits divided by the mean squared error of the trait calculated using a variance-covariance matrix obtained from the given phylogeny and a Brownian motion model of trait evolution (Münkemüller *et al.* 2012). Pagel's λ is a tree transformation approach that assesses the degree of fit of trait data to a Brownian motion model with values ranging between 1 and 0. When λ and $K \geq 1$, the phylogenetic relationships can explain the observed trait values whereas values of λ and $K < 1$ indicate that the phylogenetic relationships alone cannot explain observed trait values.

Character coding

Ancestral state reconstruction was done for absolute and relative proboscis length, as they potentially have different evolutionary mechanisms: long proboscid flies interact with the floral tube of the flowers they visit via the absolute length of their proboscis. Therefore, the strength of interactions between flowers and nemestrinid pollinators likely affect the evolution of proboscis length. Relative proboscis length was also considered, as species with larger bodies will generally have larger morphological traits. Additionally, constraints or functional costs may be larger for small bodied species with a long absolute proboscis than a large bodied species with an equally long proboscis.

Absolute and relative proboscis length were both analysed as continuous traits, whereas absolute proboscis length was additionally reconstructed as a discrete trait. To allow

reconstruction of traits with a continuous distribution and thereby including all the subtle changes in length, we reconstruct proboscis lengths and the trait value of the most recent common ancestor as continuous trait. In addition, discrete trait reconstruction was used to evaluate the frequency and directionality of transitions between long and short proboscides. For discrete trait reconstruction, the frequency distribution of absolute proboscis length was used to identify cut-off values where a clear gap in the distribution is observed. An absence of species with mean proboscides of 14-16 mm and 23-29 mm, were observed and used as the two cut-off points to generate three discrete states (Fig. 4.2). Absolute proboscis length was considered short if it was shorter than 15 mm, long if it was between 15 mm to 29 mm and very long if it was longer than 29 mm. Modelling the evolution of proboscis length based on discrete character states may have disadvantages such as the loss of biological information contained in the dataset and the potential for subjective bias in defining character states (Parins-Fukuchi 2018). We, however, aimed at evaluating how often proboscis lengths undergo large changes, specifically to the exaggerated trait values and back to short values. Therefore, we use the cut off 15mm and 23mm to evaluate the frequency and directionality of transitions between short, long and very long proboscides by reconstructing the absolute proboscis length as a discrete trait.

Continuous trait evolution

To examine the proboscis length of the most recent common ancestor, nine models of continuous evolution available for the Geiger package (Pennell *et al.* 2014) in R were compared using weighted AIC. This included the commonly used Brownian motion (BM) and Ornstein-Ulhenbeck (OU) models. The trait value for the most recent common ancestor of the ingroup and its associated 95% confidence interval was estimated for absolute and relative proboscis length using the fastANC function from the phytools package (Revell

2012). This was done on 100 randomly selected trees from the posterior distribution of the Bayesian Inference analysis from chapter three, to account for phylogenetic uncertainty. To visualize the evolution of proboscis length as a continuous trait along the tree we used the contMap function from the phytools package (Revell 2012).

Discrete trait evolution

To assess the frequency and direction of transitions, the evolution of absolute proboscis length was analysed as a discrete multistate trait (ie. ≤ 15mm, short; 15 – 23mm, long and ≥ 23 mm, very long) using maximum-likelihood ancestral state reconstruction under three models. An equal-rates (ER) model, symmetrical rates (SYM) model and an all rate different (ARD) model were compared using the fitdiscrete function from the Geiger package (Pennell *et al.* 2014). The weighted AIC values of the three models were compared to select the most appropriate model for the tree topology. Character-state transition probabilities and rates were calculated using Bayesian stochastic character mapping on 100 phylogenetic trees with the make.simmap function in the phytools package (Revell 2012). The Bayesian stochastic mapping analysis was run for 100 simulation replicates, with the model that had the lowest AIC value.

Diversification rate

To test if there were shifts in the diversification rate in the clade, we first ran a Bayesian analysis of macroevolutionary mixtures (BAMM) analysis using BAMM 2.5.0 (Rabosky *et al.* 2013). The priors configured using BAMMtools (Rabosky *et al.* 2014) in R were used with an expected number of diversification rate shifts of one, which is the default. To test the relationship between log10 transformed absolute proboscis length and the net diversification

rate along the phylogeny, we used a Structured rate permutations on phylogenies (STRAPP) analysis with the traitdependentBAMM function from BAMMtools (Rabosky & Huang 2016).

RESULTS

Morphological traits and allometry

Proboscis lengths across the sampled species show an almost continuous distribution of lengths ranging from 4mm to 53mm (Fig. 4.2) while body length varied from 7mm to 20mm. Relative proboscis length likewise shows a continuous distribution with ratios of proboscis to body length ranging from 0.29 to 3.64 in our dataset (Fig. 4.2). There is a strong relationship between body length and proboscis length across morphospecies ($F_{1,56} = 15.25$, $R^2 = 0.20$, $p < 0.001$; Fig. 4.3). This relationship between body length and proboscis length is maintained after controlling for phylogenetic relationships ($F_{1,56} = 17.53$, $R^2 = 0.22$, $p < 0.001$). The slope of evolutionary allometry across the clade is close to isometry (1.08). When the static allometry of individual morphospecies is examined, the significant relationship between body length and proboscis length is only maintained for morphospecies with a long proboscis (Table 4.1, Fig. 4.4). *Prosoeca longipennis* and *P. peringueyi* were the only morphospecies that showed a static slope > 1 (positive allometry), whereas *P. torquata* had a slope close to isometry (Table 4.1). None of the morphospecies with a proboscis shorter than 17mm showed a significant relationship between body length and proboscis length (Table 4.1, Fig. 4.4).

Phylogenetic signal

The observed trait values for both proboscis length and body length cannot be explained by phylogenetic relationships under a model of Brownian motion evolution. Little to no

phylogenetic signal was detected for proboscis length (lambda = 0.54; K = 0.29), body length (lambda = 0.83; K = 0.52) or relative proboscis length (lambda = 0.46; K = 0.24). Body length showed the highest Pagel's lambda value, but the corresponding K value was well below the threshold of phylogenetic signal.

Continuous trait evolution

A lambda model provided the best fit (AICw = 0.643) of the seven models of absolute proboscis length evolution. Most of the remaining AIC weight was shared between the remaining models (OU, AICw = 0.097; delta, AICw = 0.071), whereas the BM model and the EB model received low support (AICw = 0.046 and 0.017, respectively). The proboscis length of the most recent common ancestor averaged over 100 trees was estimated to have been 11.88 mm, with the 95% confidence interval: 11.39 – 12.37 mm, which is only slightly longer than the median length of the extant species (median: 10.95mm).

A Lambda model also provided the best fit (AICw = 0.507) of the seven models of relative proboscis length. The majority of the remaining AIC weight was made up by the OU model (AICw = 0.140), while the EB and BM model each received little support (EB AICw = 0.970E-04; BM AICw = 0.003). The relative proboscis length of the most recent common ancestor averaged over the 100 trees was estimated to have been 1.08 times the length of the body (Fig 4.5), with 95% confidence interval: 1.05 – 1.12, longer than the median length of the current ingroup (median: 0.861 times the length of the body).

Discrete trait evolution

A comparison of the weighted AIC of the ER, SYM and ARD models revealed that the ER model fits best for the multistate analysis of absolute length (AICw = 0.826). The SYM

model (AICw = 0.151) had the second highest weighted AIC for the multistate analysis. Discrete trait estimation of absolute proboscis length on 100 phylogenetic trees suggests, with weak support, that the ancestral state of this group is a short proboscid morphospecies (0.92 PP) (Fig. S4.1). Stochastic mapping inferred changes between all three states with transitions away from short are most common (Table S4.2, Fig. 4.6).

Diversification rates

The 95% credible set of rate shift configurations sampled with BAMM included 14 distinct configurations. A model with no shifts in the diversification rate received the most support (PP = 0.917). The STRAPP analysis similarly showed that the evolution of proboscis length was not associated with changes in net diversification rates ($r = 0.019$, $p = 0.899$).

DISCUSSION

The sampled nemestrinids vary considerably in absolute proboscis length as well as relative proboscis length (ie. proboscis length over body length) (Fig. S4.2). Morphospecies exhibited proboscides ranging from less than half of the body length to more than three times the length of the body (Fig. S4.2). While proboscis length is correlated with body length across species, static allometry relationships are only present in long proboscid morphospecies and absent in short proboscid morphospecies. This suggests that different selective pressures may be acting on the traits of morphospecies with long and short proboscides. Both absolute and relative proboscis length are evolutionarily labile traits that are not phylogenetically conserved. The lack of phylogenetic signal indicates that the evolution of proboscis length and body length cannot be explained by the phylogenetic relationships among morphospecies. Morphospecies with long proboscides clearly do not form a monophyletic clade, illustrated by the multiple

sister morphospecies pairs exhibiting vastly different trait values (eg. *M. brevirostris*, proboscis mean = 10.71mm and *M. longirostris*, proboscis mean = 53.24mm). Multiple independent transitions from short to longer proboscides and *vice versa* occurred in the evolutionary history of this group. There is a pattern of more frequent transitions from short to longer proboscides, indicating directionality in absolute proboscis length evolution. We found no support for an association between proboscis length and the rate of diversification. Based on these results, we find that proboscis length is a labile trait.

The recurrent and reversible pattern of proboscis lengthening suggests that different ecological conditions select strongly for both long and short proboscides. Different proboscis lengths are likely advantageous under different environmental conditions and may explain why there is no association between proboscis length and diversification. All clades include long and short proboscid morphospecies except for clade B (Fig. S4.1) which only contains short proboscid morphospecies. Absolute proboscis length shows multiple transitions between very long, long and short proboscides in all directions with transitions away from longer lengths being rarer than those towards increased lengths. Bidirectional changes in the trait values of leg length and eye-stalk length have also been found in *Rediviva* and diopsid flies, respectively (Baker & Wilkinson 2001; Kahnt *et al.* 2017), suggesting that reversals away from exaggerated traits may be relatively common.

Phylogeny does not appear to explain proboscis length changes in this clade of nemestrinids unlike the phylogenetically conserved pattern seen in orchid bees (Ramírez *et al.* 2010) or in bumble bees (Kawakita *et al.* 2004). While the phylogeny on which these results are based may not be complete (chapter three), it is clear that a single origin of long proboscides can be rejected. A similar pattern of relative trait evolution is present in the oil-bee genus *Rediviva* which exhibits multiple independent origins of leg length longer than body length from a short legged ancestor (Kahnt *et al.* 2017).

We found a positive association between body length and proboscis length across the clade, as well as in the static allometry of morphospecies with long proboscides. There was, however, no significant association between body and proboscis length for morphospecies with shorter proboscides. Increases in proboscis length from a short ancestral morphospecies may be facilitated by a change in the slope or the intercept of the allometric relationship. Directional selection, such as that imposed by long-tubed flowers (Anderson & Johnson 2008; Pauw *et al.* 2009), is generally thought to favour steeper static slopes (Voje 2016), while stabilizing selection is thought to favour shallower slopes (Pélabon *et al.* 2011). Steeper static allometries may also evolve if there is a resource trade-off between proboscis length and body size (Bonduriansky & Day 2003). The different pathways by which proboscis length changes across the phylogeny may be the result of different selection pressures experienced by different species.

Both costs and benefits should be considered when interpreting selective forces acting on ecologically important traits such as proboscis length. Access to increased floral resources associated with long proboscides likely presents a foraging advantage over competitors (Martins & Johnson 2013; Johnson *et al.* 2017; Klumpers *et al.* 2019). However, long proboscides may decrease feeding efficiency by increasing handing times (Karolyi *et al.* 2013; Bauder *et al.* 2015), be expensive to maintain (Krenn 2019) or require specific ecological conditions to be energetically worthwhile. Like for stalk-eyed flies (Swallow *et al.* 2000), larger exaggerated traits, may also have a greater negative impact on aerial performance in nemestrinids. Thus, to ensure efficient flying and foraging, a long proboscis may additionally necessitate a tight allometric relationship between body size and proboscis length. Selection on size may also be indirect, if larval host size variation results in differences in absolute proboscis lengths, as is seen in acorn weevils (Bonal *et al.* 2011), or

may be the result of abiotic selection on body size as with many other groups of insects (Chown & Gaston 2010).

Our findings suggest that diversification proceeded gradually, without a significant shift in the diversification rate. This pattern is the opposite of what has been found for flower spur evolution in groups like Antirrhineae and *Aquilegia* (Whittall & Hodges 2007; Fernández-Mazuecos *et al.* 2019). In flower spurs, increased specialization of visiting pollinators is often associated with increased diversification rates (Armbruster & Muchhala 2009), while increased generalization of the nemestrinid feeding niche (with increasing proboscis length) appears to have no effect on diversification rate. Much like proboscis length in this study, the foreleg length in *Rediviva* bees (which can be extremely exaggerated) was also not associated with a change in diversification rate (Kahnt *et al.* 2017). The lack of association between proboscis length and diversification rate may be due to low statistical power as a result of the relatively low number of morphospecies included in the phylogeny (Rabosky & Huang 2016). A larger, more inclusive phylogeny, potentially including other families or orders of insects, with a quantitative analysis may provide a more suitable basis on which to examine the association between diversification rate and proboscis length. Additionally, an analysis inclusive of a measure of error for trait means may provide a more accurate representation of trait evolution (Martins & Hansen 1997) as some morphospecies show considerable variation around the mean, especially in proboscis length. Further studies examining the mechanisms behind changes in allometry and body size will aid our understanding of the development of these exaggerated traits.

ACKNOWLEDGMENTS

We thank Dr. Iliana Medina for help with the analysis of continuous trait evolution. We also thank the KwaZulu-Natal museum for access to their Nemestrinidae collection for additional measurements. This study was funded by National Research Foundation of South Africa – Foundational Biodiversity Programme (NRF-FBIP grant # 110440 to A.G.E), National Research Foundation of South Africa (SARChI grant # 46372 to S.D.J).

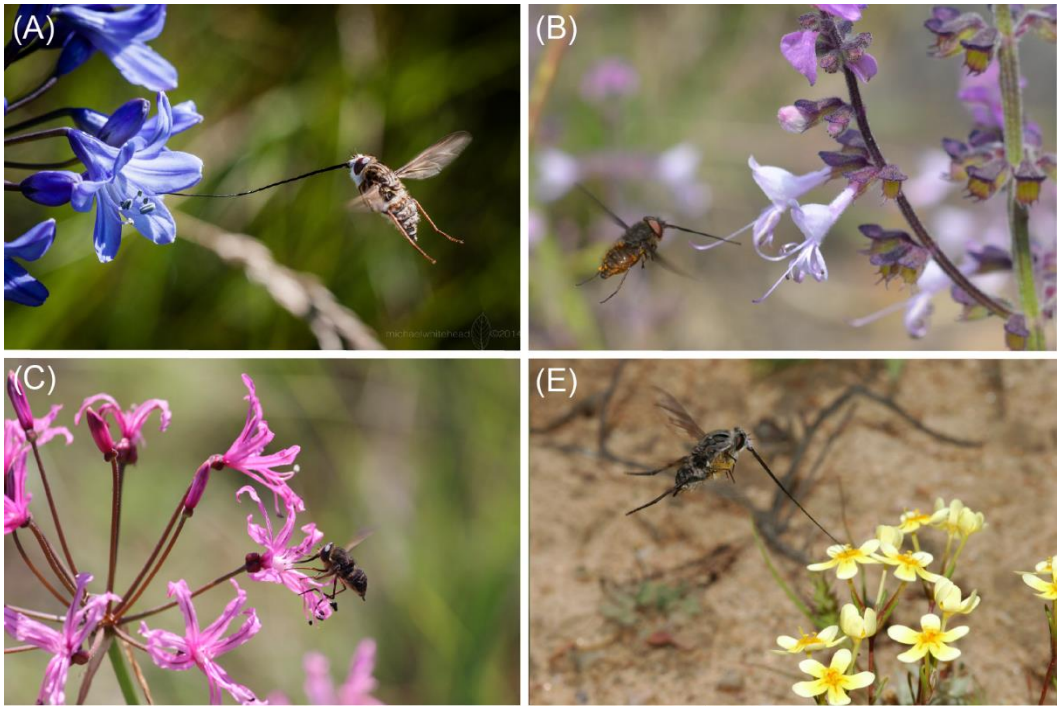


Figure 4.1 Photographs of *in situ* adult *Prosoeca* species visiting flowers. (A) *Prosoeca ganglbaueri* visiting an *Agapanthus sp.* (B) *Prosoeca sp. 6* visiting a Lamiaceae sp. (C) *Prosoeca umbrosa* visiting *Nerine angustifolia* (D) *Prosoeca peringueyi* visiting a *Zaluzianskya sp.* Photo credits: (A) Michael Whitehead, (B) Ruth Cozien, (C) Genevieve Theron, (D) Steven Johnson.

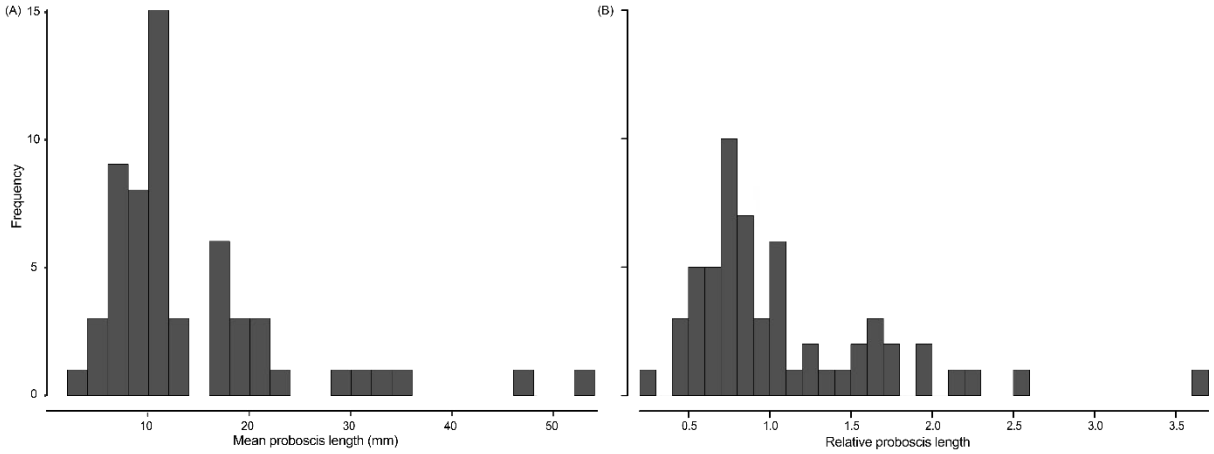


Figure 4.2 Frequency distribution of morphospecies (A) mean proboscis length (mm) and (B) relative proboscis length.

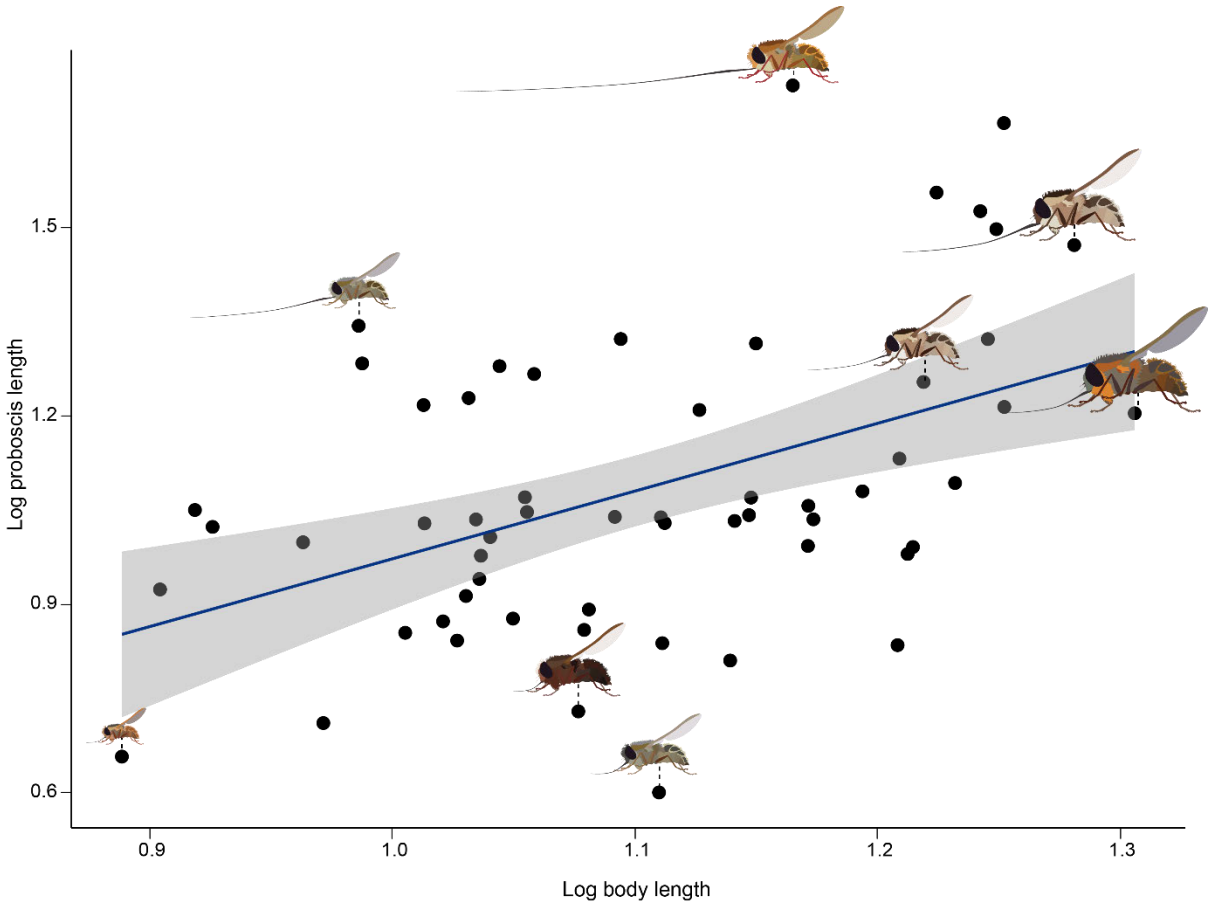


Figure 4.3 Regression of mean body length and mean proboscis length. Grey shading indicates SE. Fly illustrations are to scale relative to each other and are connected to the morphospecies that they represent on the graph by a dotted line.

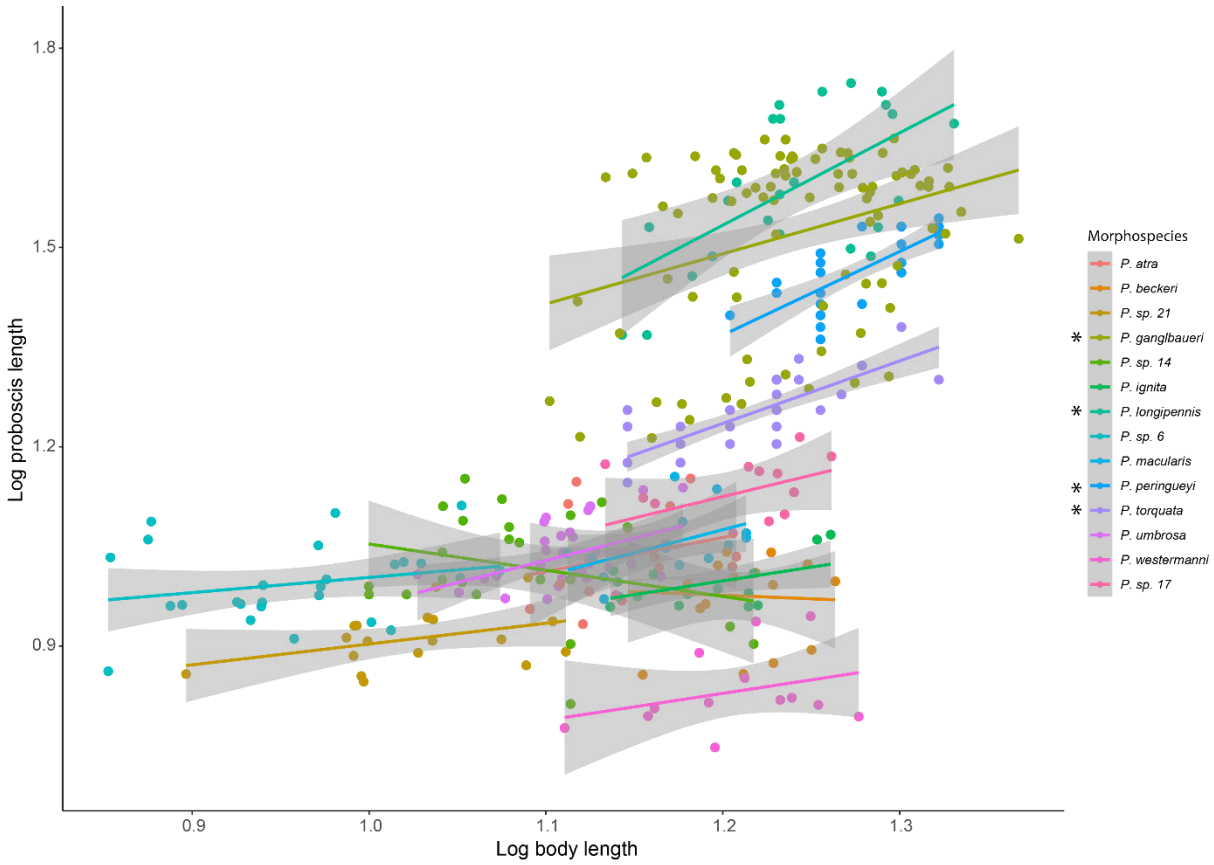


Figure 4.4 Regressions of log10 body length and log10 proboscis length for 14 morphospecies. Grey shading indicates SE. Asterisks on the figure legend indicate a significate relationship between body and proboscis length for the relevant morphospecies.

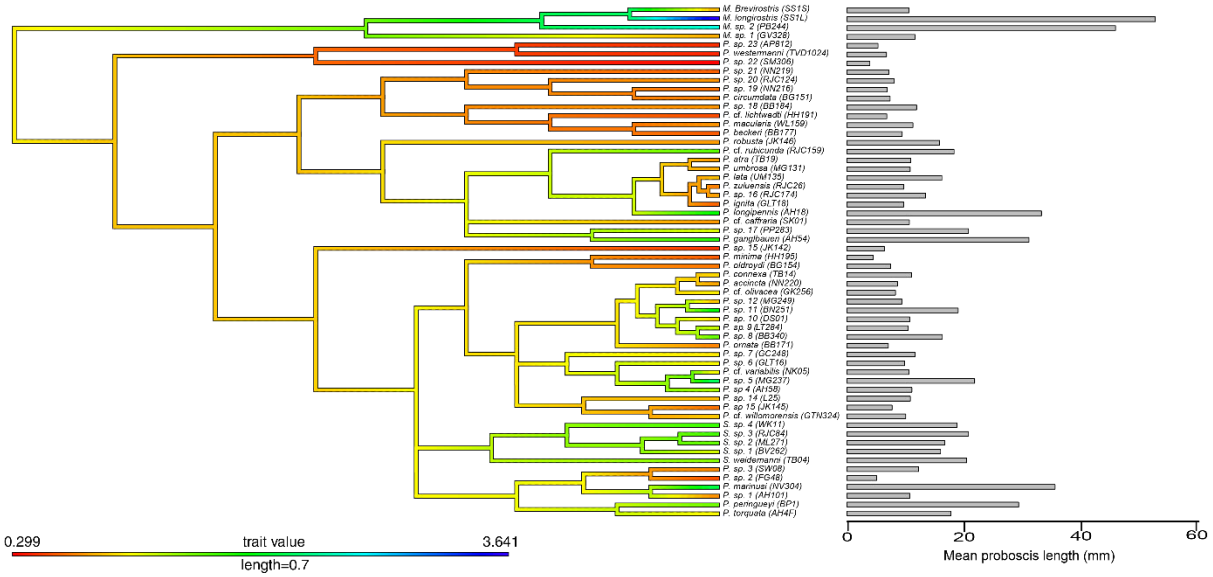


Figure 4.5 Continuous character evolution mapping of relative proboscis length with mean absolute proboscis length indicated in grey bars.

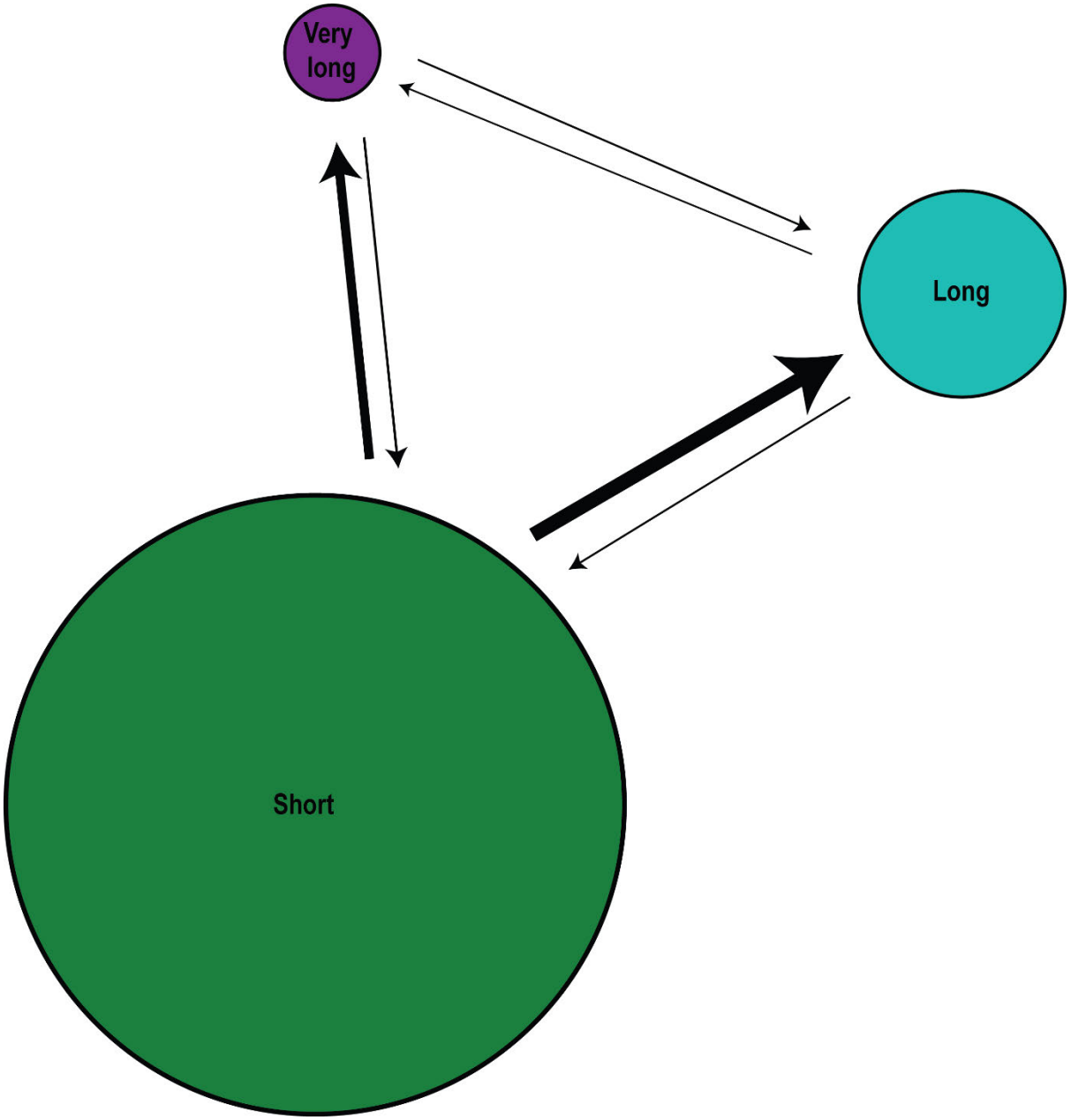


Figure 4.6 Evolutionary proboscis state transitions within the nemestrinid clade containing *Moegistorhynchus*, *Prosoeca* and *Stenobasipteron*. The size of the circles is proportional to total number of species sampled within each state: short, 39; long, 13; very long, 6. Arrow thickness is proportional to the number of transitions in each direction averaged over 100 phylogenetic trees.

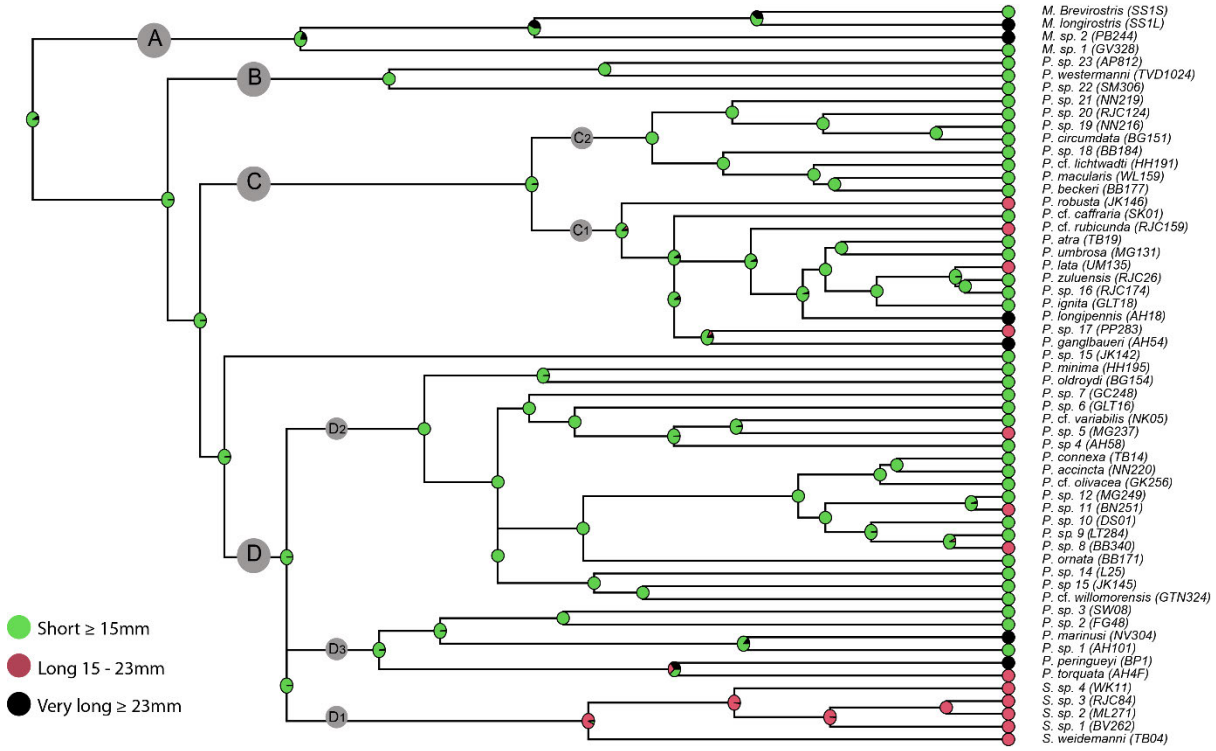


Figure S4.1 Stochastic character mapping of absolute proboscis length with three character states.

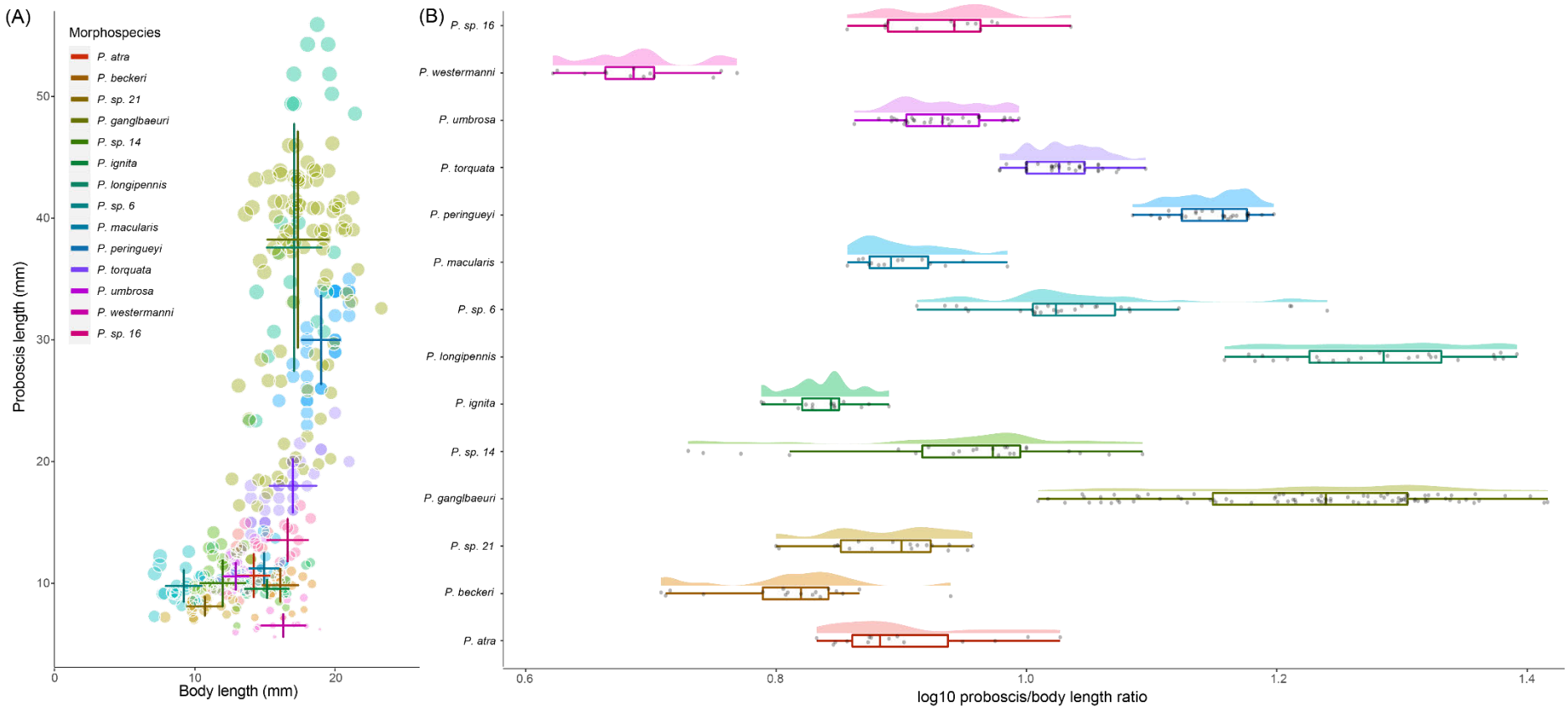


Figure S4.2 Variation in morphological traits measured for the 14 morphospecies for which allometry was assessed. (A) Scatterplot of proboscis length and body length of all measured individuals. Median and standard deviation indicated with dark vertical and horizontal bars. (B) Raincloud plot indicating the distribution of specimen data for log10 of relative proboscis length.

Table 4.1 Morphological measurements and linear regression statistics of morphospecies used to examine static allometry.

Morphospecies	Mean proboscis length Mean ± SE	Mean body length Mean ± SE	Intercept	Slope	R value	P value
Prosoeca atra	11.02 (0.49)	14.03 (0.31)	0.4489	0.5121	0.01	0.355
Prosoeca beckeri	9.56 (0.34)	16.31 (0.30)	1.1054	-0.1073	0.06	0.842
Prosoeca ganglbaueri	31.44 (1.02)	17.74 (0.26)	0.5857	0.7540	0.09	0.00229
Prosoeca ignita	9.81 (0.21)	16.39 (0.40)	0.5030	0.4126	0.20	0.0538
Prosoeca longipennis	33.58 (1.32)	17.47 (0.41)	-0.1332	1.3890	0.34	0.00159
Prosoeca macularis	11.41 (0.35)	14.84 (0.30)	0.2446	0.6921	0.16	0.0857
Prosoeca peringueyi	29.66 (0.69)	19.10 (0.26)	-0.1418	1.2582	0.52	5.9e-06
Prosoeca torquata	17.96 (0.38)	16.56 (0.29)	1.1182	1.0170	0.59	5.48e-08
Prosoeca umbrosa	10.94 (0.20)	12.90 (0.16)	0.2931	0.6689	0.19	0.00643
Prosoeca westermanni	6.85 (0.27)	16.15 (0.43)	0.3417	0.4059	0.02	0.296
Prosoeca sp. 6	9.98 (0.26)	9.19 (0.26)	0.7747	0.2286	0.02	0.236815
Prosoeca sp. 14	10.95 (0.46)	12.35 (0.39)	1.4483	-0.3946	0.22	0.220861
Prosoeca sp. 16	13.56 (0.50)	16.18 (0.40)	0.3528	0.6434	0.13	0.127
Prosoeca sp. 20	8.20 (0.19)	10.72 (0.31)	0.5900	0.3131	0.10	0.10953

Table S4.1 Localities, specimen details and morphological measurements of accessions used for the phylogenetic analyses.

Sequence ID	Morphospecies	Locality	Latitude	Longitude	Proboscis length			Body			Relative proboscis length
					Mean ± SE	Range	N	Mean ± SE	Range	N	
SS1S	*Moegistorhynchus brevirostris*	Silverstroomstrand	-33.528550	18.446101	10.71		1	12.95		1	0.83
SS1L	*Moegistorhynchus longirostris*	Silverstroomstrand	-33.528550	18.446101	53.24 (8.28)	32.06-80.78	5	14.62 (0.46)	13.31-16.2	5	3.64
GV328	*Moegistorhynchus sp. 1*	Groenvlei	-30.319344	18.067708	11.76		1	14.06		1	0.84
PB244	*Moegistorhynchus sp.2*	Picketberg	-32.79855	18.665922	46.38		1	17.87		1	2.60
NN220	*Prosoeca accincta*	Naudes Neck Pass	-30.732387	28.138476	8.73 (0.21)	7.73-9.14	6	10.86 (0.39)	9.74-12.61	7	0.80
TB19	*Prosoeca atra*	Mbotyi	-31.465393	29.730393	11.02 (0.49)	8.57-14.2	14	14.03 (0.31)	12.33-16.13	15	0.79
BB177	*Prosoeca beckeri*	Boosmansbos	-33.929018	20.86353	9 56 (0.34)	7.19-11.94	17	16.31 (0.30)	14.01-19.02	19	0.59
SK01	*Prosoeca* cf. *caffraria*	Vernon Crooke Nature Reserve	-30.274244	30.594822	10.79 (0.61)	9.66-12.96	5	13.84 (1.0)	11.74-16.52	5	0.78
HH191	*Prosoeca* cf. *lichtwardti*	Hottentots Holland Nature Reserve	-34.069298	19.047765	6 90 (0.56)	6.35-7.45	2	12.92 (0.26)	12.66-13.18	2	0.53
GK256	*Prosoeca* cf. *olivacea*	Graskop	-24.932787	30.808752	8.40 (0.41)	6.77-10.19	8	8.02 (0.22)	7.07-8.82	8	1.05
RJC159	*Prosoeca* cf. *rubicunda*	Swartberg	33.400	22.355	18,49 (1.39)	17.1-19.88	2	11,44 (1.19)	10.25-12.63	2	1.62
NK05	*Prosoeca* cf. *variabilis*	Nkandla Forest Reserve	-28.743981	31.139830	10.70 (0.53)	9.25-13.01	8	10.31 (0.44)	8.64-12.57	8	1.04
GTN324	*Prosoeca* cf. *willomoresis*	Grahamstown	-33.311184	26.527556	10.17 (0.50)	9.25-11.45	4	10.97 (0.22)	10.52-11.56	4	0.93

BG151	*Prosoeca circumdata*	PMB Botanical Gardens	-29.607538	30.34779	7.47 (0.75)	5.92-10.03	5	10.49 (0.36)	9.68-11.5	5	0.71
TB14	*Prosoeca connexa*	Mbotyi	-31.465391	29.730391	11.15 (0.71)	10.24-13.25	4	11.36 (1.15)	8.11-13.47	4	0.98
AH54	*Prosoeca ganglbaueri*	Naudes Nek	-30.77598	28.22365	31.44 (1.02)	14.25-46.15	100	17.74 (0.26)	12.65-23.3	75	1.77
GLT18	*Prosoeca ignita*	Haenertsberg	-23.880965	29.99877	9.81 (0.21)	9.1-11.69	15	16.39 (0.40)	13.69-18.24	16	0.60
UM135	*Prosoeca lata*	Umtamvuna Nature Reserve	-31.00363	30.17356	16.39 (1.35)	15-19.08	3	17.88 (1.35)	15.33-20	3	0.92
AH18	*Prosoeca longipennis*	Riversdale	-34.036519	21.455702	33.58 (1.32)	17.26-56.57	80	17.47 (0.41)	13.90-21.42	24	1.92
WL159	*Prosoeca macularis*	Waylands Nature Reserve	-33.408801	18.413050	11.41 (0.35)	9.35-14.29	14	14.84 (0.30)	12.95-16.34	14	0.77
NV304	*Prosoeca marinusi*	Niewoutville Botanical Garderns	-31.398152	19.141070	35.95 (1.33)	31.59-41.72	7	16.76 (0.55)	14.68-19.1	7	2.14
HH195	*Prosoeca minima*	Hottentots Holland Nature Reserve	-34.069298	19.047765	4 55 (0.32)	3.13-6.43	9	7.73 (0.36)	5.87-9.68	9	0.59
BG154	*Prosoeca oldroydi*	PMB Botanical Gardens	-29.607538	30.34779	7 55 (0.29)	7.16-8.42	4	11.21 (0.23)	10.54-11.56	4	0.67
BB171	*Prosoeca ornata*	Boosmansbos	-33.929018	20.86353	7 29 (0.28)	6.28-7.80	5	10.12 (0.30)	9.05-10.83	5	0.71
BP1	*Prosoeca peringueyi*	Botterkloof	-31.823601	19.256734	29.66 (0.69)	23.00-35.00	29	19.10 (0.26)	16.00-21.00	29	1.55
JK146	*Prosoeca robusta*	Jonaskop	-33.957261	19.520465	16.01 (0.96)	13.18-17.27	4	20.23 (0.71)	18.88-22.22	4	0.79
AH4F	*Prosoeca torquata*	Ariehoek	-30.244758	18.050258	17.96 (0.38)	14.00-24.00	35	16.56 (0.29)	14.00-21.00	35	1.08
MG131	*Prosoeca umbrosa*	Mount Gilboa	-29.308598	30.309024	10.94 (0.20)	9.35-13.76	33	12.90 (0.16)	10.65-15.05	34	0.85
TVD1024	*Prosoeca westermanni*	Prince Alfred Hamlet	-33.236666	19.55	6.85 (0.27)	5.59-8.81	13	16.15 (0.43)	12.9-19.92	14	0.42
RJC26	*Prosoeca zuluensis*	Bisley Nature Reserve	-29.658909	30.387038	9.85 (0.28)	9.07-11.33	7	14,84 (0.44)	13.7-16.53	7	0.66
AH101	*Prosoeca sp. 1*	Nieuwoudtville	-31.387089	19.176518	10.86 (0.38)	9.22-12.39	7	14.91 (0.48)	13.22-16.38	7	0.73

FG48	*Prosoeca sp. 2*	Namaqualand	-30.371906	18.092013	5 15 (0.59)	4.55-5.74	2	9.37 (0.55)	8.82-9.91	2	0.55
SW08	*Prosoeca sp. 3*	Soetwater			12.40		1	17.06		1	0.73
AH58	*Prosoeca sp. 4*	Naudes Nek	-30.73242	28.14157	11.23 (0.43)	10.03-13.1	7	8.29 (0.25)	6.97-9.27	8	1.35
MG237	*Prosoeca sp. 5*	Mount Gilboa	-29.276283	30.281003	22.06 (1.31)	19.66-25.60	4	9.68 (0.58)	8.15-11.6	5	2.28
GLT16	*Prosoeca sp. 6*	Haenertsberg	-23.880965	29.99877	9 98 (0.26)	7.28-12.93	27	9.19 (0.26)	7.12-11.86	26	1.09
GC248	*Prosoeca sp. 7*	Giants Castle Nature Reserve	-29.276286	30.281006	11.77		1	11.34		1	1.04
BB340	*Prosoeca sp. 8*	Bulembo border post	-25.943825	31.106867	16.50		1	10.3		1	1.60
LT284	*Prosoeca sp. 9*	Long Tom Pass	-25.149548	30.619543	10.56		1	8.43		1	1.25
DS01	*Prosoeca sp. 10*	Dullstroom	-25.396739	30.126015	10.85 (0.77)	10.08-11.62	2	10.83 (0.83)	10-11.65	2	1.00
BN251	*Prosoeca sp. 11*	Bushmans Nek Pass	-29.843928	29.210244	19.22 (0.98)	17.75-21.08	3	9.72 (0.50)	8.89-10.63	3	1.98
MG249	*Prosoeca sp. 12*	Mount Gilboa	-29.278267	30.287912	9 50 (0.22)	8.82-10.50	8	10.88 (0.25)	10.08-12.27	8	0.87
JK145	*Prosoeca sp. 13*	Jonaskop	-33.968685	19.518869	7.80		1	12.05		1	0.65
L25	*Prosoeca sp. 14*	Langkloof	-30.55937	18.128272	10.95 (0.46)	8.00-14.19	17	12.35 (0.39)	9.99-16.50	20	0.89
JK142	*Prosoeca sp. 15*	Jonaskop	-33.957259	19.520465	6.47 (1.07)	5.4-7.54	2	13.78 (1.13)	12.65-14.91	2	0.47
RJC174	*Prosoeca sp. 16*	Eshowe	28.8910833	31.4349	13,56 (0.50)	9.47-16.39	13	16.18 (0.40)	13.6-18.26	14	0.84
PP283	*Prosoeca sp. 17*	Paardeplaats Nature Reserve	-25.095679	30.562498	21.02 (0.83)	17.85-22.99	6	17.60 (0.72)	14.75-19.41	6	1.19
BB184	*Prosoeca sp. 18*	Boosmansbos	-33.929018	20.86353	12.03 (0.23)	11.22-12.52	5	15.62 (0.58)	13.37-18.08	7	0.77
NN216	*Prosoeca sp. 19*	Naudes Neck Pass	-30.736268	28.13939	6 96 (0.27)	6.32-7.92	6	10.63 (0.46)	9.14-11.99	6	0.65

RJC124	*Prosoeca sp. 20*	Saddleback Pass	-25.79536	31.094741	8 20 (0.19)	7.02-10.07	18	10.72 (0.31)	7.88-12.92	18	0.76
NN219	*Prosoeca sp. 21*	Naudes Neck Pass	-30.736268	28.13939	7 24 (0.36)	6.04-8.52	6	12.00 (0.28)	11.06-12.76	5	0.60
SM306	*Prosoeca sp. 21*	Silvermine Nature Reserve	-34.085932	18.418983	3 99 (0.14)	3.85-4.13	2	12.67 (0.22)	12.45-12.88	2	0.30
AP812	*Prosoeca sp. 23*	Cape Peninsula	-34.354020	18.493001	5 37		1	11.93		1	0.45
BV262	*Stenobasipteron sp. 1*	Bridal Veil Falls	-25.082029	30.723459	16.21 (0.41)	15.39-16.68	3	13.38 (0.23)	12.93-13.71	3	1.21
ML271	*Stenobasipteron sp. 2*	Mountainlands Nature Reserve	-25.7735	31.08543	16.93 (1.15)	14.83-20.73	5	10.75 (0.58)	8.92-12.45	5	1.58
RJC84	*Stenobasipteron sp. 3*	Abel Erasmus Pass	-24.55128	30.76479	21.02		1	12.42		1	1.69
WK11	*Stenobasipteron sp. 4*	Serala Forest Reserve	-24.055322	30.005245	19.03		1	11.07		1	1.72
TB04	*Stenobasipteron weidermanni*	Grahamstown	-33.328778	26.500194	20.68 (0.65)	19.24-23.69	6	14.12 (0.29)	13.2-15.27	6	1.46

Table S4.2 Number of transitions between states of proboscis length using a discrete character state evolution analysis on 100 phylogenetic trees.

	Number of morphospecies (n)	Short	Long	Very long
Short (<15mm)	39		1.25	1.53
Long (15-23mm)	13	9.25		1.23
Very long (>23mm)	6	5.8	1.06	

Chapter 5

Repeated biome shifts following a Miocene Fynbos origin for southern African tangle-vein flies in the Nemestrininae subfamily (Nemestrinidae)

Genevieve L. Theron, Bruce C. Anderson, Allan G. Ellis, Steven D. Johnson & Timotheüs van der Niet

To be submitted to: Molecular Phylogenetics and Evolution

ABSTRACT

While transitions between broad ecological niches, such as biomes, can trigger speciation events directly, the prevalence of niche conservatism may result in diversification occurring largely within broad niches. Studying the history of biome occupancy in lineages can thus shed light on the role of major ecological shifts in radiations. South Africa is ecologically diverse, containing nine biomes distributed within winter and summer rainfall regions, as well as three global biodiversity hotspots. Here we investigate the timing of diversification events and reconstruct biome shifts across a clade of pollinating long-tongued flies (Nemestrinidae), comprising of *Moegistorhynchus*, *Stenobasipteron* and *Prosoeca*, that occur in seven of the nine biomes of South Africa. Using COI substitution rates across a phylogeny of 58 morphospecies, we reconstruct a Miocene (19 MYA) origin for the root of this clade. We use the dated phylogeny in combination with reconstruction of biome shifts to infer a Fynbos origin for the clade, with successive, bidirectional transitions between Fynbos and Grassland. Shifts from these biomes into Forest and Succulent Karoo were also present, albeit less frequent. Furthermore, multiple but relatively few independent colonisations of biomes were reconstructed, whereas the majority of speciation occurred within biomes rather than via shifts between biomes. Our results mirror those of similar biogeographical patterns of some groups of plants and insects and suggests a shared response to biotic or abiotic factors of the biome. Thus, as is the case for plants, our results indicate that the region in which the Fynbos biome is located was an historical repository of insect diversity from which colonisation of other biomes took place.

Keywords: biogeography, biome transitions, molecular dating, Nemestrinidae, speciation

INTRODUCTION

Biomes can be viewed as representing broad ecological niches based on differences in plant community composition and structure, as well as climatic variables (Rutherford *et al.* 2006). Niche conservatism, the tendency of species within a lineage to inhabit similar ecological niches as they diversify (Chapple & Keogh 2004; Crisp *et al.* 2009; Morinière *et al.* 2016; Laver *et al.* 2017), is frequent at the level of biomes (Crisp *et al.* 2009). This is likely due to the substantial ecological differences that exist between biomes, which may require profound physiological and morphological adaptation, inhibiting biome shifts, at least in plants (Simon *et al.* 2009; Donoghue & Edwards 2014). It is thought that evolutionary biome switching may require some degree of pre-adaptation which facilitates the colonisation of new biomes (Crisp & Cook 2012; Stock & Verboom 2012; Edwards & Donoghue 2013). Inherited physiological traits such as climate tolerances may thus limit species' geographic range and ability to adapt to new environments. Therefore, studying and understanding the patterns and processes associated with transitions in biomes are key to explaining local species diversity (Gosz 1992).

Indeed, speciation driven by ecological shifts within a particular biome appears to be much more frequent than speciation driven by ecological shifts among biomes, at least in plants, suggesting that biome conservatism is common (Crisp *et al.* 2009). However, once transitions between biomes occur, lineages may be introduced to novel ecological opportunities that promote adaptive radiation (Yoder *et al.* 2010; Pirie *et al.* 2019). In some clades, however, diversification appears to be linked to repeated biome transitions rather than to within-biome radiations (Mitchell *et al.* 2014; Cardillo *et al.* 2017). A recent meta-analysis suggested that methodological issues associated with the use of single biome occupancy (ie. monomorphic coding) may have caused the underestimation of biome transitions compared to when species were allowed to occupy multiple biomes (Dale *et al.* 2021). Additionally,

case studies using well resolved, densely sampled phylogenies suggest that these repeated transitions are more common than once thought (Holstein *et al.* 2011; Schmerler *et al.* 2012; Cardillo *et al.* 2017; Smissen & Rowe 2018). Studies on plants have shaped ideas about the relative importance of biome transitions for understanding diversification. This is perhaps because biomes are often defined on the basis of vegetation structure or composition, while very little is known about evolutionary biome transitions in animals such as insects (but see Kim & Farrell, 2015; Breeschoten *et al.*, 2016; de Jager & Ellis, 2017).

Although biome definitions are often based on plants, the direct or indirect association of the majority of insects with plants during some stage of their life cycle may suggest that biomes are also suitable geographic units for understanding spatial patterns of insect distribution (Siemann *et al.* 1998). Most insect species and higher-level taxa are limited to particular biomes due to physiological limitations and/or presence of obligate mutualistic interaction partners (Wiens & Graham 2005; Waterman *et al.* 2011; Duffy & Johnson 2017). Thus, studying the occupancy of ecologically distinct geographic units such as biomes in a phylogenetic context provides a useful framework to understand the relative importance of changes in ecological niches for insect diversification. This framework can also provide insights into the factors driving diversification and the rates of ecological adaptation to novel environments through evolutionary time (Donoghue & Edwards 2014).

The biota of South Africa is highly diverse; they are distributed across nine biomes (Mucina & Rutherford 2006) encompassing regions of mainly winter- or summer-rainfall (Tyson & Preston-Whyte 2000) and three biodiversity hotspots (Hrdina & Romportl 2017). Distinct biomes have been identified for South Africa based on floristic composition and climatic variables. These biomes range from Forests with complete canopy cover, to xeric Succulent Karoo with sparse, low growing vegetation (Mucina & Rutherford 2006). All nine biomes have undergone much spatial and compositional change over geological history, with the

currently recognizable biomes originating between the Cenozoic and Pleistocene (Nilsson *et al.* 1996; Scott *et al.* 1997; Verboom *et al.* 2009; Bond 2015). While much of South Africa is thought to have once been host to forests and woodlands, closed-canopy forests, up until the Miocene, are now mostly small, isolated remnant patches, embedded in a matrix of grass or shrub dominated biomes (Nilsson *et al.* 1996; Mucina & Rutherford 2006; Bond 2015).

Nemestrinidae, a small, early diverging dipteran family centered in the Mediterranean, eastern Australia, southern Africa, Chile and Argentina (Bernardi 1973), occurs in the majority of South African biomes. Southern Africa houses *ca.* 44 (although chapter three suggests that species diversity may be underestimated by half) of the *ca.* 277 described nemestrinid species worldwide (Barraclough 2006, 2017). All extant adult Nemestrinidae are thought to be important flower visitors, although this has only been shown in Chile and southern Africa (Goldblatt *et al.* 2000; Devoto & Medan 2006). Their other life stages are, however, less well understood in the southern African genera, but it is assumed that all Nemestrinidae have parasitic larvae (Bernardi 1973).

Nemestrinidae is an ancient family of flies, with many fossils found from the Jurassic (Ansorge & Mostovski 2000). Due to the easily recognizable diagonal wing vein, a substantial number of nemestrinid fossils has been identified. The family is thought to have originated in the Triassic, and *ca.* 26 species of nemestrinid fossils have been described (Bernardi 1973; Mostovski 1998; Mostovski & Martínez-Delclòs 2000; Wedmann 2007; Liu & Huang 2019a, 2019b; Wedmann *et al.* 2021). While southern Africa represents a hotspot of nemestrinid diversity (Bernardi 1973), no southern African nemestrinid fossils have been identified yet.

Prosoeca, Stenobasipteron and *Moegistorhynchus* form a clade of closely related southern African genera in the Nemestrinidae. *Moegisorhynchus* currently has four described species and is limited to the western region (winter-rainfall area) of South Africa, while

Stenobasipteron, with three species, occurs in the eastern part of the South Africa (summer-rainfall) and parts of Zimbabwe. *Prosoeca* is the largest of the three genera with 37 described taxa. The genus is widely distributed throughout southern Africa including Zimbabwe and Namibia (Barraclough 2017). The taxonomy of these three genera requires substantial revision, as multiple undescribed species have been attributed to all three of these genera. Additionally, *Moegistorhynchus* was found to be sister to a paraphyletic *Prosoeca*, with *Stenobasipteron* being nested within *Prosoeca* (chapter three). Collectively, this clade likely occurs in almost all of South Africa's nine biomes (Barraclough 2006). While *Stenobasipteron weidermanni* is the only nemestrinid species in this clade thought to be truly Forest dwelling (Potgieter & Edwards 2005), many other species are associated with Forest ecotones. Species from the southern African clade comprising *Moegistorhynchus*, *Prosoeca* and *Stenobasipteron* are distributed across multiple biomes and provide an ideal system for examining patterns of dispersal and speciation over evolutionary time, in association with ecologically distinct biomes.

Stenobasipteron and *Moegistorhynchus* have no known fossils, but two fossil species have been designated, albeit with doubt, as closely related to *Prosoeca* in *Palembolus* (Bequaert & Carpenter 1936; Mostovski & Martínez-Delclòs 2000). The oldest of the two fossils, *Prosoeca* (*Palembolus*) *saxea*, is from the Lower Cretaceous, suggesting an origin of at least 125 million years ago (MYA) for *Palembolus* (Mostovski & Martínez-Delclòs 2000). The two *Prosoeca (Palembolus)* fossils were found in Florissant, USA (Bequaert & Carpenter 1936) and near Cuenca, Spain (Mostovski & Martínez-Delclòs 2000), outside the current, southern African, distribution of extant *Prosoeca* species (Barraclough 2017). *Palembolus* was considered a subgenus of *Prosoeca* by Bequaert & Carpenter (1936), but Bernardi (1973) rather considered it a separate genus. Bernardi (1973) suggested that the uninformative wing venation characters and the disjunct distribution of *Prosoeca* and *Palembolus* cast doubt on

the subgenus classification, as extant *Prosoeca* are an exclusively southern African radiation. It is therefore unclear if the southern African *Prosoeca* originated as far back as the fossils suggest or if the extant genus represents an independent and/or younger lineage from these fossils.

The aim of this study was to evaluate the timing of diversification events and to reconstruct biome shifts for the clade comprising of *Moegistorhynchus*, *Prosoeca* and *Stenobasipteron*. We ask the following questions: i) When did this clade of Nemestrinidae originate? ii) In which South African biome did the most recent common ancestor of this clade originate? iii) Were individual biomes colonised once or were there multiple transitions to each biome? iv) Lastly, we ask whether species diversity is primarily the product of shifts between biomes or within biome diversifications?

METHODS

Taxon sampling

A total of 58 morphospecies of *Prosoeca*, *Stenobasipteron* and *Moegistorhynchus* were sampled widely across the different biomes between January 2014 and February 2020 from the known southern African range (Fig. 5.1, Table S5.1). A total of 49 *Prosoeca*, four *Moegistorhynchus* and five *Stenobasipteron* morphospecies were sampled (chapter three).

Divergence time estimation

To estimate the dates of diversification in this clade of Nemestrinidae, we made use of COI substitution rates, since fossil calibration could not be reliably implemented in this analysis. Although there are many fossil nemestrinid specimens, the majority of these represent lineages that are not adequately sampled in our phylogenetic analysis. Furthermore, it is

currently unclear whether *Prosoeca* (*Palembolus*) fossils should be considered a subgenus of *Prosoeca* or as a separate, extinct genus *Palembolus* (Bequaert & Carpenter 1936; Bernardi 1973). Incorrect fossil placement within the phylogeny has been suggested to lead to large errors in divergence date estimation (Lee 1999; Phillips *et al.* 2009). Additionally, the distinction between crown and stem placement of a particular fossil is crucial for accurate calibration (Doyle & Donoghue 1993). Therefore, given that phylogenetic relationships are currently poorly understood within Nemestrinidae (chapter three) and the uncertainty surrounding the placement of the *Prosoeca* (*Palembolus*) fossils, within or outside of this clade, we decided to not use fossil calibration for this analysis. Instead of fossil calibration, we thus make use of a molecular clock to date the phylogeny. The COI alignment with 58 accessions of 1401 base pairs (bp) was used for molecular dating as it is the only partition for which substitution rates are reasonably well-established in insects (Papadopoulou *et al.* 2010). PartitionFinder (Lanfear *et al.* 2017) was used on the CIPRES Science Gateway v3.3 (Miller *et al.* 2010) to identify the best partitioning scheme and substitution model for COI. We used Beauti v. 1.10.4 to create an XML file for BEAST v. 1.10.4 (Drummond *et al.* 2012) for divergence time estimation, implementing a Markov chain Monte Carlo (MCMC) of 100 million generations, sampling trees and parameters every 10,000 generations with a burn-in of 25%. We used a strict clock model and a Yule pure-birth speciation model with a defined starting tree (Drummond *et al.* 2006). The trimmed Majority-rule tree with a single accession per morphospecies from the Bayesian phylogenetic analysis of chapter three was used as the starting tree and this topology was constrained throughout the analysis. Prior to the dating analysis, all outgroups were removed, retaining 58 tips with *Prosoeca*, *Stenobasipteron* and *Moegistorhynchus* morphospecies. Previously determined COI substitution rates for insects were used to calibrate the constrained topology with a mean rate of 0.0177 substitutions per million years (Papadopoulou *et al.* 2010). A normal prior was set with a standard deviation of

35%, to account for both higher and lower substitution rates for mtDNA reported among insect taxa (Brower 1994; Clarke *et al.* 2001). We inspected the results using Tracer (Rambaut *et al.* 2018) to confirm that the effective sample size (ESS) was >100 for all parameters, and used TreeAnotator to summarize a maximum clade credibility tree (Drummond *et al.* 2012).

Biome allocation

To assign morphospecies to geographical areas, we used the South African biomes as designated by Mucina & Rutherford (2006). These nine biomes are fine-scale-resolution, ecologically meaningful, allopatric geographic units of interest for this clade of nemestrinids. Adult Nemestrinidae show tight associations with the plants they feed on as they are dependent on nectar for energy (Manning & Goldblatt 2000; Barraclough 2006), suggesting a likely association with biomes. While the adult stage represents only one portion of the nemestrinid life cycle and may or may not be the main determining factor of biome occupancy, little is known about the other life stages of South African nemestrinids. It likely that the putative hosts are plant herbivores, which may also have tight associations with these biomes. We used high accuracy GPS locations mapped in QGIS v3.10 (QGIS Development Team 2014) as well as in-field vegetation assessment for fine scale assessment of biome type. All individuals collected in the field were sorted to morphospecies corresponding to the morphospecies in the phylogenetic tree (chapter three). Morphospecies were then mapped in accordance with biomes as delineated by the South African National Biodiversity Institute (http://bgis.sanbi.org/SpatialDataset/Detail/669) using their GPS locations (Fig. 5.1; table S5.1). Only individuals from this project were used for mapping to ensure the accuracy of the GPS localities and morphospecies designation used. Since biome occupation was designated using coordinates from specimens, together with a fine-scale biome map which features

several interweaved biomes, the accuracy of the GPS localities is important for accurate biome allocation. Thus, museum specimens were not used because of the potentially uninformative or misleading nature of identifications and their relevant geographic label data (eg. old collection localities often have the name of the nearest town instead of precise locations) of many nemestrinids museum specimens (Barraclough 2006). Several morphospecies are consistently found on the border of two different biome types (ie. biome ecotones) from in-field observations, namely Forest and either Grassland or Savanna respectively, whereas other morphospecies traverse two interwoven biomes and can be found in either biome, as is the case for many species occurring in both Succulent Karoo and Fynbos. In all these cases (n = 12 morphospecies) we coded the character state as being polymorphic for both biomes in which they occurred.

Biome transitions

To reconstruct the biome state of the most recent common ancestor and to estimate if individual biomes have been colonised multiple times independently, we used the dated tree from the BEAST analysis. The RASP 4.2 (Yu *et al.* 2020) program was used to compare and implement models from BioGeoBEARS (Matzke 2013). We compared the performance of the six widely used biogeographical models (Matzke 2013) to test the fit of the models available: the dispersal–extinction–cladogenesis (DEC) model (Ree & Smith 2008), a maximum-likelihood form of the dispersal–vicariance (DIVALIKE) model (Ronquist 1997) and a maximum-likelihood form of the Bayesian biogeographical inference (BAYAREALIKE) model (Landis *et al.* 2013), each with and without the J parameter by their weighted AIC values. The models each differ in the cladogenetic events allowed, while the J parameter accounts for jump dispersal. Jump dispersal or founder-event speciation is a process by which long-distance dispersal occurs, potentially followed by radiation outside the

parental range. The maximum number of areas was set to two per lineage, which corresponds to the maximum number of areas occupied by any extant morphospecies in the tree. To account for phylogenetic uncertainty, this analysis was repeated a further 10 times on random trees drawn from the posterior of the Bayesian inference analysis from chapter three, using the sample function from R (R Core Team 2017). These 10 trees were dated using the chronos function in APE with a strict clock and a tolerance parameter of 0.0062.

RESULTS

The time-calibrated tree estimated from the BEAST analysis (Fig. S5.1) indicated that the age of the root of the clade is 18.9 MY (range 10.7 – 34.4). The *Moegistorhynchus* clade diversified in the Miocene (clade A crown age: 17.3 MY; range 9.6 – 31.9). Clade B diverged from the rest of the *Prosoeca* + *Stenobasipteron* clade 18.4 MYA (range 10.2 – 33.5) while clade C and D diverged from each other 17.4 MYA (range 9.8 – 31.9).

The BioGeoBEARS analysis suggested that the BAYAREALIKE+J model fit our dataset the best (AICw = 0.92, Table S2). Nine out of ten trees on which the analysis was repeated supported a Fynbos origin for the clade (mean = 78%) and a single tree showed weak support for either a Fynbos or Forest/Grassland origin (15%) (Fig. 5.2). The average number of shifts based on repeating the analysis with ten randomly chosen trees suggests that Fynbos (12.32), Grassland (9.90) and Forest (7.50) were the dominant source biomes from which dispersal took place, while all the seven biomes acted as destinations (Fig. 5.3, Table 5.1). The majority of speciation events occurred within Fynbos, Grassland and Forest (Fig. 5.3, Table 5.1).

DISCUSSION

In this study we examined ancestral biome occupation and transitions among biomes, representing broad ecological niches, in a clade of southern African nemestrinids. We provide evidence that this clade diversified from its most recent common ancestor *ca* 19 MYA and therefore that the extant species are relatively young compared to the old age suggested by the *Prosoeca* (*Palembolus*) fossils. The support for the BAYAREALIKE+J model indicates that the diversification of *Prosoeca*, *Stenobasipteron* and *Moegistorhynchus* as a clade can be characterized as a series of founder events into different biomes with subsequent radiation (Matzke 2013). We find reasonably strong support for a Fynbos origin of this clade. It is clear from the sampled topologies that morphospecies from particular biomes are not confined to strict monophyletic groups (Fig. 5.2). This suggests that the evolution of biome occupancy is relatively labile in this group.

The use of a limited dataset (ie. not including historical records), resulted in many morphospecies having single localities. Thus, some morphospecies are coded to occur in a single biome when they may be polymorphic. This represents a major caveat since it may lead to false conclusions about shifts between and within biomes (Dale *et al.* 2021). Dale *et al.*, (2021) suggested that the use of monomorphic coding generally leads to an underestimate of shifts rather than an overestimate. However, the possibility that a large increase in polymorphism, with increased sampling, would lead to a decrease in shifts between biomes cannot be excluded. It is unlikely that with further sampling, the ancestral biome of this clade will change as all the known species in *Moegistorhynchus* (ie. the sister genus to *Procoeca* + *Stenobasipteron*) (Barraclough 2006) and clade B are Fynbos species. While biomes, defined mainly on the basis of plants, may not be ideal as units for analysis for highly mobile insects, direct reliance by nemestrinids and their putative hosts on plants means that many species

may stick to a single biome. Therefore, these results provide insights into the patterns of colonisation of these broad plant-based ecological niches over evolutionary time.

Our results suggest that the timing of diversification in this clade of southern African Nemestrininae is more recent than any known fossil from this family found to date, including the *Prosoeca* (*Palembolus)* fossils (Bequaert & Carpenter 1936; Mostovski & Martínez-Delclòs 2000). This finding could be explained by a scenario in which the colonisation of Africa by this lineage of nemestrinids was relatively recent. The use of COI calibration instead of fossils may have contributed to the young age found in this analysis, although molecular clock dating is generally thought to overestimate the age of clades, whereas fossils are thought to underestimate the age of clades (Benton & Ayala 2003). Other analyses, such as a fossil-calibrated tree of bee flies, have been found to yield divergence dates consistent with estimates from a COI molecular clock (Li *et al.* 2021), suggesting that in certain groups different molecular dating methods may converge on similar results.

We find that *Moegistorhynchus* diverged from the *Prosoeca* and *Stenobasipteron* clade in the Fynbos *ca* 19 MYA, roughly coinciding with the aridification of the west coast of South Africa after the initial period of uplift *ca* 30 MYA (Siesser 1980). Other insect lineages such as the oil-collecting bee genus *Rediviva*, the lycaenid butterfly genus *Chrysoritis* and the bee-fly species *Megapalpus capensis* have been shown to have originated in the cape Fynbos region in a similar time period (*ca* 29 MYA, *ca* 32 MYA and *ca* 15 MYA, respectively) (de Jager & Ellis 2017; Kahnt *et al.* 2017; Talavera *et al.* 2020). The origin of *Rediviva* was estimated with the use of fossil-calibrated data while the origin of *Megapalpus* and *Chrysoritis* were estimated using COI calibration (de Jager & Ellis 2017; Kahnt *et al.* 2017; Talavera *et al.* 2020). This, suggests that these Cape Fynbos insect lineages may all have shared a similar biogeographic history (Morrone & Crisci 1995).

Some of the first colonisations of individual biomes coincide roughly with the estimated origin of the biomes themselves. The oldest recognisable Fynbos plant lineages are at least 60 My old, whereas known nemestrinid food plants such as *Pelargonium*, diversified *ca* 10 MYA (Verboom *et al.* 2009). The Succulent Karoo is considered to be a young biome, emerging with the increased summer aridity along the west coast *ca* 10 MYA (Linder 2003; Verboom *et al.* 2009). The oldest Succulent Karoo plant lineages are estimated to be *ca* 17 MY (Verboom *et al.* 2009) and the oldest Succulent Karoo nemestrinid lineage that diverged from a Fynbos ancestor in our analysis is 12.6 MYA (6.7 – 23.0). This transition is substantially earlier than that of *Megapalpus capensis* which expanded into the Succulent Karoo biome *ca* 3.5 MYA (de Jager & Ellis 2017). *Prosoeca* appears to have transitioned to the Grassland at least 8.1 MYA (4.4 – 14.7) which coincides with the estimated age of this biome (ie. *ca* 12 MY) in southern Africa (Bytebier *et al.* 2011).

The frequency and direction of transitions varied considerably among biomes with several evolutionary transitions between the seven biomes, suggesting that biome occupation is relatively labile for this lineage of nemestrinids. While shifts between biomes were numerous, the results suggest that the majority of speciation events took place within biomes. The dominance of speciation events within biomes coincides with what is thought to be the dominant pattern in plants. The presence of several biome transitions is similar to what has been found in lineages such as the *Pseudomys* division of Australian Rodents as well as Marsupials, which show a pattern of repeated biome transitions (Mitchell *et al.* 2014; Smissen & Rowe 2018), whereas *Rediviva* oil collecting bees show a relatively limited number of only two transitions from the winter-rainfall to summer-rainfall region (Kahnt *et al.* 2017). Multiple bidirectional transitions occurred between Fynbos and Grassland, but we also found evidence for multiple transitions from these two biomes to the other biomes. These bidirectional transitions contrast with the unidirectional transitions from the Cape

Fynbos region towards the Grasslands of the Drakensberg that have been reported for various plant genera such as *Disa* (Galley *et al.* 2007). Other lineages such as *Satyrium*, *Aloe* and *Kniphofia* have been suggested, without formal analysis, to show the opposite pattern of evolutionary movement from the Drakensberg to the Cape (Galley *et al.* 2007).

The evolutionary transition from open biomes into those with a canopy, such as Savanna and closed canopy forest, were less common than shifts between different open biomes. Only ~26% of morphospecies were found to have any association with biomes that had a canopy layer, although with further sampling this may prove to be an underestimate of the diversity. Another possible reason for the scarcity of species in biomes with a canopy may be a scarcity of suitable larval hosts or food plants in these biomes. Power *et al*. (2017) suggests that for plants, a transition from a closed canopy habitat to open vegetation and *vice versa* impose contrasting adaptive challenges. Forests have become increasingly limited in South Africa through natural processes, and more recently by anthropogenic causes. Therefore, the possibility of extremely rare or extinct Forest species cannot be ruled out. The Fynbos and Grassland are both occupied by fire-driven, seasonally dry, open vegetation, often with an overlap in plant genera composition between the two biomes (Rutherford *et al.* 2006). These similarities may potentially have made transitions between these two biomes physiologically easier than between other biomes (Power *et al.* 2017).

This study provides the first evidence for a Cape Fynbos origin during the Miocene of a clade of southern African Nemestrininae. We show that biome transitions are common and labile in a clade of pollinating flies that traverses seven different biomes. Bidirectional transitions between Fynbos and Grassland were particularly common in this group. These results suggest that the transition from open biomes to semi-closed and completely closed canopy biomes are less frequent than between different open biomes. Increased locality and taxon sampling of this clade of nemestrinids would be beneficial to improve biome characterization and to

confirm the rarity of speciation events within biomes other than Fynbos and Grassland. Furthermore, a broader sampling of the Nemestrinidae family may allow for the inclusion of fossil calibration points for a more accurate dating analysis.

ACKNOWLEDGMENTS

This study was funded by National Research Foundation of South Africa – Foundational Biodiversity Programme (NRF-FBIP grant # 110440 to A.G.E), National Research Foundation of South Africa (SARChI grant # 46372 to S.D.J).

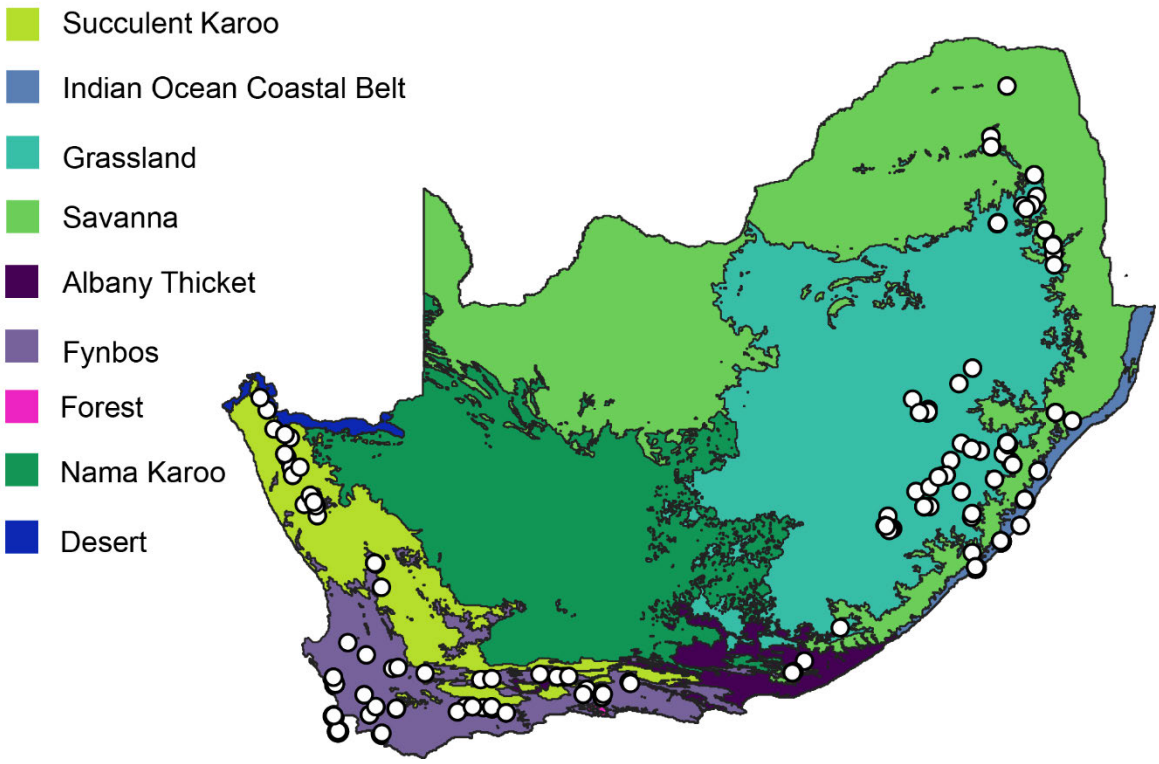


Figure 5.1 The distribution of all Nemestrinidae sampled for this study in the biomes of South Africa and eSwatini as designated by Mucina & Rutherford, 2006.

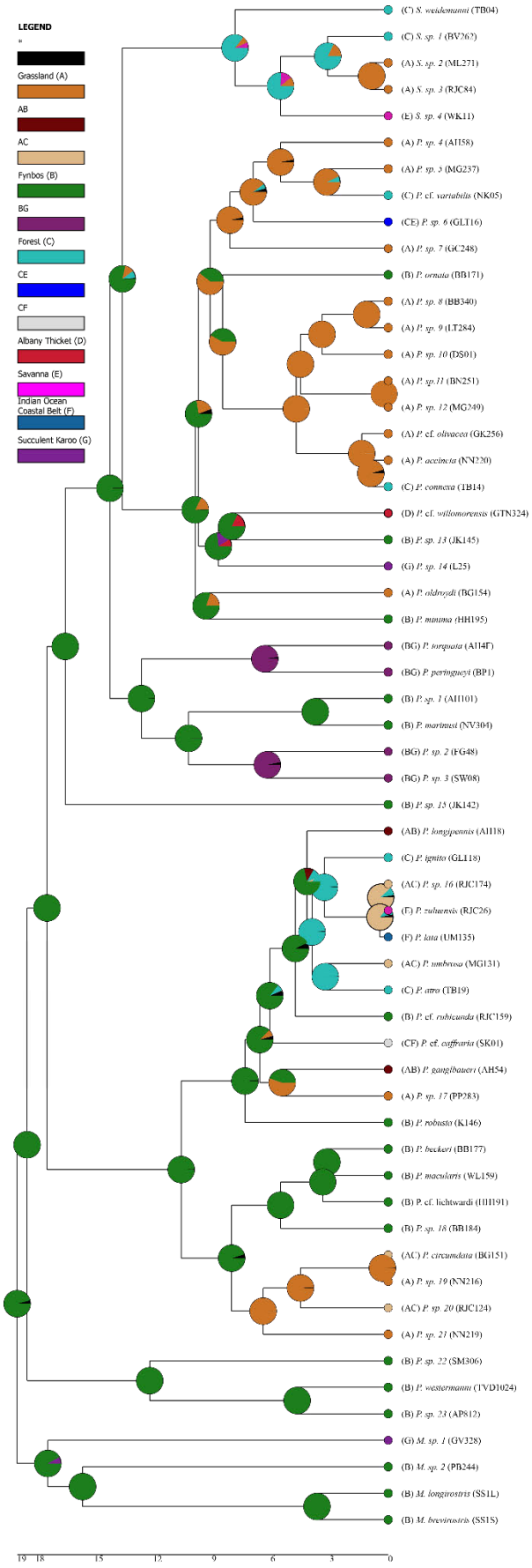

LEGEND
*
Grassland (A)
AB
AC
Fynbos (B)
BG
Forest (C)
CE
CF
Albany Thicket (D)
Savanna (E)
Indian Ocean Coastal Belt (F)
Succulent Karoo (G)
(C) S. weidemanni (TB04)
(C) S. sp. 1 (BV267)
(A) S. sp. 2 (ML271)
(A) S. sp. 3 (RJC84)
(E) S. sp. 4 (WK11)
(A) P. sp. 4 (AH158)
(A) P. sp. 5 (MG237)
(C) P. cf. variabilis (NK05)
(CE) P. sp. 6 (GLT16)
(A) P. sp. 7 (GC248)
(B) P. ornata (BB171)
(A) P. sp. 8 (BB340)
(A) P. sp. 9 (LT284)
(A) P. sp. 10 (DS01)
(A) P. sp.11 (BN251)
(A) P. sp. 12 (MG240)
(A) P. cf. olivacea (GK256)
(A) P. accincta (NN220)
(C) P. connexa (TB14)
(D) P. cf. willomorensis (GTN324)
(B) P. sp. 13 (JK145)
(G) P. sp. 14 (L25)
(A) P. oldroydi (BG154)
(B) P. minima (HH195)
(BG) P. torquata (AH4F)
(BG) P. peringueyi (BP1)
(B) P. sp. 1 (AH101)
(B) P. maritima (NV304)
(BG) P. sp. 2 (FG48)
(BG) P. sp. 3 (SW08)
(B) P. sp. 13 (JK142)
(AB) P. longipennis (AH18)
(C) P. ignita (GL118)
(AC) P. sp. 16 (RJC174)
(F) P. zuluensis (RJC26)
(F) P. lata (UM135)
(AC) P. umbrosa (MG131)
(C) P. atra (TB19)
(B) P. cf. rubicunda (RJC159)
(CF) P. cf. caffraria (SK01)
(AB) P. ganglbaueri (AH54)
(A) P. sp. 17 (PP283)
(B) P. robusta (K146)
(D) P. beckeri (BB177)
(B) P. macularis (WL159)
(B) P. cf. lichtwardi (HH191)
(B) P. sp. 18 (BB184)
(AC) P. circumdata (BG151)
(A) P. sp. 19 (NN216)
(AC) P. sp. 20 (RJC124)
(A) P. sp. 21 (NN219)
(B) P. sp. 22 (SM306)
(B) P. westermanni (TVD1024)
(B) P. sp. 23 (AP812)
(G) M. sp. 1 (GV328)
(B) M. sp. 2 (PB244)
(B) M. longirostris (SS1L)
(D) M. brevirostris (SS1S)
19 18 15 12 9 6 3 0

Figure 5.2 BAYAREALIKE + J reconstruction of ancestral biome reconstruction on one of the ten phylogenetic trees. Single letters corrospond to biomes: (A) Grassland, (B) Fynbos, (C) Forest, (D) Albany Thicket, (E) Savanna, (F) Indian Ocean Coastal Belt, (G) Succulent Karoo. Letter combinations corrospond to polymorphic species that can be found in both biomes.

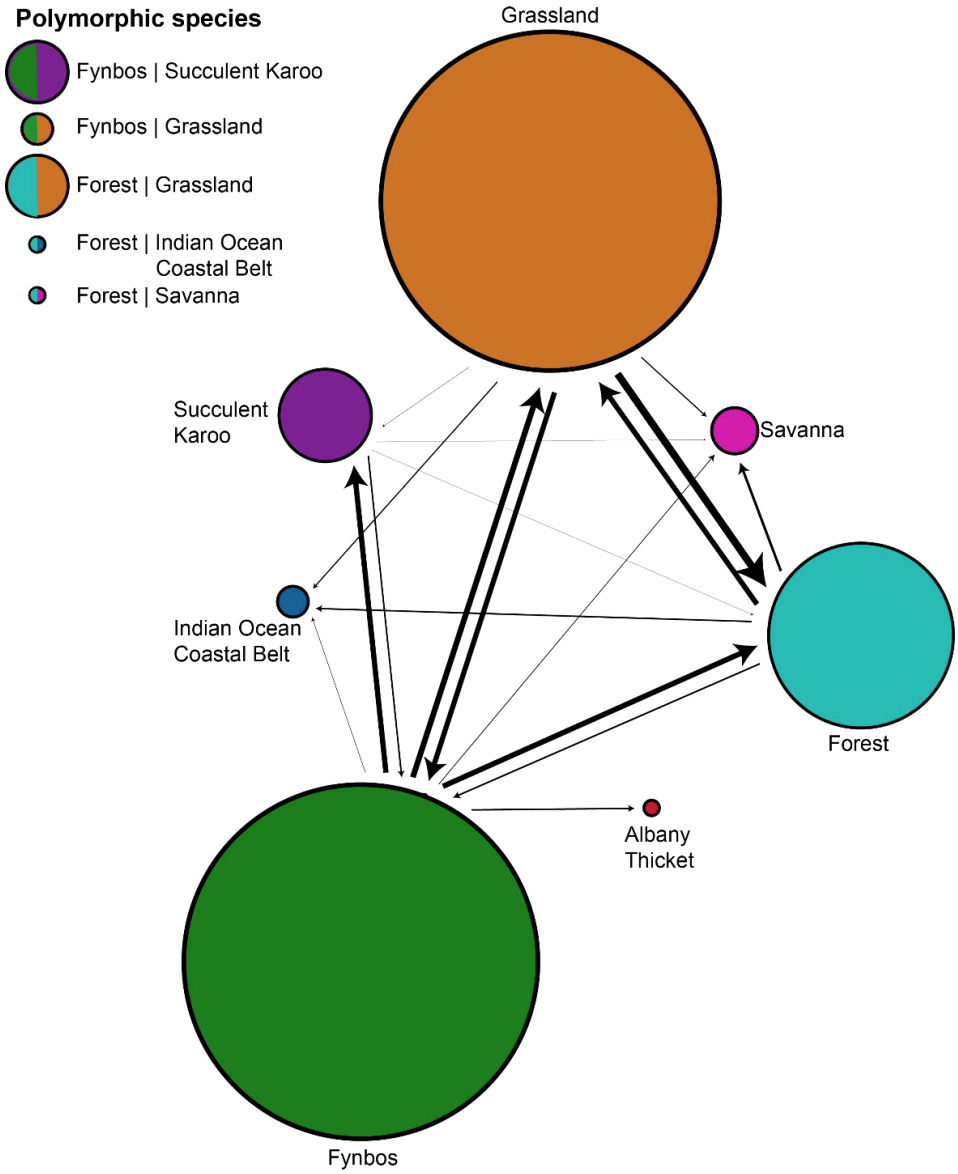


Figure 5.3 Evolutionary biome transitions within the nemestrinid clade containing *Moegistorhynchus*, *Prosoeca* and *Stenobasipteron*. The size of the circles is proportional to the total number of species sampled within each biome (polymorphic species are included in both biome totals in which they occur): Grassland, 22; Fynbos, 23; Forest, 12; Albany thicket, 1; savanna, 3; Indian Ocean Coastal Belt, 2; Succulent Karoo, 5. The size of the bi-coloured circles is proportional to the number of polymorphic species in each set of biomes. Arrow thickness is proportional to the number of transitions in each direction averaged over 10 phylogenetic trees, and absence of an arrow indicates the lack of any events.

Table 5.1 Average number of transitions for each biome as a source or destination. Bold numbers indicate within biome speciation without a change in biome. Non-integer numbers result from averaging transitions for 10 phylogenetic trees to account for phylogenetic uncertainty.

Destination	Source						
	Grassland	Fynbos	Forest	Albany Thicket	Savanna	Indian Ocean Coastal belt	Succulent Karoo
Grassland	**14.5**	3.7	3.8	0	0	0	0
Fynbos	3.35	**18.4**	1	0	0	0	0.2
Forest	4.85	3.45	**3.7**	0	0	0	0.1
Albany Thicket	0	1	0	**0**	0	0	0
Savanna	0.8	0.5	1.5	0	**0**	0	0.1
Indian Ocean Coastal Belt	0.8	0.3	0.5	0	0	**0**	0
Succulent Karoo	0.2	3.4	0	0	0	0	**2.3**

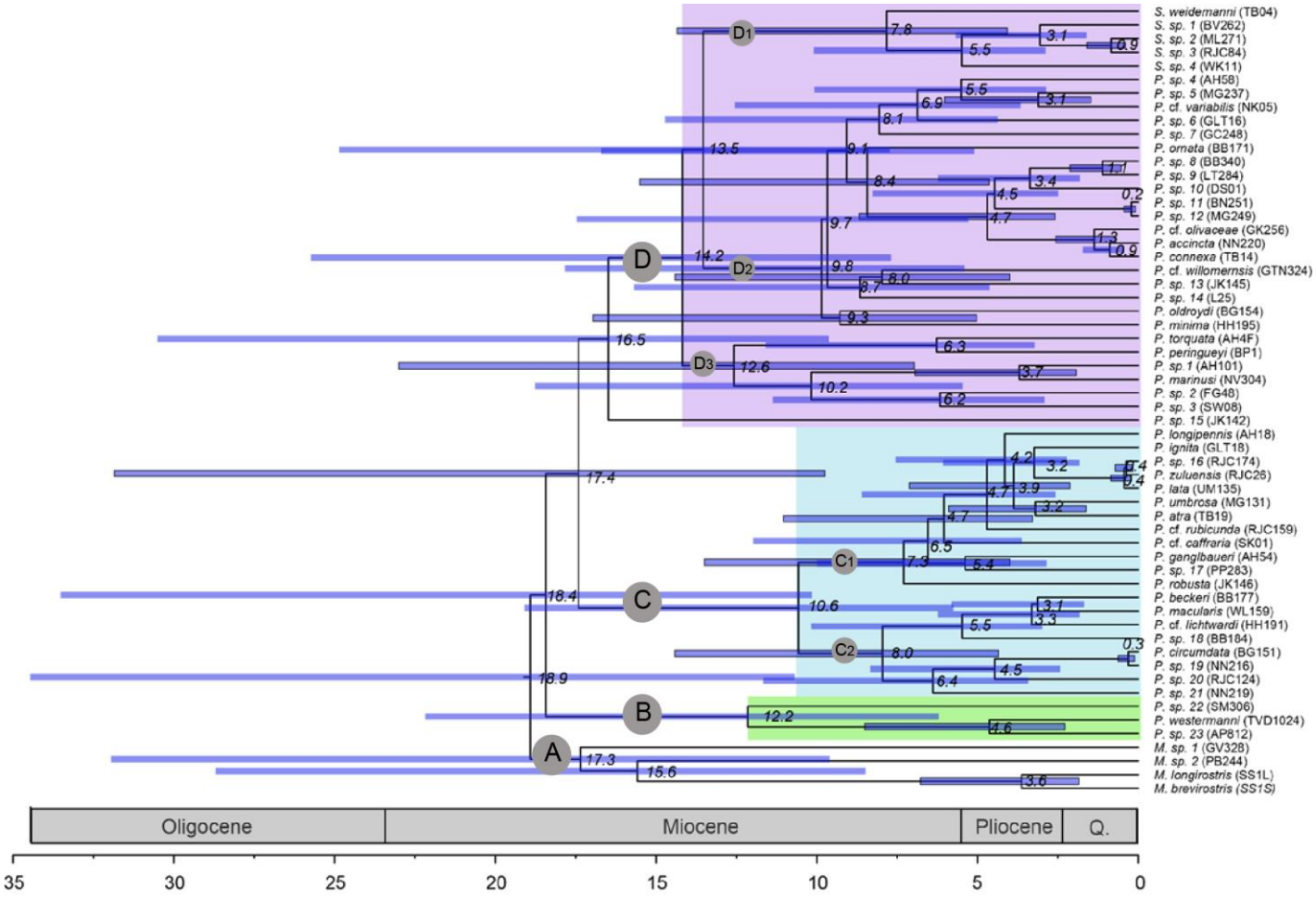


Figure S5.1 Dated maximum clade credibility tree with the estimated crown age next to relevant nodes and 95% confidence interval indicated with blue bars.

Table S5.1 Table of localities and specimen details of accessions used for the dating and biogeography analyses. Numbers of localities included refers to mapping species into biomes in Fig. 5.1 for further biogeographic analyses.

Sequence ID	Species	Locality	Lattitude	Longitude	Number of localities included
SS1S	*Moegistorhynchus brevirostris*	Silverstroomstrand	-33.528550	18.446101	1
SS1L	*Moegistorhynchus longirostris*	Silverstroomstrand	-33.528550	18.446101	1
GV328	*Moegistorhynchus sp. 1*	Groenvlei	-30.319344	18.067708	1
PB244	*Moegistorhynchus sp.2*	Picketberg	-32.79855	18.665922	1
NN220	*Prosoeca accincta*	Naudes Neck Pass	-30.732387	28.138476	2
TB19	*Prosoeca atra*	Mbotyi	-31.465393	29.730393	5
BB177	*Prosoeca beckeri*	Boosmansbos	-33.929018	20.86353	5
SK01	*Prosoeca* cf. *caffraria*	Vernon Crooke Nature Reserve	-30.274244	30.594822	2
HH191	*Prosoeca* cf. *lichtwardti*	Hottentots Holland Nature Reserve	-34.069298	19.047765	1
GK256	*Prosoeca* cf. *olivacea*	Graskop	-24.932787	30.808752	1
RJC159	*Prosoeca* cf. *rubicunda*	Swartberg	33.400	22.355	2
NK05	*Prosoeca* cf. *variabilis*	Nkandla Forest Reserve	-28.743981	31.139830	1
GTN324	*Prosoeca* cf. *willomoresis*	Grahamstown	-33.311184	26.526693	1
BG151	*Prosoeca circumdata*	PMB Botanical Gardens	-29.607538	30.34779	2
TB14	*Prosoeca connexa*	Mbotyi	-31.465391	29.730391	2
AH54	*Prosoeca ganglbaueri*	Naudes Nek	-30.77598	28.22365	43
GLT18	*Prosoeca ignita*	Haenertsberg	-23.880965	29.99877	1

UM135	*Prosoeca lata*	Umtamvuna Nature Reserve	-31.00363	30.17356	1
WL159	*Prosoeca macularis*	Waylands Nature Reserve	-33.408801	18.413050	1
NV304	*Prosoeca marinusi*	Niewoutville Botanical Garderns	-31.398152	19.141070	1
HH195	*Prosoeca minima*	Hottentots Holland Nature Reserve	-34.069298	19.047765	1
BG154	*Prosoeca oldroydi*	PMB Botanical Gardens	-29.607538	30.34779	1
BB171	*Prosoeca ornata*	Boosmansbos	-33.929018	20.86353	1
BP1	*Prosoeca peringueyi*	Botterkloof	-31.823601	19.256734	8
JK146	*Prosoeca robusta*	Jonaskop	-33.957261	19.520465	1
AH4F	*Prosoeca torquata*	Ariehoek	-30.244758	18.050258	11
MG131	*Prosoeca umbrosa*	Mount Gilboa	-29.308598	30.309024	7
TVD1024	*Prosoeca westermanni*	Price Alfred Hamlet	-33.236666	19.55	1
RJC26	*Prosoeca zuluensis*	Bisley Nature Reserve	-29.658909	30.387038	1
AH101	*Prosoeca sp. 1*	Nieuwoudtville	-31.387089	19.176518	1
FG48	*Prosoeca sp. 2*	Namaqualand	-30.371906	18.092013	6
AH58	*Prosoeca sp. 4*	Naudes Nek	-30.73242	28.14157	3
MG237	*Prosoeca sp. 5*	Mount Gilboa	-29.276283	30.281003	1
GLT16	*Prosoeca sp. 6*	Haenertsberg	-23.880965	29.99877	2
GC248	*Prosoeca sp. 7*	Giants Castle Nature Reserve	-29.276286	30.281006	1
BB340	*Prosoeca sp. 8*	Bulembo border post	-25.943825	31.106867	1
LT284	*Prosoeca sp. 9*	Long Tom Pass	-25.149548	30.619543	1
DS01	*Prosoeca sp. 10*	Dullstroom	-25.396739	30.126015	1
BN251	*Prosoeca sp. 12*	Bushmans Nek Pass	-29.843928	29.210244	1
MG249	*Prosoeca sp. 12*	Mount Gilboa	-29.278267	30.287912	1
JK145	*Prosoeca sp. 13*	Jonaskop	-33.968685	19.518869	1

L25	*Prosoeca sp. 14*	Langkloof	-30.55937	18.128272	1
JK142	*Prosoeca sp. 15*	Jonaskop	-33.957259	19.520465	1
RJC174	*Prosoeca sp. 16*	Eshowe	28.8910833	31.4349	5
PP283	*Prosoeca sp. 17*	Paardeplaats Nature Reserve	-25.095679	30.562498	3
BB184	*Prosoeca sp. 18*	Boosmansbos	-33.929018	20.86353	1
NN216	*Prosoeca sp. 19*	Naudes Neck Pass	-30.736268	28.13939	2
RJC124	*Prosoeca sp. 20*	Saddleback Pass	-25.79536	31.094741	5
NN219	*Prosoeca sp. 21*	Naudes Neck Pass	-30.736268	28.13939	2
SM306	*Prosoeca sp. 22*	Silvermine Nature Reserve	-34.085932	18.418983	1
AP812	*Prosoeca sp. 23*	Cape Peninsula	-34.354020	18.493001	1
BV262	*Stenobasipteron sp. 1*	Bridal Veil Falls	-25.082029	30.723459	1
ML271	*Stenobasipteron sp. 2*	Mountainlands Nature Reserve	-25.7735	31.08543	2
RJC84	*Stenobasipteron sp. 3*	Abel Erasmus Pass	-24.55128	30.76479	1
WK11	*Stenobasipteron sp. 4*	Serala Forest Reserve	-24.055322	30.005245	1
TB04	*Stenobasipteron weidermanni*	Grahamstown	-33.328778	26.500194	4

Table S5.2 Weighted AIC values for the six biogeographic models tested in RASP using BioGeoBEARS.

Model	AICw
DEC	1.4e-10
DEC+J	0.071
DIVALIKE	1.7e-11
DIVALIKE+J	0.0047
BAYAREALIKE	3.7e-17
BAYAREALIKE+J	0.92

Chapter 6

Conclusions

Understanding the relationships within and between lineages is an essential first step to conserving and understanding the biology, ecology and evolution of taxa (Funk *et al.* 2002; Mace 2004). This thesis contributes to the understanding of both the diversity of Nemestrinidae in southern Africa and the relationships within and among three southern African genera: *Moegistorhynchus*, *Prosoeca* and *Stenobasipteron* (hereafter together referred to as the "MPS clade" for convenience). The phylogeny presented in this thesis, which includes 58 morphospecies from three genera, is the most comprehensive molecular phylogeny of southern African nemestrinids to date. Although the taxon sampling in this study did not include representatives of all southern African Nemestrinidae, evidence discussed below does suggest monophyly of the MPS clade. Furthermore, ancestral character state reconstruction of proboscis length and biome has provided insights into the evolutionary history of the MPS clade.

The Nemestrinidae are thought to be an ancient dipteran family, with a Triassic origin (Ansorge & Mostovski 2000), but virtually nothing is known about their evolution in time and space. Wiegmann *et al.* (2011) on the other hand suggest a Jurassic origin for the Nemestrinidae family using a fossil calibrated time tree of Diptera. The results presented in this thesis suggest a relatively young southern African radiation of the MPS clade (*ca.* 19 MYA) within the Nemestrinidae. Analyses presented in this thesis indicate that the MPS clade originated in the Fynbos biome from a short proboscid (*ca.* 12mm) most recent common ancestor and then radiated into five diverse clades that occupy the majority of South African biomes. The estimated age of the MPS clade is somewhat younger than the origin of some of the oldest plant clades that make up the characteristic Fynbos vegetation (*ca.* 61 MYA) (Verboom *et al.* 2009), and corresponds to timing of diversification of various other insect lineages (de Jager & Ellis 2017; Kahnt *et al.* 2017; Talavera *et al.* 2020). This suggests that these insect lineages may all have shared a similar biogeographic history (Morrone &

Crisci 1995). A shared history of these insect lineages through space and time may indicate that they have diverged in response to similar climatic or biotic variables that emerged during the Miocene. The estimated timing of the first transitions of *Prosoeca + Stenobasipteron* into biomes like the Grassland by short proboscid ancestors roughly corresponds to the estimated origin of the Grassland biome (Bytebier *et al.* 2011). My results cannot exclude the possibility that the MPS clade diverged from lineages that include other southern African genera (ie. *Nycterimyia*, *Trichopsidea* and *Atriadops*), as none of these genera were represented in the phylogeny. However, based on morphological synapomorphies and preliminary phylogenetic results, it is more likely that the MPS clade is sister to one of the genera that are part of Nemestrininae or *Hirmoneura*. The remaining three Nemestrininae genera and *Hirmoneura* have no extant species in southern Africa, and no southern African Nemestrinidae fossils are available. This suggests that Southern Africa was likely colonised by nemestrinids from elsewhere. The current sampling does not allow inferences about the number of times southern Africa was colonised by various nemestrinid lineages and from where they dispersed.

Patterns of species diversity of Southern African Nemestrininae are uneven across biomes. The Grassland and Fynbos biomes are comparatively species-rich in nemestrinid species. Several explanations, which are not mutually exclusive, have been proposed to explain uneven patterns of diversity in other clades. First, if the amount of time that a lineage was present in an area differs, uneven diversity may arise, even if diversification rates are similar across regions (McPeek & Brown 2007; Sundaram *et al.* 2019). The perceived climatic stability in the Fynbos biome (Dynesius & Jansson 2000; Sandel *et al.* 2011) may have allowed time for the high diversity of nemestrinid species to accumulate. However, the more recently colonized Grassland biome (*ca.* 8 MYA) also has a high number of morphospecies, suggesting that factors in addition to time must also have played an important role in the

radiation of these taxa. Second, in studies of plant and bird lineages, vegetation structure has been shown to affect diversification rates (Onstein *et al.* 2016; Pinto-Ledezma *et al.* 2017). In particular, open biomes with no canopy layer, such as the Fynbos and Grassland, have been suggested to have higher diversification rates than biomes without a closed canopy in plants (Smith & Donoghue 2008). Vegetation structure, however, is unlikely to have affected diversification in the clade studied here, as no shifts in diversification were detected across the phylogeny, even though representatives occur in closed and open habitats. Thirdly, other aspects of biomes, related to availability of biotic niches that provide opportunities for diversification, may explain uneven diversity patterns (Johnson 2010). Florally diverse biomes, such as Fynbos and Grassland, may therefore have offered increased opportunities for diversifying on floral feeding niches. Alternatively, given that Nemestrinidae likely rely on larval hosts for part of their life cycle (Potgieter 1929; Bernardi 1973; Barraclough 2017), host diversity may also be associated with increased opportunities for diversification, as has been suggested by a general association between parasite diversity and host diversity (Kamiya *et al.* 2014). However, in the absence of any information on hosts in southern African nemistrinids, this explanation cannot yet be evaluated. Overall, it appears that there may be no single explanation for the disparity in diversity among biomes, but rather an interplay of multiple mechanisms may be driving patterns of diversity.

While there are disparities in the patterns of diversity among biomes and proboscis states, all biomes investigated have nemestrinid species with long and short probosides, suggesting that there is no clear association between certain biomes and a particular proboscis length. There is also no evidence to suggest that diversification rates are associated with transition to a particular proboscis length or certain biome. Biome occupation and proboscis length appear to be relatively labile, with multiple shifts between states and exhibit a degree of

directionality. The directionality found in transitions should, however, be interpreted with care as phylogenetic uncertainty and increased sampling are both likely to affect this result.

Within the Nemestrinidae, the clade of southern African nemestrinids is particularly diverse in morphological traits, including proboscis length (Bernardi 1973). Selection pressures from various sources have potentially led to traits showing transitions among almost all states. The effects of different selection pressures on different species may be illustrated by the variation in static allometry slopes between morphospecies with long and short proboscides. Long proboscides with positive allometric slopes may generally be the result of directional selection towards longer traits, for increased nectar acquisition (Voje 2016). Short proboscides, with negative allometric slopes, may rather have experienced stabilising selection to match the tube length of the flowers they visit (Pélabon *et al.* 2011). While studies on long proboscid species have suggested that directional selection for longer traits is imposed by floral tubes (Anderson & Johnson 2009; Pauw *et al.* 2009), very little work has been done on short proboscid species. Therefore, future work focused on the selection pressures imposed on short proboscides will provide insights into the factors influencing the lengths of traits in these short proboscid species.

While it is unclear exactly what has caused the high diversity of southern African nemistrinids, it is clear that the diversity in southern Africa has been severely underestimated. The historically overlooked *Prosoeca peringueyi* complex is a clear example of the taxonomic neglect of the genera in the MPS clade. The two species in this complex exhibited conspicuous morphological differences which, in conjunction with molecular evidence, were used to justify the description of a new species (chapter 2). The use of phylogenetic reconstruction and molecular species delimitation allowed us to estimate that the southern African species may eventually reach double the number of species currently described, as

estimates suggest that almost half of the diversity sampled is undescribed. Additionally, *Prosoeca* may ultimately be among the most diverse nemestrinid genera in the world, alongside *Nemestrinus* and *Trichopthalma* which have *ca.* 66 and 63 species described, respectively. Many of the undescribed species in the MPS clade are potentially important mutualists and parasites that are involved in ecological processes in the environment (Molbo *et al.* 2003; Smith *et al.* 2006; Suzán *et al.* 2009; Solodovnikov & Shaw 2017). Interpretation of mutualistic interactions is determined by proper understanding of taxonomy. An example of how taxonomy affects the understanding of the specificity of interactions, and its implications for studies of co-evolution, is highlighted by fig-wasp mutualisms. These fig-wasp mutualisms were thought to exhibit exceptionally high degrees of reciprocal specialization. However, molecular tools have repeatedly demonstrated that many fig species are pollinated by multiple cryptic wasp species and that many of these mutualisms are more diffuse than a tight, one to one relationship (Michaloud *et al.,* 1996; Kerdelhue *et al.,* 1999; Molbo *et al.*, 2003; Sutton *et al.*, 2017). A high degree of specialization is thought be an indication of elevated extinction risk (Bond 1994) while increased generalization may buffer against extinction (Pauw 2007). In the MPS clade, more than 150 plants species, which form part of nine well-studied pollination guilds, are each thought to rely exclusively on a single nemestrinid species for pollination (Manning & Goldblatt 1996, 1997, 2000; Potgieter & Edwards 2005; Anderson & Johnson 2009; Newman *et al.* 2014). However, in at least two of these guilds the pollinator is thought to be represented by a species complex, rather than by a single species (Barraclough 2017; chapter 2). The presence of cryptic species and numerous undescribed, ecologically important, nemestrinid pollinators hampers conservation efforts as well as attempts to study the ecology and biology of the southern African Nemestrininae (Sangster 2009). With anthropogenic activities increasing the risk of species loss, the

importance of knowledge about the diversity and ecology of southern African Nemestrininae is essential to develop effective conservation strategies (Ceballos & Ehrlich 2009).

Understanding the taxonomic relationships among taxa provides an accurate view of the management units of conservation (ie. the species) and a foundation for examining ecological and biological hypotheses. Thus, the paraphyly of *Prosoeca* with respect to *Stenobasipteron* illustrates the need for further taxonomic attention of these two genera. Without a comprehensive molecular phylogeny for the Nemestrinidae it cannot be known with certainty that *Stenobasipteron*, *Moegistorhynchus* and *Prosoeca* form a monophyletic clade. However, morphological synapomorphies, such as a well-developed proboscis, the morphology of the female genitalia and the geographic distribution of unsampled genera (Bernardi 1973), suggest that it is unlikely that the monophyly of the MPS clade will be refuted with more comprehensive taxon sampling. My preliminary results of the relationships between the MPS clade and the outgroups suggest that the subfamily Nemestrininae is potentially not well-defined and that *Hirmoneura* may need to be included in Nemestrininae or that Nemestrininae may need to be split further, although additional sampling is required to reinforce these results. The affinity of *Moegistorhynchus* and *Prosoeca* + *Stenobasipteron* to the remaining, unsampled, genera of the Nemestrinidae is still unknown, thus hindering any conclusive statements about the relationships of the genera of the MPS clade with each other and within the family. As the need for extensive integrative taxonomic study of this clade and subfamily is clear, the relatively well-supported phylogeny I have constructed can form the starting point for the further exploration of systematics, taxonomy, biology and evolution.

While the sampling across the five clades (A-E) in the phylogenetic tree reconstructed in chapter three was unbiased and extensive, the taxon and character sampling for individual taxa is by no means complete for *Moegistorhynchus*, *Stenobasipteron* or *Prosoeca*. It has

been noted that increased taxon sampling increases the accuracy of phylogenetic reconstructions, estimates of diversification rates (Cusimano *et al.* 2010; Brock *et al.* 2011) as well as reconstructions of ancestral character states (Maddison 1995). A lack of information, such as too few gene regions or species sampled, may lead to phylogenetic hypotheses with weak or conflicting support. This may be expressed as multiple alternative topologies, polytomies and/or low support values (Fitzjohn *et al.* 2009; Diniz-Filho *et al.* 2013). The possibility that the effects of phylogenetic uncertainty and unsampled taxa may have led to erroneous conclusions and reconstructions of character states (Rangel *et al.* 2015) in this thesis cannot be excluded. Thus, increased sampling of taxa and the use of several more gene regions or implementation of next generation sequencing techniques may have improved the resolution of the reconstructed phylogeny. The accuracy of the ancestral reconstruction also relies on a sufficiently realistic model of evolution and accurate trait measurements (Joy *et al.* 2016). Within-species variation in trait measurements may affect the accuracy of ancestral reconstructions and can be introduced by instrument error, sampling variation or natural trait variation within a population. The inclusion of a measure of error (eg. standard error) instead of simple trait means can potentially improve parameter estimation of reconstructions (Ives *et al.* 2007). The results of this thesis nonetheless contribute a relatively well-resolved phylogeny on which further research can be built and provides valuable insights into the patterns of evolution within and among species of the MPS clade of nemestrinids.

Possible directions for future research

This work has the potential for expansion, with further study of the systematics, taxonomy, ecology and evolution of the MPS clade as well as the Nemestrinidae family.

Extensive inter- and intraspecific sampling of southern African species as well as worldwide genera for a family-level phylogeny would be ideal to corroborate the patterns I have found

here. Such a phylogeny may provide deeper insight into the monophyly of the MPS clade as well as the monophyly of the Nemestrininae and might be useful to identify the sister group of the MPS clade. Knowledge of the sister lineage to the MPS clade would allow investigation into the colonisation of the Fynbos from elsewhere. A comprehensive family level phylogeny would also enable precise placement of many of the available fossils and thus be ideal to confirm the estimated timing of divergence. Furthermore, it may be possible to test if a well-developed proboscis, as opposed to a vestigial proboscis, may have led to an increase in diversification rates within Nemestrinidae, as those genera with well-developed probosides tend to be more species-rich compared to those without (Bernardi 1973).

Further sampling of nemestrinids in different seasons and new localities, together with extensive taxonomic work, will contribute towards an accurate estimate of the diversity of Nemestrinidae in southern Africa through several new species descriptions. Furthermore, extensive taxonomic work and additional species delimitation analyses will contribute towards the resolution of various species complexes, such as the *P. ganglbaueri* and *P. longipennis* complexes. Thorough taxonomic study focused on identifying morphological synapomorphies associated with molecular clades identified in the phylogeny, may ultimately change the delineations of *Prosoeca* and *Stenobasipteron*.

Detailed study of the ecology and biology of various species may provide useful insights into the developmental mechanisms and selection controlling changes in proboscis length between species. A focus on larval host use, the influence of abiotic factors and selection on proboscis length by flower tube depth may be ideal to identify potential drivers of morphological change. Detailed study of the floral resource use patterns of co-occurring, sympatric species such as *P. peringueyi* and *P. torquata* may provide insights into aspects affecting their co-existence, such as competition and niche partitioning. Furthermore, understanding patterns of

flower visitation and pollination by these two *Prosoeca* species may provide an indication of pollination redundancy and thus extinction risk, in a guild of plants that were thought to be pollinated by a single species.

Investigation into the role of nemestrinid species as potential environmental indicators may aid future conservation efforts. Nemestrinidae in southern Africa may represent ideal environmental indicator species as they likely have a complex life cycle that depends on abiotic conditions and biotic interaction with larval hosts and numerous plant species as adults. Thus, a single nemestrinid species likely has an impact on several other organisms in the environment and *vice versa*. An understanding of both their larval and adult ecology would be key for monitoring and conservation prioritisation.

REFERENCES

ADAMSKI, D., BOEGE, K., LANDRY, J.F., SOHN, J.C., 2009. Two new species of *Wockia heinemann* (Lepidoptera: Urodidae) from coastal dry-forests in western México. *Proc. Entomol. Soc. Washingt.* **111**, 166–182.

ALBERT, J.S., 2020. Phylogenetic Character Reconstruction, in: Kaas, J.H. (Ed.), Evolutionary Neuroscience. Academic Press, pp. 17–28.

AMORI, G., GIPPOLITI, S., LUISELLI, L., 2011. Do biodiversity hotspots match with rodent conservation hotspots? *Biodivers. Conserv.* **20**, 3693–3700.

ANDERSON, B., JOHNSON, S.D., 2008. The Geographical Mosaic of Coevolution in a Plant-Pollinator Mutualism. *Evolution.* **62**, 220–225.

ANDERSON, B., JOHNSON, S.D., 2009. Geographical covariation and local convergence of flower depth in a guild of fly-pollinated plants. *New Phytol.* **182**, 533–540.

ANDERSON, B., PAUW, A., COLE, W.W., BARRETT, S.C.H., 2016. Pollination, mating and reproductive fitness in a plant population with bimodal floral-tube length. *J. Evol. Biol.* **29**, 1631–1642.

ANDERSON, B., ROS, P., WIESE, T.J., ELLIS, A.G., 2014. Intraspecific divergence and convergence of floral tube length in specialized pollination interactions. *Proc. R. Soc. B Biol. Sci.* **281**, 1–9.

ANDERSON, B., TERBLANCHE, J.S., ELLIS, A.G., 2010. Predictable patterns of trait mismatches between interacting plants and insects. *BMC Evol. Biol.* **10**.

ANSORGE, J., MOSTOVSKI, M.B., 2000. Redescription of *Prohirmoneura jurassica* Handlirsch 1906 (Diptera : Nemestrinidae) from the Lower Tithonian lithographic limestone of Eichstatt (Bavaria). *Neues Jahrb. Fur Geol. Und Palaontologie-*

Monatshefte **1906**, 235–243.

ARMBRUSTER, W.S., MUCHHALA, N., 2009. Associations between floral specialization and species diversity: Cause, effect, or correlation? *Evol. Ecol.* **23**, 159–179.

BAKER, R.H., WILKINSON, G.S., 2001. Phylogenetic analysis of sexual dimorphism and eye-span allometry in stalk-eyed flies (Diopsidae). *Evolution.* **55**, 1373–1385.

BARRACLOUGH, D., SLOTOW, R., 2010. The South African keystone pollinator *Moegistorhynchus longirostris* (Wiedemann, 1819) (Diptera: Nemestrinidae): Notes on biology, biogeography and proboscis length variation. *African Invertebr.* **51**, 397–403.

BARRACLOUGH, D.A., 2006. An overview of the South African tangle-veined flies (Diptera: Nemestrinidae), with an annotated key to the genera and a checklist of species. *Zootaxa* **63**, 39–63.

BARRACLOUGH, D.A., 2017. Nemestrinidae (Tangle-veined Flies). *Man. Afrotropical Diptera* **2**, 995–1003.

BARRACLOUGH, D.A., COLVILLE, J.F., KAROLYI, F., KRENN, H.W., 2018. A striking new species of *Prosoeca* Schiner, 1867 (Diptera: Nemestrinidae): An important pollinator from the Bokkeveld Plateau, Northern Cape Province, South Africa. *Zootaxa* **4497**, 411–421.

BAUDER, J.A.-S., KAROLYI, F., 2019. Superlong Proboscises as Co-adaptations to Flowers, in: Krenn, H.W. (Ed.), Insect Mouthparts, Zoological Monographs. Springer International Publishing, Cham.

BAUDER, J.A.S., MORAWETZ, L., WARREN, A.D., KRENN, H.W., 2015. Functional constraints on the evolution of long butterfly proboscides: Lessons from Neotropical skippers (Lepidoptera: Hesperiidae). *J. Evol. Biol.* **28**, 678–687.

BENTON, M.J., AYALA, F.J., 2003. Dating the tree of life. *Science.* **300**, 1698–1700.

BEQUAERT, J.C., 1932. The Nemestrinidae (Diptera) in the V. v. Roder Collection. *Zool. Anzeiger Leipzig.* **100**, 13–33.

BEQUAERT, J.C., CARPENTER, F., 1936. The Nemestrinidae of the Miocene of Florissant, Colorado, and Their Relations to the Recent Fauna. *J. Paleontol.* **10**, 395–409.

BERLOCHER, S.H., FEDER, J.L., 2002. Sympatric Speciation in Phytophagous Insects: Moving Beyond Controversy? *Annu. Rev. Entomol.* **47**, 773–815.

BERNARDI, N., 1973. The genera of the family Nemestrinidae (Diptera: Brachycera). *Arq. Zool.*

BEZZI, M., 1924. The South African Nemestrinidae (Diptera) as represented in the South African Museum. *Ann. South African Museum.* **19**, 164–190.

BICKFORD, D., LOHMAN, D.J., SODHI, N.S., NG, P.K.L., MEIER, R., WINKER, K., INGRAM, K.K., DAS, I., 2007. Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* **22**, 148–155.

BLAIR, C.P., ABRAHAMSON, W.G., JACKMAN, J.A., 2005. Cryptic speciation and host-race formation in a purportedly generalist tumbling flower beetle. *Evolution.* **59**, 304–316.

BLOMBERG, S.P., GARLAND, T., IVES, A.R., 2003. Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits Are More Labile. *Evolution.* **57**, 717.

BONAL, R., ESPELTA, J.M., VOGLER, A.P., 2011. Complex selection on life-history traits and the maintenance of variation in exaggerated rostrum length in acorn weevils. *Oecologia.* **167**, 1053–1061.

BOND, W.J., 1994. Do mutualisms matter? Assessing the impact of pollinator and disperser

disruption on plant extinction. *Philos. Trans. - R. Soc. London, B.* **344**, 83–90.

BOND, W.J., 2015. Fires in the Cenozoic: A late flowering of flammable ecosystems. *Front. Plant Sci.* **5**, 1–11.

BONDURIANSKY, R., DAY, T., 2003. The evolution of static allometry in sexually selected traits. *Evolution.* **57**, 2450–2458.

BORRELL, B.J., 2005. Long tongues and loose niches: Evolution of euglossine bees and their nectar flowers. *Biotropica.* **37**, 664–669.

BOSC, C., HUI, C., ROETS, F., PAUW, A., 2019. Importance of biotic niches versus drift in a plant-inhabiting arthropod community depends on rarity and trophic group. *Ecography.* **42**, 1926–1935.

BREESCHOTEN, T., DOORENWEERD, C., TARASOV, S., VOGLER, A.P., 2016. Phylogenetics and biogeography of the dung beetle genus *Onthophagus* inferred from mitochondrial genomes. *Mol. Phylogenet. Evol.* **105**, 86–95.

BROCK, C.D., HARMON, L.J., ALFARO, M.E., 2011. Testing for temporal variation in diversification rates when sampling is incomplete and nonrandom. *Syst. Biol.* **60**, 410–419.

BROOKS, T.M., MITTERMEIER, R.A., DA FONSECA, G.A.B., GERLACH, J., HOFFMANN, M., LAMOREUX, J.F., MITTERMEIER, C.G., PILGRIM, J.D., RODRIGUES, A.S.L., 2006. Global Biodiversity Conservation Priorities. *Science.* **313**, 58–61.

BROSI, B.J., BRIGGS, H.M., 2013. Single pollinator species losses reduce floral fidelity and plant reproductive function. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13044–13048.

BROWER, A.V.Z., 1994. Rapid morphological radiation and convergence among races of

the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6491–6495.

BROWN, J.W., MILLER, S.E., HORAK, M., 2003. Studies on New Guinea moths. 2. Description of a new species of *Xenothictis Meyrick* (Lepidoptera: Tortricidae: Archipini). *Proc. Entomol. Soc. Washingt.* **105**, 1043–1050.

BROWN, B. V., BORKENT, A., ADLER, P.H., AMORIM, D. de S., BARBER, K., BICKEL, D., BOUCHER, S., BROOKS, S.E., BURGER, J., BURINGTON, Z.L., CAPELLARI, R.S., COSTA, D.N.R., CUMMING, J.M., CURLER, G., DICK, C.W., EPLER, J.H., FISHER, E., GAIMARI, S.D., GELHAUS, J., GRIMALDI, D.A., HASH, J., HAUSER, M., HIPPA, H., IBÁÑEZ-BERNAL, S., JASCHHOF, M., KAMENEVA, E.P., KERR, P.H., KORNEYEV, V., KORYTKOWSKI, C.A., KUNG, G.A., KVIFTE, G.M., LONSDALE, O., MARSHALL, S.A., MATHIS, W., MICHELSEN, V., NAGLIS, S., NORRBOM, A.L., PAIERO, S., PAPE, T., PEREIRA-COLAVITE, A., POLLET, M., ROCHEFORT, S., RUNG, A., RUNYON, J.B., SAVAGE, J., SILVA, V.C., SINCLAIR, B.J., SKEVINGTON, J.H., STIREMAN, J.O., SWANN, J., THOMPSON, F.C., VILKAMAA, P., WHEELER, T., WHITWORTH, T., WONG, M., WOOD, D.M., WOODLEY, N., YAU, T., ZAVORTINK, T.J., ZUMBADO, M.A., 2018. Comprehensive inventory of true flies (Diptera) at a tropical site. *Commun. Biol.* **1**, 1–8.

BYTEBIER, B., ANTONELLI, A., BELLSTEDT, D.U., LINDER, H.P., 2011. Estimating the age of fire in the Cape flora of South Africa from an orchid phylogeny. *Proc. Biol. Sci.* **278**, 188–195.

CARDILLO, M., WESTON, P.H., REYNOLDS, Z.K.M., OLDE, P.M., MAST, A.R., LEMMON, E.M., LEMMON, A.R., BROMHAM, L., 2017. The phylogeny and

biogeography of *Hakea* (Proteaceae) reveals the role of biome shifts in a continental plant radiation. *Evolution.* **71**, 1928–1943.

CARDOSO-GUSTAVSON, P., SAKA, M.N., PESSOA, E.M., PALMA-SILVA, C., PINHEIRO, F., 2018. Unidirectional transitions in nectar gain and loss suggest food deception is a stable evolutionary strategy in *Epidendrum* (Orchidaceae): Insights from anatomical and molecular evidence. *BMC Plant Biol.* **18**, 1–13.

CARSTENS, B.C., PELLETIER, T.A., REID, N.M., SATLER, J.D., 2013. How to fail at species delimitation. *Mol. Ecol.* **22**, 4369–4383.

CATERINO, M.S., CHO, S., SPERLING, F.A.H., 2000. The current state of insect molecular systematics: A thriving tower of babel. *Annu. Rev. Entomol.* **45**, 1–54.

CEBALLOS, G., EHRLICH, P.R., 2009. Discoveries of new mammal species and their implications for conservation and ecosystem services. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 3841–3846.

CHANGBUNJONG, T., BHUSRI, B., SEDWISAI, P., WELUWANARAK, T., NITIYAMATAWAT, E., CHAREONVIRIYAPHAP, T., RUANGSITTICHAI, J., 2018. Species identification of horse flies (Diptera: Tabanidae) in Thailand using DNA barcoding. *Vet. Parasitol.* **259**, 35–43.

CHAPPLE, D.G., KEOGH, J.S., 2004. Parallel adaptive radiations in arid and temperate Australia: Molecular phylogeography and systematics of the *Egernia whitii* (Lacertilia: Scincidae) species group. *Biol. J. Linn. Soc.* **83**, 157–173.

CHOWN, S.L., GASTON, K.J., 2010. Body size variation in insects: A macroecological perspective. *Biol. Rev.* **85**, 139–169.

CHRISTENHUSZ, M.J.M., BYNG, J.W., 2016. The number of known plants species in the world and its annual increase. *Phytotaxa* **261**, 201–217.

CISNEROS, G., 2015. Analyzing the geographic mosaic of coevolution of the long-tonged pollinator *Prosoeca ganglbaueri* Lichtwardt, 1910 (Diptera, Nemestrinidae) in the Drakensberg mountains in South Africa by using nuclear and mitochondrial markers. Leiden.

CLARKE, T.E., LEVIN, D.B., KAVANAUGH, D.H., REIMCHEN, T.E., 2001. Rapid evolution in the *Nebria gregaria* group (Coleoptera: Carabidae) and the paleogeography of the queen charlotte islands. *Evolution.* **55**, 1408–1418.

CRISP, M.D., ARROYO, M.T.K., COOK, L.G., GANDOLFO, M.A., JORDAN, G.J., MCGLONE, M.S., WESTON, P.H., WESTOBY, M., WILF, P., LINDER, H.P., 2009. Phylogenetic biome conservatism on a global scale. *Nature.* **458**, 754–756.

CRISP, M.D., COOK, L.G., 2012. Phylogenetic niche conservatism: What are the underlying evolutionary and ecological causes? *New Phytol.* **196**, 681–694.

CUSIMANO, N., BARRETT, M.D., HETTERSCHEID, W.L.A., RENNER, S.S., 2010. A phylogeny of the areae (Araceae) implies that typhonium, sauromatum, and the australian species of typhonium are distinct clades. *Taxon.* **59**, 439–447.

DALE, E.E., LARCOMBE, M.J., LEE, W.G., 2021. The effect of single biome occupancy on the estimation of biome shifts and the detection of biome conservatism. *PLoS One.* **16**, 1–15.

DARRIBA, D., TABOADA, G.L., DOALLO, R., POSADA, D., 2015. Europe PMC Funders Group jModelTest 2 : more models , new heuristics and high- performance computing. *Nat. Methods.* **9**, 6–9.

DAYRAT, B., 2005. Towards integrative taxonomy. *Biol. J. Linn. Soc.* **85**, 407–415.

DE JAGER, M.L., ELLIS, A.G., 2017. Evolutionary history of a keystone pollinator parallels the biome occupancy of angiosperms in the Greater Cape Floristic Region. *Mol.*

Phylogenet. Evol. **107**, 530–537.

DE JAGER, M.L., NEWMAN, E., THERON, G., BOTHA, P., BARTON, M., ANDERSON, B., 2016. Pollinators can prefer rewarding models to mimics: consequences for the assumptions of Batesian floral mimicry. *Plant Syst. Evol.* **302**, 409–418.

DEVOTO, M., MEDAN, D., 2006. Diversity, distribution and floral specificity of tangle-veined flies (Diptera: Nemestrinidae) in north west Patagonia, Argentina. *Rev. Chil. Hist. Nat.* **79**, 29–40.

DINIZ-FILHO, J.A.F., LOYOLA, R.D., RAIA, P., MOOERS, A.O., BINI, L.M., 2013. Darwinian shortfalls in biodiversity conservation. *Trends Ecol. Evol.* **28**, 689–695.

DONOGHUE, M.J., EDWARDS, E.J., 2014. Biome Shifts and Niche Evolution in Plants. *Annu. Rev. Ecol. Evol. Syst.* **45**, 547–572.

DOYLE, J.A., DONOGHUE, M.J., 1993. Phylogenies and Angiosperm Diversification. *Paleobiology*. **19**, 141–167.

DRUMMOND, A.J., HO, S.Y.W., PHILLIPS, M.J., RAMBAUT, A., 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, 699–710.

DRUMMOND, A.J., SUCHARD, M.A., XIE, D., RAMBAUT, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973.

DUFFY, K.J., JOHNSON, S.D., 2017. Specialized mutualisms may constrain the geographical distribution of flowering plants. *Proc. R. Soc. B Biol. Sci.* **284**, 20171841.

DYNESIUS, M., JANSSON, R., 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 9115–9120.

EASTWOOD, R., PIERCE, N.E., KITCHING, R.L., HUGHES, J.M., 2006. Do Ants

Enhance Diversification in Lycaenid Butterflies? Phylogeographic Evidence From a Model Myrmecophile, Jalmenus Evagoras. *Evolution.* **60**, 315–327.

EDWARDS, E.J., DONOGHUE, M.J., 2013. Is it easy to move and easy to evolve? Evolutionary accessibility and adaptation. *J. Exp. Bot.* **64**, 4047–4052.

ELLIS, A.G., ANTHONY VERBOOM, G., VAN DER NIET, T., JOHNSON, S.D., PETER LINDER, H., 2014. Speciation and extinction in the Greater Cape Floristic Region, in: Fynbos. Oxford University Press, pp. 119–141.

EVENHUIS, N.L., 2020. The insect and spider collections of the world website [WWW Document]. URL http://hbs.bishopmuseum.org/ codens/

FAITH, D.P., 1992. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **61**, 1–10.

FERNÁNDEZ-MAZUECOS, M., BLANCO-PASTOR, J.L., JUAN, A., CARNICERO, P., FORREST, A., ALARCÓN, M., VARGAS, P., GLOVER, B.J., 2019. Macroevolutionary dynamics of nectar spurs, a key evolutionary innovation. *New Phytol.* **222**, 1123–1138.

FITZJOHN, R.G., MADDISON, W.P., OTTO, S.P., 2009. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst. Biol.* **58**, 595–611.

FONTANETO, D., FLOT, J.F., TANG, C.Q., 2015. Guidelines for DNA taxonomy, with a focus on the meiofauna. *Mar. Biodivers.* **45**, 433–451.

FORBES, A.A., BAGLEY, R.K., BEER, M.A., HIPPEE, A.C., WIDMAYER, H.A., 2018. Quantifying the unquantifiable: Why Hymenoptera, not Coleoptera, is the most speciose animal order. *BMC Ecol.* **18**, 1–11.

FOREST, F., 2009. Calibrating the tree of life: Fossils, molecules and evolutionary timescales. *Ann. Bot.* **104**, 789–794.

FRECKLETON, R.P., HARVEY, P.H., PAGEL, M., 2002. Phylogenetic analysis and comparative data: A test and review of evidence. *Am. Nat.* **160**, 712–726.

FUJITA, M.K., LEACHÉ, A.D., BURBRINK, F.T., MCGUIRE, J.A., MORITZ, C., 2012. Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecol. Evol.* **27**, 480–488.

FUNK, D.J., OMLAND, K.E., 2003. Species-Level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. *Annu. Rev. Ecol. Evol. Syst.* **34**, 397–423.

GALLEY, C., BYTEBIER, B., BELLSTEDT, D.U., LINDER, H.P., 2007. The Cape element in the Afrotemperate flora: From Cape to Cairo? *Proc. R. Soc. B Biol. Sci.* **274**, 535–543.

GARCÍA-ROBLEDO, C., KUPREWICZ, E.K., BAER, C.S., CLIFTON, E., HERNÁNDEZ, G.G., WAGNER, D.L., 2020. The Erwin equation of biodiversity: From little steps to quantum leaps in the discovery of tropical insect diversity. *Biotropica.* **52**, 590–597.

GEERTS, S., PAUW, A., 2009. Hyper-specialization for long-billed bird pollination in a guild of South African plants: the Malachite Sunbird pollination syndrome. *South African J. Bot.* **75**, 699–706.

GLOR, R.E., 2010. Phylogenetic insights on adaptive radiation. *Annu. Rev. Ecol. Evol. Syst.* **41**, 251–270.

GOLDBLATT, P., BERNHARDT, P., MANNING, J.C., 2000. Adaptive Radiation of Pollination Mechanisms in *Ixia* (Iridaceae: Crocoideae). *Ann. Missouri Bot. Gard.* **87**, 564.

GOLDBLATT, P., MANNING, J.C., 2007. Pollination of *Romulea syringodeoflora* (Iridaceae: Crocoideae) by a long-proboscid fly, *Prosoeca sp*. (Diptera: Nemestrinidae). *South African J. Bot.* **73**, 56–59.

GOLDBLATT, P., NÄNNI, I., BERNHARDT, P., MANNING, J.C., 2004. Floral Biology of *Hesperantha* (Iridaceae: Crocoideae): How Minor Shifts in Floral Presentation Change the Pollination System. *Ann. Missouri Bot. Gard.* **91**, 186–206.

GOSZ, J.R., 1992. Ecological Functions in a Biome Transition Zone: Translating Local Responses to Broad-Scale Dynamics, in: Hansen, A.J., di Castri, F. (Eds.), Landscape Boundaries: Consequenecs for Biotic Diversity and Ecological Flows. Springer, pp. 55–75.

GOULD, S.J., 1966. Allometry and size in ontogeny and phylogeny. *Biol. Rev. Camb. Philos. Soc.* **41**, 587–640.

GOULD, S.J., 1971. Geometric Similarity in Allometric Growth: A Contribution to the Problem of Scaling in the Evolution of Size. *Am. Nat.* **105**, 113–136.

GREATHEAD, D.J., 1967. The genus *Trichopsidea* Westwood, with a discussion of its relation to other genera of Nemestrinidae (Diptera). *J. Nat. Hist.* **1**, 305–313.

HABER, W.A., FRANKIE, G.W., 1989. A Tropical Hawkmoth Community: Costa Rican Dry Forest Sphingidae. *Biotropica.* **21**, 155–172.

HAENNI, J.P., BORER, M., 2007. First report of a Nemestrinidae (Diptera) parasitoid of the Mantodea. *Stud. dipterologica.* **14**, 61–65.

HALL, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*

HANSEN, D.M., VAN DER NIET, T., JOHNSON, S.D., 2012. Floral signposts: Testing the

significance of visual “nectar guides” for pollinator behaviour and plant fitness. *Proc. R. Soc. B Biol. Sci.* **279**, 634–639.

HARDIN, G., 1960. The Competitive Exclusion Principle. *Science.* **131**, 1292–1297.

HARVEY, P., PAGEL, M., 1992. The comparative method in evolutionary biology. Oxford University Press.

HAVERKAMP, A., BING, J., BADEKE, E., HANSSON, B.S., KNADEN, M., 2016. Innate olfactory preferences for flowers matching proboscis length ensure optimal energy gain in a hawkmoth. *Nat. Commun.* **7**, 11644.

HAWKINS, B.A., PORTER, E.E., 2003. Does herbivore diversity depend on plant diversity? The case of California butterflies. *Am. Nat.* **161**, 40–49.

HAYES, K.A., 2021. Taxonomic shortcuts lead to long delays in species discovery, delineation, and identification. *Biol. Invasions.* **8**.

HEBERT, P.D.N., CYWINSKA, A., BALL, S.L., DEWAARD, J.R., 2003a. Biological identifications through DNA barcodes. *Proc. R. Soc. B Biol. Sci.* **270**, 313–321.

HEBERT, P.D.N., PENTON, E.H., BURNS, J.M., JANZEN, D.H., HALLWACHS, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci.* **101**, 14812–14817.

HEBERT, P.D.N., RATNASINGHAM, S., DEWAARD, J.R., 2003b. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. B Biol. Sci.* **270**, 96–99.

HEBERT, P.D.N., RATNASINGHAM, S., ZAKHAROV, E. V., TELFER, A.C., LEVESQUE-BEAUDIN, V., MILTON, M.A., PEDERSEN, S., JANNETTA, P., DEWAARD, J.R., 2016. Counting animal species with DNA barcodes: Canadian

insects. *Philos. Trans. R. Soc. B Biol. Sci.* **371**.

HENRY, C.S., 1994. Singing and cryptic speciation insects. *Trends Ecol. Evol.* **9**, 388–392.

HERRE, E.A., 2006. Barcoding helps biodiversity fly. *Proc. Natl. Acad. Sci.* **103**, 3949–3950.

HODGES, S.A., ARNOLD, M.L., 1995. Spurring plant diversification: are floral nectar spurs a key innovation? *Proc. R. Soc. London. Ser. B Biol. Sci.* **262**, 343–348.

HOFFMAN, M., KOENIG, K., BUNTING, G., COSTANZA, J., WILLIAMS, K.J., 2016. Biodiversity Hotspots (version 2016.1) [dataset]. Zanodo.

HOLSTEIN, N., RENNER, S.S.S., WIENS, J., GRAHAM, C., CRISP, M., ARROYO, M., COOK, L., GANDOLFO, M., JORDAN, G., MCGLONE, M., WESTON, P., WESTOBY, M., WILF, P., LINDER, H., PETERSON, A., SOBERÓN, J., SÁNCHEZ-CORDERO, V., DAUBY, G., DUMINIL, J., HEUERTZ, M., HARDY, O., BORN, C., ALVAREZ, N., MCKEY, D., OSSARI, S., WICKINGS, E., HOSSAERT-MCKEY, M., CHEVALLIER, M.-H., LOWE, A., HARRIS, D., DORMONTT, E., DAWSON, I., DEBOUT, G., DOUCET, J., HARDY, O., MUCHUGI, A., LENGKEEK, A., KADU, C., MULUVI, G., NJAGI, E., DAWSON, I., DAVIS, C., BELL, C., FRITSCH, P., MATHEWS, S., COUVREUR, T., CHATROU, L., SOSEF, M., RICHARDSON, J., KIER, G., MUTKE, J., DINERSTEIN, E., RICKETTS, T., KÜPER, W., KREFT, H., BARTHLOTT, W., KÜPER, W., SOMMER, J., LOVETT, J., BARTHLOTT, W., PICKFORD, M., KINGSTON, J., HILL, A., MARINO, B., BAKKER, E. van Z., MERCER, J., AXELROD, D., RAVEN, P., JACOBS, B., CERLING, T., HARRIS, J., MACFADDEN, B., LEAKEY, M., QUADE, J., EISENMANN, V., EHLERINGER, J., SEPULCHRE, P., RAMSTEIN, G., FLUTEAU, F., SCHUSTER, M., TIERCELIN, J.-J., BRUNET, M., ZACHOS, J., PAGANI, M., SLOAN, L., THOMAS, E., BILLUPS,

K., DEMENOCAL, P., DUPONT, L., JAHNS, S., MARRET, F., NING, S., GASSE, F., ELENGA, H., MALEY, J., VINCENS, A., FARRERA, I., MALEY, J., COLYN, M., GAUTIER-HION, A., VERHEYEN, W., SOSEF, M., LEAL, M., MARRET, F., MALEY, J., SCOURSE, J., MALEY, J., JEFFREY, C., ZIMMERMANN, A., COWLING, S., COX, P., JONES, C., MASLIN, M., PEROS, M., SPALL, S., FJELDSÅ, J., LOVETT, J., SALZMANN, U., HOELZMANN, P., BHATT, D., KUMAR, A., BLEHER, B., POTGIETER, C., JOHNSON, D., BÖHNING-GAESE, K., WALTER, H., BRECKLE, S.-W., LORENZEN, E., MASEMBE, C., ARCTANDER, P., SIEGISMUND, H., FREITAG, S., ROBINSON, T., ARAYA, Y., SILVERTOWN, J., GOWING, D., MCCONWAY, K., LINDER, H., MIDGLEY, G., HIJMANS, R., CAMERON, S., PARRA, J., JONES, P., JARVIS, A., GAUSSEN, H., MADDISON, W., MADDISON, D., LEWIS, P., HOLSTEIN, N., RENNER, S.S.S., KOCYAN, A., ZHANG, L.-B., SCHAEFER, H., RENNER, S.S.S., SCHAEFER, H., HEIBL, C., RENNER, S.S.S., HORNE, E., KUMPATLA, S., PATTERSON, K., KUPTA, M., THOMPSON, S., HAMILTON, M., SHAW, J., LICKEY, E., SCHILLING, E., SMALL, R., YOKOYAMA, J., SUZUKI, M., IWATSUKI, K., HASEBE, M., TABERLET, P., GIELLY, L., PAUTOU, G., BOUVET, J., VOLZ, S., RENNER, S.S.S., SCHUSTER, K., BALTHAZAR, M., ENDRESS, P., QIU, Y.-L., INOUE, H., NOJIMA, H., OKAYAMA, H., MADDISON, W., MADDISON, D., STAMATAKIS, A., LUDWIG, T., MEIER, H., POSADA, D., CRANDALL, K., STAMATAKIS, A., HOOVER, P., ROUGEMONT, J., DRUMMOND, A., RAMBAUT, A., 2011. A dated phylogeny and collection records reveal repeated biome shifts in the African genus *Coccinia* (Cucurbitaceae). *BMC Evol. Biol.* **11**, 28.

HRDINA, A., ROMPORTL, D., 2017. Evaluating Global Biodiversity Hotspots – Very Rich and Even More Endangered. *J. Landsc. Ecol.* **10**.

HTTP://BGIS.SANBI.ORG/SPATIALDATASET/DETAIL/669, 2011. South African National Parks. CCAB - Current Biome Delineations 2011 [vector geospatial dataset]. *Available from Biodivers. GIS website*.

HUXLEY, J.S., 1924. Constant Differential Growth-Ratios and their Significance. *Nature.* **114**, 895–896.

IVES, A.R., MIDFORD, P.E., GARLAND, T., 2007. Within-species variation and measurement error in phylogenetic comparative methods. *Syst. Biol.* **56**, 252–270.

JINBO, U., KATO, T., ITO, M., 2011. Current progress in DNA barcoding and future implications for entomology. *Entomol. Sci.* **14**, 107–124.

JOHNSON, J.A., WATSON, R.T., MINDELL, D.P., 2005. Prioritizing species conservation: does the Cape Verde kite exist? *Proceeding R. Soc. B* **272**, 1365–71.

JOHNSON, S.D., 2006a. Pollinator-driven speciation in plants, in: Harder, L., Barrett, S.C.H. (Eds.), Ecology and Evolution of Flowers. Oxford University Press Inc., New York, pp. 297–301.

JOHNSON, S.D., 2006b. Pollination by long-proboscid flies in the endangered African orchid *Disa scullyi*. *South African J. Bot.* **72**, 24–27.

JOHNSON, S.D., LINDER, H.P., 1995. Systematics and evolution of the *Disa draconis* complex (Orchidaceae). *Bot. J. Linn. Soc.* **118**, 289–307.

JOHNSON, S.D., MORE, M., AMORIM, F.W., HABER, W.A., FRANKIE, G.W., STANLEY, D.A., COCUCCI, A.A., RAGUSO, R.A., MORE, M., AMORIM, F.W., HABER, W.A., FRANKIE, G.W., STANLEY, D.A., COCUCCI, A.A., RAGUSO, R.A., 2017. The long and the short of it: A global analysis of hawkmoth pollination niches and interaction networks. *Funct. Ecol.* **31**, 101–115.

JONES, O.R., PURVIS, A., BAUMGART, E., QUICKE, D.L.J., 2009. Using taxonomic revision data to estimate the geographic and taxonomic distribution of undescribed species richness in the Braconidae (Hymenoptera: Ichneumonoidea). *Insect Conserv. Divers.* **2**, 204–212.

JOY, J.B., LIANG, R.H., MCCLOSKEY, R.M., NGUYEN, T., POON, A.F.Y., 2016. Ancestral Reconstruction. *PLoS Comput. Biol.* **12**, 1–20.

KAHNT, B., MONTGOMERY, G.A., MURRAY, E., KUHLMANN, M., PAUW, A., MICHEZ, D., PAXTON, R.J., DANFORTH, B.N., 2017. Playing with extremes: Origins and evolution of exaggerated female forelegs in South African *Rediviva* bees. *Mol. Phylogenet. Evol.* **115**, 95–105.

KAMIYA, T., O'DWYER, K., NAKAGAWA, S., POULIN, R., 2014. Host diversity drives parasite diversity: Meta-analytical insights into patterns and causal mechanisms. *Ecography (Cop.).* **37**, 689–697.

KANMIYA, K., 1987. New records of endoparasitic Nemestrinidae and Tachinidae (Diptera) on Tettigoniidae. *Makunagi.* **15**, 1–4.

KAROLYI, F., MORAWETZ, L., COLVILLE, J.F., HANDSCHUH, S., METSCHER, B.D., KRENN, H.W., 2013. Time management and nectar flow: Flower handling and suction feeding in long-proboscid flies (Nemestrinidae: *Prosoeca*). *Naturwissenschaften.* **100**, 1083–1093.

KAROLYI, F., SZUCSICH, N.U., COLVILLE, J.F., KRENN, H.W., 2012. Adaptations for nectar-feeding in the mouthparts of long-proboscid flies (Nemestrinidae: *Prosoeca*). *Biol. J. Linn. Soc.* **107**, 414–424.

KAWAKITA, A., SOTA, T., ITO, M., ASCHER, J.S., TANAKA, H., KATO, M., ROUBIK, D.W., 2004. Phylogeny, historical biogeography, and character evolution in bumble bees

(Bombus: Apidae) based on simultaneous analysis of three nuclear gene sequences. *Mol. Phylogenet. Evol.* **31**, 799–804.

KEMBEL, S.W., COWAN, P.D., HELMUS, M.R., CORNWELL, W.K., MORLON, H., ACKERLY, D.D., BLOMBERG, S.P., WEBB, C.O., 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics.* **26**, 1463–1464.

KEMP, J.E., ELLIS, A.G., 2017. Significant local-scale plant-insect species richness relationship independent of abiotic effects in the temperate cape floristic region biodiversity hotspot. *PLoS One.* **12**, 1–16.

KERDELHUE, C., LE CLAINCHE, I., RASPLUS, J.Y., 1999. Molecular Phylogeny of the *Ceratosolen* Species Pollinating Ficus of the Subgenus *Sycomorus sensu stricto*: Biogeographical History and Origins of the Species-Specificity Breakdown Cases. *Mol. Phylogenet. Evol.* **11**, 401–414.

KIM, S. Il, FARRELL, B.D., 2015. Phylogeny of world stag beetles (Coleoptera: Lucanidae) reveals a Gondwanan origin of Darwin's stag beetle. *Mol. Phylogenet. Evol.* **86**, 35–48.

KLIMOV, P.B., MIRONOV, S. V., OCONNOR, B.M., 2017. Convergent and unidirectional evolution of extremely long aedeagi in the largest feather mite genus, *Proctophyllodes* (Acari: Proctophyllodidae): Evidence from comparative molecular and morphological phylogenetics. *Mol. Phylogenet. Evol.* **114**, 212–224.

KLINGENBERG, C.P., 2005. Developmental Constraints, Modules, and Evolvability, in: Hallgrímsson, B., Hall, B.K. (Eds.), Variation. Academic Press, Burlington, pp. 219–247.

KLUMPERS, S.G.T., STANG, M., KLINKHAMER, P.G.L., 2019. Foraging efficiency and size matching in a plant–pollinator community: the importance of sugar content and tongue length. *Ecol. Lett.* **22**, 469–479.

KNOWLTON, N., JACKSON, J.B.C., 1994. New taxonomy and niche partitioning on coral reefs: jack of all trades or master of some? *Trends Ecol. Evol.* **9**, 7–9.

KOZLOV, M. V, ZHU, J., PHILIPP, P., FRANCKE, W., ZVEREVA, E.L., HANSSON, B.S., LÖFSTEDT, C., 1996. Pheromone specificity in *Eriocrania semipurpurella* (Stephens) and *E. sangii* (Wood) (Lepidoptera: Eriocraniidae) based on chirality of semiochemicals. *J. Chem. Ecol.* **22**, 431–454.

KRENN, H.W., 2019. Insect Mouthparts, Zoological Monographs. Springer International Publishing, Cham.

KRENN, H.W., PLANT, J.D., SZUCSICH, N.U., 2005. Mouthparts of flower-visiting insects. *Athropod Struct. Devvelopment.* **34**, 1–40.

KUMAR, S., STECHER, G., LI, M., KNYAZ, C., TAMURA, K., 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549.

KUNTE, K., 2007. Allometry and functional constraints on proboscis lengths in butterflies. *Funct. Ecol.* **21**, 982–987.

LABANDEIRA, C.C., KVACEK, J., MOSTOVSKI, M.B., 2007. Pollination drops, pollen, and insect pollination of Mesozoic gymnosperms. *Taxon.* **56**, 663–695.

LABARBERA, M., 1989. Analyzing body size as a factor in ecology and evolution. *Annu. Rev. Ecol. Syst.* **20**, 97–117.

LANDIS, M.J., MATZKE, N.J., MOORE, B.R., HUELSENBECK, J.P., 2013. Bayesian analysis of biogeography when the number of areas is large. *Syst. Biol.* **62**, 789–804.

LANFEAR, R., FRANDSEN, P.B., WRIGHT, A.M., SENFELD, T., CALCOTT, B., 2017. Partitionfinder 2: New methods for selecting partitioned models of evolution for

molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* **34**, 772–773.

LAVER, R.J., NIELSEN, S. V., ROSAUER, D.F., OLIVER, P.M., 2017. Trans-biome diversity in Australian grass-specialist lizards (Diplodactylidae: *Strophurus*). *Mol. Phylogenet. Evol.* **115**, 62–70.

LEE, M.S.Y., 1999. Molecular clock calibrations and metazoan divergence dates. *J. Mol. Evol.* **49**, 385–391.

LEIGH, J.W., BRYANT, D., 2015. POPART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116.

LI, X., TEASDALE, L.C., BAYLESS, K.M., ELLIS, A.G., WIEGMANN, B.M., LAMAS, C.J.E., LAMBKIN, C.L., EVENHUIS, N.L., NICHOLLS, J.A., HARTLEY, D., SHIN, S., TRAUTWEIN, M., ZWICK, A., LESSARD, B.D., YEATES, D.K., 2021. Phylogenomics reveals accelerated late Cretaceous diversification of bee flies (Diptera: Bombyliidae). *Cladistics.* **37**, 276–297.

LIBERLES, D.A. (Ed.), 2007. Ancestral Sequence Reconstruction. Oxford University Press.

LICHTWARDT B. 1920. Über afrikanische Nemestriniden (Dipt.). Entomologische Mitteilungen 9:97–101.

LINDER, H.P., 2003. The radiation of the Cape flora, southern Africa. *Biol. Rev. Camb. Philos. Soc.* **78**, 597–638.

LINDER, H.P., 2005. Evolution of diversity: the Cape flora. *Trends Plant Sci.* **10**, 536–41.

LINDER, H.P., VERBOOM, G.A., 2015. The Evolution of Regional Species Richness: The History of the Southern African Flora. *Annu. Rev. Ecol. Evol. Syst.* **46**, 393–412.

LIU, Y., HUANG, D., 2019a. New materials of Nemestrinidae (Diptera: Brachycera) from the mid-Cretaceous Burmese amber. *Cretac. Res.* **104**, 104173.

LIU, Y., HUANG, D., 2019b. A large new nemestrinid fly from the Lower Cretaceous Yixian Formation at Liutiaogou, Ningcheng County, Inner Mongolia, NE China. *Cretac. Res.* **96**, 107–112.

MADDISON, W.P., 1995. Calculating the probability distributions of ancestral states reconstructed by parsimony on phylogenetic trees. *Syst. Biol.* **44**, 474–481.

MAECHLER, M., 2004. The diptest Package. *Compute.* 1–5.

MAGLIANESI, M. a., BÖHNING-GAESE, K., SCHLEUNING, M., 2015. Different foraging preferences of hummingbirds on artificial and natural flowers reveal mechanisms structuring plant-pollinator interactions. *J. Anim. Ecol.* **84**, 655–664.

MANNING, J.C., GOLDBLATT, P., 1996. The *Prosoeca peringueyi* (Diptera: Nemestrinidae) pollination guild in southern Africa: long-tongued flies and their tubular flowers. *Ann. Missouri Bot. Gard.* **83**, 67–86.

MANNING, J.C., GOLDBLATT, P., 1997. The *Moegistorhynchus longirostris* (Diptera: Nemestrinidae) pollination guild: Long-tubed flowers and a specialized long-proboscid fly pollination system in southern Africa. *Plant Syst. Evol.* **206**, 51–69.

MANNING, J.C., GOLDBLATT, P., 2000. The Long-Proboscid Fly Pollination System in Southern Africa. *Ann. Missouri Bot. Gard.* **87**, 146–170.

MANNING, J.C., GOLDBLATT, P., 2012. Plants of the Greater Cape Floristic Region 1: The Core Cape Subregion, Strelitzia. Strelitzia 29, South African National Biodiversity Institute, Pretoria.

MARTINS, D.J., JOHNSON, S.D., 2013. Interactions between hawkmoths and flowering plants in east africa: Polyphagy and evolutionary specialization in an ecological context. *Biol. J. Linn. Soc.* **110**, 199–213.

MARTINS, E.P., HANSEN, T.F., 1997. Phylogenies and the comparitive method. *Am. Nat.* **149**, 646–667.

MATZKE, N.J., 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Front. Biogeogr.* **5**, 242–248.

MAYHEW, P.J., 2007. Why are there so many insect species? Perspectives from fossils and phylogenies. *Biol. Rev.* **82**, 425–454.

MAZIJK, R., CRAMER, M.D., VERBOOM, G.A., 2021. Environmental heterogeneity explains contrasting plant species richness between the South African Cape and southwestern Australia. *J. Biogeogr.* jbi.14118.

MENGUAL, X., STÅHLS, G., ROJO, S., 2008. First phylogeny of predatory flower flies (Diptera, Syrphidae, Syrphinae) using mitochondrial COI and nuclear 28S rRNA genes: Conflict and congruence with the current tribal classification. *Cladistics.* **24**, 543–562.

MICHALOUD, G., CARRIERE, S., KOBBI, M., 1996. Exceptions to the One : One Relationship Between African Fig Trees and Their Fig Wasp Pollinators: Possible Evolutionary Scenarios. *J. Biogeogr.* **23**, 513–520.

MILLER, J.T., JOLLEY-ROGERS, G., MISHLER, B.D., THORNHILL, A.H., 2018. Phylogenetic diversity is a better measure of biodiversity than taxon counting. *J. Syst. Evol.* **56**, 663–667.

MILLER, M.A., PFEIFFER, W., SCHWARTZ, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gatew. Comput. Environ. Work. GCE 2010*.

MINNAAR, C., ANDERSON, B., DE JAGER, M.L., KARRON, J.D., 2019a. Plant-pollinator interactions along the pathway to paternity. *Ann. Bot.* **123**, 225–245.

MINNAAR, C., DE JAGER, M.L., ANDERSON, B., 2019b. Intraspecific divergence in floral-tube length promotes asymmetric pollen movement and reproductive isolation. *New Phytol.* **224**, 1160–1170.

MISHLER, B.D., DONOGHUE, M.J., 1982. Species Concepts: A Case for Pluralism. *Syst. Zool.* **31**, 491.

MITCHELL, K.J., PRATT, R.C., WATSON, L.N., GIBB, G.C., LLAMAS, B., KASPER, M., EDSON, J., HOPWOOD, B., MALE, D., ARMSTRONG, K.N., MEYER, M., HOFREITER, M., AUSTIN, J., DONNELLAN, S.C., LEE, M.S.Y., PHILLIPS, M.J., COOPER, A., 2014. Molecular phylogeny, biogeography, and habitat preference evolution of Marsupials. *Mol. Biol. Evol.* **31**, 2322–2330.

MOLBO, D., MACHADO, C.A., SEVENSTER, J.G., KELLER, L., HERRE, E.A., 2003. Cryptic species of fig-pollinating wasps: Implications for the evolution of the fig-wasp mutualism, sex allocation, and precision of adaptation. *Proc. Natl. Acad. Sci.* **100**, 5867–5872.

MORINIÈRE, J., VAN DAM, M.H., HAWLITSCHEK, O., BERGSTEN, J., MICHAT, M.C., HENDRICH, L., RIBERA, I., TOUSSAINT, E.F.A., BALKE, M., 2016. Phylogenetic niche conservatism explains an inverse latitudinal diversity gradient in freshwater arthropods. *Sci. Rep.* **6**, 1–12.

MORITA, S.I., 2011. Repeatability and precision in proboscis length measurements for long proboscid flies. *Zootaxa*. **58**, 49–58.

MORRONE, J.J., CRISCI, J. V., 1995. Historical biogeography: Introduction to methods. *Annu. Rev. Ecol. Syst.* **26**, 373–401.

MOSTOVSKI, M.B., 1998. A Revision of the Nemestrinid Flies (Diptera, Nemestrinidae) Described by Rohdendorf, and a Deascription of New Taxa of the Nemestrinidae from

the Upper Jurassic of Kazakhstan. *Paleontogical J.* **32**, 369–375.

MOSTOVSKI, M.B., MARTÍNEZ-DELCLÒS, X., 2000. New Nemestrinoidea (Diptera: Brachycera) from the Upper Jurassic - Lower Cretaceous of Eurasia, taxonomy, and palaeobiology. *Entomol. Probl.* **31**, 137–148.

MOULTON, J.K., WIEGMANN, B.M., 2004. Evolution and phylogenetic utility of CAD (rudimentary) among Mesozoic-aged Eremoneuran Diptera (Insecta). *Mol. Phylogenet. Evol.* **31**, 363–378.

MUCINA, L., RUTHERFORD, M.C., 2006. The vegetation of South Africa, Lesotho and Swaziland. *Strelitzia. 19* 1–30.

MÜNKEMÜLLER, T., LAVERGNE, S., BZEZNIK, B., DRAY, S., JOMBART, T., SCHIFFERS, K., THUILLER, W., 2012. How to measure and test phylogenetic signal. *Methods Ecol. Evol.* **3**, 743–756.

MYERS, N., MITTERMEIER, R.A., MITTERMEIER, C.G., DA FONSECA, G.A.B., KENT, J., 2000. Biodiversity hotspots for conservation priorities. *Nature.* **403**, 853–858.

NEW, T.R., 1999. Untangling the web: Spiders and the challenges of invertebrate conservation. *J. Insect Conserv.* **3**, 251–256.

NEWMAN, E., MANNING, J.C., ANDERSON, B., 2014. Matching floral and pollinator traits through guild convergence and pollinator ecotype formation. *Ann. Bot.* **113**, 373–84.

NICHOLSON, D.B., ROSS, A.J., MAYHEW, P.J., 2014. Fossil evidence for key innovations in the evolution of insect diversity. *Proc. R. Soc. B Biol. Sci.* **281**.

NILSSON, S., GRAFSTRÖM, E., COETZEE, J., 1996. On the origin of the sarcolaenaceae with reference to pollen morphological evidence. *Grana.* **35**, 321–334.

OLLERTON, J., JOHNSON, S.D., HINGSTON, A.B., 2006. Geographical Variation in Diversity and Specificity of Pollination Systems, in: Ollerton, J., Waser, N.M. (Eds.), Plant-Pollinator Interations: From Specialization to Generalization. University of Chicago Press, Chicago and London, pp. 282–309.

ORME, D., 2013. The caper package : comparative analysis of phylogenetics and evolution in R. *R Packag. version 0.5, 2* 1–36.

ORR, M.C., HUGHES, A.C., CHESTERS, D., PICKERING, J., ZHU, C.D., ASCHER, J.S., 2021. Global Patterns and Drivers of Bee Distribution. *Curr. Biol.* **31**, 451-458.e4.

PAGEL, M., 1999. Inferring the historical patterns of biological evolution. *Nature.* **401**, 877–884.

PAPADOPOULOU, A., ANASTASIOU, I., VOGLER, A.P., 2010. Revisiting the insect mitochondrial molecular clock: The mid-aegean trench calibration. *Mol. Biol. Evol.* **27**, 1659–1672.

PARADIS, E., CLAUDE, J., STRIMMER, K., 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics.* **20**, 289–290.

PARINS-FUKUCHI, C., 2018. Use of Continuous Traits Can Improve Morphological Phylogenetics. *Syst. Biol.* **67**, 328–339.

PAUW, A., 2007. Collapse of a pollination web in small conservation areas. *Ecology.* **88**, 1759–1769.

PAUW, A., 2013. Can pollination niches facilitate plant coexistence? *Trends Ecol. Evol.* **28**, 30–7.

PAUW, A., COCUCCI, A.A., SÉRSIC, A.N., 2020. The least effective pollinator principle: specialized morphology despite generalized ecology. *Plant Biol.* **22**, 924–931.

PAUW, A., STOFBERG, J., WATERMAN, R.J., 2009. Flies and Flowers in Darwin'S Race. *Evolution.* **63**, 268–279.

PÉLABON, C., ARMBRUSTER, W.S., HANSEN, T.F., 2011. Experimental evidence for the Berg hypothesis: Vegetative traits are more sensitive than pollination traits to environmental variation. *Funct. Ecol.* **25**, 247–257.

PENNELL, M.W., EASTMAN, J.M., SLATER, G.J., BROWN, J.W., UYEDA, J.C., FITZJOHN, R.G., ALFARO, M.E., HARMON, L.J., 2014. Geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics.* **30**, 2216–2218.

PHILLIPS, M.J., BENNETT, T.H., LEE, M.S.Y., 2009. Molecules, morphology, and ecology indicate a recent, amphibious ancestry for echidnas. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 17089–17094.

PIRIE, M.D., KANDZIORA, M., NÜRK, N.M., LE MAITRE, N.C., MUGRABI DE KUPPLER, A., GEHRKE, B., OLIVER, E.G.H., BELLSTEDT, D.U., 2019. Leaps and bounds: geographical and ecological distance constrained the colonisation of the Afrotemperate by Erica. *BMC Evol. Biol.* **19**, 1–12.

POTGIETER, C.J., EDWARDS, T.J., 2005. The *Stenobasipteron wiedmanni* (Diptera, Nemestrinidae) pollination guild in Eastern Southern Africa. *Ann. Missouri Bot. Gard.* **92**, 254–267.

POTGIETER, C.J., EDWARDS, T.J., VAN STADEN, J., 2009. Pollination of *Plectranthus* spp. (Lamiaceae) with sigmoid flowers in southern Africa. *South African J. Bot.* **75**, 646–659.

POTGIETER, J.T., 1929. A Contribution to the Biology of the Brown Swarm Locust *Locustana pardalina* (Wlk.) and its natural Enemies. *Pan-Afr. Agric. Vet. Conf.*

Pretoria, Un. S. Afr. Dept. Agric. 265–308.

POWER, S.C., ANTHONY VERBOOM, G., BOND, W.J., CRAMER, M.D., 2017. Environmental correlates of biome-level floristic turnover in South Africa. *J. Biogeogr.* **44**, 1745–1757.

PRESCOTT, H.W., 1960. Suppression of Grasshoppers by Nemestrinid Parasites (Diptera). *Ann. Entomol. Soc. Am.* **53**, 513–521.

PROCHEŞ, Ş., FOREST, F., VELDTMAN, R., CHOWN, S.L., COWLING, R.M., JOHNSON, S.D., RICHARDSON, D.M., SAVOLAINEN, V., 2009. Dissecting the plant-insect diversity relationship in the Cape. *Mol. Phylogenet. Evol.* **51**, 94–99.

PUILLANDRE, N., LAMBERT, A., BROUILLET, S., ACHAZ, G., 2012a. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* **21**, 1864–1877.

PUILLANDRE, N., MODICA, M. V., ZHANG, Y., SIROVICH, L., BOISSELIER, M.C., CRUAUD, C., HOLFORD, M., SAMADI, S., 2012b. Large-scale species delimitation method for hyperdiverse groups. *Mol. Ecol.* **21**, 2671–2691.

QGIS DEVELOPMENT TEAM, 2014. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org. *Qgisorg*.

R CORE TEAM, 2017. R Development Core Team. *R A Lang. Environ. Stat. Comput.*

RABOSKY, D.L., GRUNDLER, M., ANDERSON, C., TITLE, P., SHI, J.J., BROWN, J.W., HUANG, H., LARSON, J.G., 2014. BAMMtools: An R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* **5**, 701–707.

RABOSKY, D.L., HUANG, H., 2016. A Robust Semi-Parametric Test for Detecting Trait-Dependent Diversification. *Syst. Biol.* **65**, 181–193.

RABOSKY, D.L., SANTINI, F., EASTMAN, J., SMITH, S.A., SIDLAUSKAS, B., CHANG, J., ALFARO, M.E., 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun.* **4**, 1–8.

RAMBAUT, A., DRUMMOND, A.J., XIE, D., BAELE, G., SUCHARD, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**, 901–904.

RAMÍREZ, S.R., ROUBIK, D.W., SKOV, C.E., PIERCE, N.E., 2010. Phylogeny, diversification patterns and historical biogeography of euglossine orchid bees (Hymenoptera: Apidae). *Biol. J. Linn. Soc.* **100**, 552–572.

RANGEL, T.F., COLWELL, R.K., GRAVES, G.R., FUČÍKOVÁ, K., RAHBEK, C., DINIZ-FILHO, J.A.F., 2015. Phylogenetic uncertainty revisited: Implications for ecological analyses. *Evolution.* **69**, 1301–1312.

RANTA, E., LUNDBERG, H., 1980. Resource Partitioning in Bumblebees: The Significance of Differences in Proboscis Length. *Oikos.* **35**, 298.

RAVAOARIMANANA, I.B., TIEDEMANN, R., MONTAGNON, D., RUMPLER, Y., 2004. Molecular and cytogenetic evidence for cryptic speciation within a rare endemic Malagasy lemur, the Northern Sportive Lemur (*Lepilemur septentrionalis*). *Mol. Phylogenet. Evol.* **31**, 440–448.

REDDY, S., DÁVALOS, L.M., 2003. Geographical sampling bias and its implications for conservation priorities in Africa. *J. Biogeogr.* **30**, 1719–1727.

REE, R.H., SMITH, S.A., 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* **57**, 4–14.

REEMER, M., STÅHLS, G., 2013. Phylogenetic relationships of Microdontinae (Diptera: Syrphidae) based on molecular and morphological characters. *Syst. Entomol.* **38**, 661–

688.

RENAUD, A.K., SAVAGE, J., ADAMOWICZ, S.J., 2012. DNA barcoding of Northern Nearctic Muscidae (Diptera) reveals high correspondence between morphological and molecular species limits. *BMC Ecol.* **12**, 1–15.

REVELL, L.J., 2012. phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223.

RONQUIST, F., 1997. Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Syst. Biol.* **46**, 195–203.

RONQUIST, F., 2004. Bayesian inference of character evolution. *Trends Ecol. Evol.* **19**, 475–481.

RONQUIST, F., SANMARTÍN, I., 2011. Phylogenetic Methods in Biogeography. *Annu. Rev. Ecol. Evol. Syst.* **42**, 441–464.

RONQUIST, F., TESLENKO, M., VAN DER MARK, P., AYRES, D.L., DARLING, A., HÖHNA, S., LARGET, B., LIU, L., SUCHARD, M.A., HUELSENBECK, J.P., 2012. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542.

RUTHERFORD, M., MUCINA, L., POWRIE, L., 2006. Biomes and bioregions of southern Africa. *Veg. South Africa Lesotho Sawzil.* 31–51.

SANDEL, B., ARGE, L., DALSGAARD, B., DAVIES, R.G., GASTON, K.J., SUTHERLAND, W.J., SVENNING, J.-C., 2011. The Influence of Late Quaternary Climate-Change Velocity on Species Endemism. *Science.* **334**, 660–664.

SANGSTER, G., 2009. Increasing numbers of bird species result from taxonomic progress, not taxonomic inflation. *Proc. R. Soc. B Biol. Sci.* **276**, 3184–3191.

SCHIPPER, J., CHANSON, J.S., CHIOZZA, F., COX, N.A., HOFFMANN, M., KATARIYA, V., LAMOREUX, J., RODRIGUES, A.S.L., STUART, S.N., TEMPLE, H.J., BAILLIE, J., BOITANI, L., LACHER, T.E., MITTERMEIER, R.A., SMITH, A.T., ABSOLON, D., AGUIAR, J.M., AMORI, G., BAKKOUR, N., BALDI, R., BERRIDGE, R.J., BIELBY, J., BLACK, P.A., BLANC, J.J., BROOKS, T.M., BURTON, J.A., BUTYNSKI, T.M., CATULLO, G., CHAPMAN, R., COKELISS, Z., COLLEN, B., CONROY, J., COOKE, J.G., DA FONSECA, G.A.B., DEROCHER, A.E., DUBLIN, H.T., DUCKWORTH, J.W., EMMONS, L., EMSLIE, R.H., FESTA-BIANCHET, M., FOSTER, M., FOSTER, S., GARSHELIS, D.L., GATES, C., GIMENEZ-DIXON, M., GONZALEZ, S., GONZALEZ-MAYA, J.F., GOOD, T.C., HAMMERSON, G., HAMMOND, P.S., HAPPOLD, D., HAPPOLD, M., HARE, J., HARRIS, R.B., HAWKINS, C.E., HAYWOOD, M., HEANEY, L.R., HEDGES, S., HELGEN, K.M., HILTON-TAYLOR, C., HUSSAIN, S.A., ISHII, N., JEFFERSON, T.A., JENKINS, R.K.B., JOHNSTON, C.H., KEITH, M., KINGDON, J., KNOX, D.H., KOVACS, K.M., LANGHAMMER, P., LEUS, K., LEWISON, R., LICHTENSTEIN, G., LOWRY, L.F., MACAVOY, Z., MACE, G.M., MALLON, D.P., MASI, M., MCKNIGHT, M.W., MEDELLÍN, R.A., MEDICI, P., MILLS, G., MOEHLMAN, P.D., MOLUR, S., MORA, A., NOWELL, K., OATES, J.F., OLECH, W., OLIVER, W.R.L., OPREA, M., PATTERSON, B.D., PERRIN, W.F., POLIDORO, B.A., POLLOCK, C., POWEL, A., PROTAS, Y., RACEY, P., RAGLE, J., RAMANI, P., RATHBUN, G., REEVES, R.R., REILLY, S.B., REYNOLDS, J.E., RONDININI, C., ROSELL-AMBAL, R.G., RULLI, M., RYLANDS, A.B., SAVINI, S., SCHANK, C.J., SECHREST, W., SELF-SULLIVAN, C., SHOEMAKER, A., SILLERO-ZUBIRI, C., DE SILVA, N., SMITH, D.E., SRINIVASULU, C., STEPHENSON, P.J., VAN STRIEN, N., TALUKDAR, B.K., TAYLOR, B.L., TIMMINS, R., TIRIRA, D.G.,

TOGNELLI, M.F., TSYTSULINA, K., VEIGA, L.M., VIÉ, J.C., WILLIAMSON, E.A., WYATT, S.A., XIE, Y., YOUNG, B.E., 2008. The status of the world's land and marine mammals: diversity, threat, and knowledge. *Science.* **322**, 225–230.

SCHLUTER, D., 1996. Causes of Adaptive Radiation. *Am. Nat.* **148**, S40–S64.

SCHMERLER, S.B., CLEMENT, W.L., BEAULIEU, J.M., CHATELET, D.S., SACK, L., DONOGHUE, M.J., EDWARDS, E.J., 2012. Evolution of leaf form correlates with tropical-temperate transitions in *Viburnum* (Adoxaceae). *Proc. R. Soc. B Biol. Sci.* **279**, 3905–3913.

SCHOLTZ, C.H., CHOWN, S.L., 1995. Insects in southern Africa: How many species are there? *S. Afr. J. Sci.* **91**, 124–126.

SCOTT, L., ANDERSON, H., ANDERSON, J., 1997. Vegetation history, in: Richardson, D.M., Pierce, S.M., Cowling, R.M. (Eds.), Vegetation of Southern Africa. Cambridge University Press, Cambridge, pp. 62–90.

SHIN, S., BAYLESS, K.M., WINTERTON, S.L., DIKOW, T., LESSARD, B.D., YEATES, D.K., WIEGMANN, B.M., TRAUTWEIN, M.D., 2018. Taxon sampling to address an ancient rapid radiation: a supermatrix phylogeny of early brachyceran flies (Diptera). *Syst. Entomol.* **43**, 277–289.

SIEMANN, E., TILMAN, D., HAARSTAD, J., RITCHIE, M., 1998. Experimental tests of the dependence of arthropod diversity on plant diversity. *Am. Nat.* **152**, 738–750.

SIESSER, W.G., 1980. Late Miocene origin of the Benguela upswelling system off Northern Namibia. *Science.* **208**, 283–285.

SIMON, C., FRATI, F., BECKENBACH, A., CRESPI, B., LIU, H., FLOOK, P., 1994. Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers. *Ann. Entomol. Soc. Am.*

87, 651–701.

SIMON, M.F., GRETHER, R., DE QUEIROZ, L.P., SKEMAE, C., PENNINGTON, R.T., HUGHES, C.E., 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 20359–20364.

SIMPSON, G.G., 1953. The Major Features of Evolution. Columbia University Press, New York.

SMISSEN, P.J., ROWE, K.C., 2018. Repeated biome transitions in the evolution of Australian rodents. *Mol. Phylogenet. Evol.* **128**, 182–191.

SMITH, M.A., WOODLEY, N.E., JANZEN, D.H., HALLWACHS, W., HEBERT, P.D.N., 2006. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *Proc. Natl. Acad. Sci.* **103**, 3657–3662.

SMITH, S.A., DONOGHUE, M.J., 2008. Rates of molecular evolution are linked to life history in flowering plants. *Science.* **322**, 86–89.

SNYMAN, L.P., NEVES, L., LEMPEREUR, L., BOSMAN, A.C., 2020. Overview of the horseflies (Diptera: Tabanidae) of South Africa: assessment of major collections for spatiotemporal analysis. *Austral Entomol.* **59**, 549–560.

SOLODOVNIKOV, A., SHAW, J.J., 2017. The remarkable Australian rove beetle genus *Myotyphlus*: its cryptic diversity and significance for exploring mutualism among insects and mammals (Coleoptera: Staphylinidae). *Austral Entomol.* **56**, 311–321.

SOLTIS, D.E., SOLTIS, P.S., DOYLE, J.J., 1998. Molecular Systematics of Plants II, Journal of Chemical Information and Modeling. Springer US, Boston, MA.

STOCK, W.D., VERBOOM, G.A., 2012. Phylogenetic ecology of foliar N and P concentrations and N:P ratios across mediterranean-type ecosystems. *Glob. Ecol. Biogeogr.* **21**, 1147–1156.

STUART, B.L., INGER, R.F., VORIS, H.K., 2006. High level of cryptic species diversity revealed by sympatric lineages of Southeast Asian forest frogs. *Biol. Lett.* **2**, 470–474.

SUTTON, T.L., DEGABRIEL, J.L., RIEGLER, M., COOK, J.M., 2017. Local coexistence and genetic isolation of three pollinator species on the same fig tree species. *Heredity.* **118**, 486–490.

SUZÁN, G., MARCÉ, E., GIERMAKOWSKI, J.T., MILLS, J.N., CEBALLOS, G., OSTFELD, R.S., ARMIÉN, B., PASCALE, J.M., YATES, T.L., 2009. Experimental evidence for reduced rodent diversity causing increased hantavirus prevalence. *PLoS One.* **4**.

SWALLOW, J.G., WILKINSON, G.S., MARDEN, J.H., 2000. Aerial performance of stalk-eyed flies that differ in eye span. *J. Comp. Physiol. - B Biochem. Syst. Environ. Physiol.* **170**, 481–487.

TALAVERA, G., KALISZEWSKA, Z.A., HEATH, A., PIERCE, N.E., 2020. Recent diversification of Chrysoritis butterflies in the South African Cape (Lepidoptera: Lycaenidae). *Mol. Phylogenet. Evol.* **148**.

TAPLEY, B., MICHAELS, C.J., GUMBS, R., BÖHM, M., LUEDTKE, J., PEARCE-KELLY, P., ROWLEY, J.J.L., 2018. The disparity between species description and conservation assessment: A case study in taxa with high rates of species discovery. *Biol. Conserv.* **220**, 209–214.

TEDESCO, P.A., BIGORNE, R., BOGAN, A.E., GIAM, X., JÉZÉQUEL, C., HUGUENY, B., 2014. Estimating how many undescribed species have gone extinct. *Conserv. Biol.*

28, 1360–1370.

THERON, G.L., GRENIER, F., ANDERSON, B.C., ELLIS, A.G., JOHNSON, S.D., MIDGLEY, J.M., VAN DER NIET, T., 2020. Key long-proboscid fly pollinator overlooked: morphological and molecular analyses reveal a new *Prosoeca* (Nemestrinidae) species. *Biol. J. Linn. Soc.* 1–13.

THOMPSON, J.D., HIGGINS, D.G., GIBSON, T.J., 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.

TOBIAS, J.A., SEDDON, N., SPOTTISWOODE, C.N., PILGRIM, J.D., FISHPOOL, L.D.C., COLLAR, N.J., 2010. Quantitative criteria for species delimitation. *Ibis.* **152**, 724–746.

TYSON, P.D., PRESTON-WHYTE, R.A., 2000. The weather and climate of southern Africa, 2nd editio. ed. Oxford University Press, Cape Town.

VAN DER NIET, T., PEAKALL, R., JOHNSON, S.D., 2014. Pollinator-driven ecological speciation in plants: New evidence and future perspectives. *Ann. Bot.* **113**, 199–211.

VERBOOM, G.A., ARCHIBALD, J.K., BAKKER, F.T., BELLSTEDT, D.U., CONRAD, F., DREYER, L.L., FOREST, F., GALLEY, C., GOLDBLATT, P., HENNING, J.F., MUMMENHOFF, K., LINDER, H.P., MUASYA, A.M., OBERLANDER, K.C., SAVOLAINEN, V., SNIJMAN, D.A., NIET, T. van der, NOWELL, T.L., 2009. Origin and diversification of the Greater Cape flora: Ancient species repository, hot-bed of recent radiation, or both? *Mol. Phylogenet. Evol.* **51**, 44–53.

VOJE, K.L., 2016. Scaling of morphological characters across trait type, sex, and environment: A meta-analysis of static allometries. *Am. Nat.* **187**, 89–98.

WATERMAN, R.J., BIDARTONDO, M.I., STOFBERG, J., COMBS, J.K., GEBAUER, G., SAVOLAINEN, V., BARRACLOUGH, T.G., PAUW, A., 2011. The Effects of Above- and Belowground Mutualisms on Orchid Speciation and Coexistence. *Am. Nat.* **177**, E54–E68.

WEDMANN, S., 2007. a Nemestrinid Fly (Insecta: Diptera: Nemestrinidae: Cf. *Hirmoneura*) From the Eocene Messel Pit (Germany). *J. Paleontol.* **81**, 1114–1117.

WEDMANN, S., HÖRNSCHEMEYER, T., ENGEL, M.S., ZETTER, R., GRÍMSSON, F., 2021. The last meal of an Eocene pollen-feeding fly. *Curr. Biol.* 1–7.

WHITTALL, J.B., HODGES, S.A., 2007. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature.* **447**, 706–709.

WIEGMANN, B.M., TRAUTWEIN, M.D., WINKLER, I.S., BARR, N.B., KIM, J.-W., LAMBKIN, C., BERTONE, M.A., CASSEL, B.K., BAYLESS, K.M., HEIMBERG, A.M., WHEELER, B.M., PETERSON, K.J., PAPE, T., SINCLAIR, B.J., SKEVINGTON, J.H., BLAGODEROV, V., CARAVAS, J., KUTTY, S.N., SCHMIDT-OTT, U., KAMPMEIER, G.E., THOMSON, F.C., GRIMALDI, D., BECKENBACH, A.T., COURTNEY, G.W., FRIEDRICH, M., MEIER, R., YEATES, D.K., 2011. Episodic radiations in the fly tree of life. *Proc. …* **108**, 5690–5695.

WIENS, J.J., GRAHAM, C.H., 2005. Niche conservatism: Integrating evolution, ecology, and conservation biology. *Annu. Rev. Ecol. Evol. Syst.* **36**, 519–539.

WILKINSON, G.S., REILLO, P.R., 1994. Female choice response to artificial selection on an exaggerated male trait in a stalk-eyed fly. *Proc. R. Soc. B Biol. Sci.* **255**, 1–6.

WINTERTON, S.L., WIEGMANN, B.M., SCHLINGER, E.I., 2007. Phylogeny and Bayesian divergence time estimations of small-headed flies (Diptera: Acroceridae) using multiple molecular markers. *Mol. Phylogenet. Evol.* **43**, 808–832.

YODER, A.D., BURNS, M.M., GÉNIN, F., 2002. Molecular evidence of reproductive isolation in sympatric sibling species of mouse lemurs. *Int. J. Primatol.* **23**, 1335–1343.

YODER, J.B., CLANCEY, E., DES ROCHES, S., EASTMAN, J.M., GENTRY, L., GODSOE, W., HAGEY, T.J., JOCHIMSEN, D., OSWALD, B.P., ROBERTSON, J., SARVER, B.A.J., SCHENK, J.J., SPEAR, S.F., HARMON, L.J., 2010. Ecological opportunity and the origin of adaptive radiations. *J. Evol. Biol.* **23**, 1581–1596.

YORK, G.T., PRESCOTT, H.W., 1952. Nemestrinid Parasites of Grasshoppers. *J. Econ. Entomol.* **45**, 5–10.

YOSHITAKE, H., KATO, T., JINBO, U., ITO, M., 2008. A new *Wagnerinus* (Coleoptera: Curculionidae) from northern Japan: Description including a DNA barcode. *Zootaxa.* **1740**, 15.

YU, Y., BLAIR, C., HE, X., 2020. RASP 4: Ancestral State Reconstruction Tool for Multiple Genes and Characters. *Mol. Biol. Evol.* **37**, 604–606.

ZACHOS, F.E., 2018. Mammals and meaningful taxonomic units: the debate about species concepts and conservation. *Mamm. Rev.* **48**, 153–159.

ZHANG, J., KAPLI, P., PAVLIDIS, P., STAMATAKIS, A., 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics.* **29**, 2869–2876.