Molecular Phylogenetics and Evolution 184 (2023) 107798

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

The phylogenetic structure and coalescent species delimitation of an endemic trapdoor spider genus, *Stasimopus* (Araneae, Mygalomorphae, Stasimopidae) in the Karoo region of South Africa

Shannon Brandt [a,*], Catherine Sole [a,*], Robin Lyle [b]

[a] Department of Zoology and Entomology, University of Pretoria, Hatfield, Pretoria, South Africa
[b] Agricultural Research Council – Plant Health and Protection, Biosystematics, Roodeplaat, Pretoria, South Africa

ARTICLE INFO

Keywords:
Trapdoor-spider
Stasimopus
Species delimitation
Phylogenetics
Karoo

ABSTRACT

The Karoo region of South Africa is a unique and sensitive ecosystem which is facing pressure for development due to economic incentives such as mining, farming and shale gas exploration. The species diversity of many taxa in the area is largely unknown. A phylogenetic analysis of the cork-lid trapdoor spider genus, *Stasimopus* (Stasimopidae) was undertaken in order to gain insight into the relationships between the species that may be present in the area. The species within *Stasimopus* are challenging to identify and define using traditional morphological methods due to a high degree of morphological conservatism within the genus. For this reason, multiple coalescent based species delimitation methods were used to attempt to determine the species present for *Stasimopus* in the region which was tested against the morphological identifications and genetic clades (based on CO1, 16S and EF-1γ). We tested single-locus methods Automatic Barcode Gap Discovery (ABGD), Bayesian implementation of Poisson Tree Processes (bPTP) and General Mixed Yule- Coalescent (GMYC), as well as multi-locus Brownie. The phylogenetic analysis of *Stasimopus* in the Karoo showed that there is a high degree of genetic diversity within the genus. The species delimitation results proved unfruitful for the genus, as they appear to delimit population structure rather than species for most methods. Alternative methods should be investigated to aid in the identification of the species in order truly understand the species diversity of the genus.

1. Introduction

Mygalomorph spiders are notoriously difficult to study for a number of reasons such as outdated literature, the challenge of sampling these cryptic organisms, sexual dimorphism (making linking the sexes morphologically difficult) and their highly conservative morphologies (Wilson et al. 2018). This combination makes species level identifications and descriptions incredibly challenging, which leads to poorly defined conservation statuses across numerous mygalomorph taxa. Resolving this has become increasingly important with the reported population declines in these spiders especially as they tend to be short range endemics exposed to an ever-changing environment.

The spider infraorder Mygalomorphae is a primitive group of spiders which have largely retained their ancestral traits. These plesiomorphic characters include longitudinal fangs, no specialisation of the spinning structures and the retention of two pairs of book lungs (Beavis et al., 2011; Godwin et al., 2018). Mygalomorph spiders live in retreats comprising vertical silk-lined burrows or chambers under rocks or on trees (Dippenaar-Schoeman, 2002; Mason et al., 2018). The females lead sedentary lifestyles, only exiting the burrow to capture prey within reach of the burrow entrance (Engelbrecht and Prendini, 2011; Satler et al., 2013). The males however, are nomadic in nature, wandering to locate females and thus being the drivers of nuclear genome dispersal. Once mature enough the offspring will leave the mothers burrow but only move a few meters before making a new burrow to live in. This behaviour leads to extensive population structuring (Bond et al., 2001; Newton et al., 2020; Satler et al., 2013). Mygalomorph spiders have life history traits of long life spans and sedentary habits, making them ideal candidates for both phylogenetic and phylogeographic studies as well as biodiversity monitoring (Ferretti et al., 2014; Kremen et al., 1993; Leavitt et al., 2015; Mason et al., 2018; Newton et al., 2020).

The focal taxon for this study is the genus *Stasimopus* (Simon, 1892) (African cork-lid trapdoor spider) of the family Stasimopidae (Opatova et al., 2019). *Stasimopus* is endemic to Southern Africa, consisting of 47

* Corresponding authors.
E-mail addresses: shannon.brandt@tuks.co.za (S. Brandt), catherine.sole@up.ac.za (C. Sole).

https://doi.org/10.1016/j.ympev.2023.107798
Received 13 June 2022; Received in revised form 30 March 2023; Accepted 18 April 2023
Available online 23 April 2023

described species (World Spider Catalog, 2022), but the true diversity in the genus could be significantly higher (Engelbrecht and Prendini, 2011). Based on the taxonomic literature, five species of *Stasimopus* are known from within the part of the Karoo accessed in this study and two species from nearby the area (Dippenaar-Schoeman et al., 2010).

Stasimopus and other mygalomorphs were historically described and taxonomically placed on the basis of morphological characters. In *Stasimopus*, features such as the shape of the palpal bulb, length of the pedipalp in relation to the first leg, spination patterns on the legs and eye patterns are used for identification (Engelbrecht and Prendini, 2012; Hendrixson and Bond, 2004). The retention of ancestral traits has led to a group of spiders with highly conservative morphologies, which makes species level identifications challenging (Hamilton et al., 2014; Newton et al., 2020; Opatova et al., 2019; Rix et al., 2018). Genetic methods have been used to aid in the identification of mygalomorphs at species level by creating phylogenies, which indicate divergent groups or separate evolutionary lineages. To define a species one should start by using species delimitation methods and then confirm the putative species by using multiple lines of evidence, such as morphological, genetic and ecological data (De Queiroz, 2007; Derkarabetian and Hedin, 2014). Coalescent-based species delimitation makes use of molecular data to determine species boundaries (De Queiroz, 2007; Fujita et al., 2012). The coalescent approach provides objectivity, as specific probabilities of speciation events are produced which is sustained by the evolutionary and population genetic history of the taxa (Carstens et al., 2013; Fujita et al., 2012; Fujita et al., 2011).

There are two broad types of delimitation approaches based on information availability: discovery approaches and validation approaches. Discovery approaches are used in instances where prior knowledge is limited and populations or taxonomy is not adequately defined and putative groups must be created (Carstens et al., 2013). This is done by the use of molecular methods or morphology (Carstens et al., 2013; Fujita et al., 2012; Satler et al., 2013). Species validation approaches require prior information such as predetermined species tree topologies and individuals already allocated to different putative species. It has been demonstrated that multiple delimitation approaches should be applied to a data set and that results which are congruent across the methodologies are considered reliable (Carstens et al., 2013; Fujita et al., 2012).

Species delimitation techniques have been employed to several mygalomorph groups with varying degrees of success (Hamilton et al., 2014; Hedin, 2015; Leavitt et al., 2015; Opatova and Arnedo, 2014; Satler et al., 2013, 2011). These largely empirical studies found that mygalomorph spiders tend to exhibit strong genetic structuring, leading to contentious species limits, thus species delimitation methods tend to oversplit the data (Hamilton et al., 2014; Leavitt et al., 2015; Opatova and Arnedo, 2014). These studies tend to rely on only a single gene region (usually mitochondrial) or only single locus methods, the delimitations can thus likely be improved by the addition of more gene regions (especially nuclear genes) or other delimiting characters such as morphology (Hedin, 2015; Satler et al., 2013). These studies did however find coalescent species delimitation useful for uncovering species complexes in cases of crypsis, such as Satler et al. (2013) for the *Aliatypus* genus in California. By uncovering cases of crypsis, or simply understanding the species diversity of a taxon, conservation efforts can be more directed and effective.

Concise and effective conservation is vital when dealing with sensitive ecosystems which have slow recovery to disturbance. An area for concern in South Africa is the Karoo region. This is a unique semi-arid area, in the interior of the country characterised by several mountain ranges. The Karoo is seeing large scale land use changes for mining, farming, renewable energy, intensive shale gas exploration and the Square Kilometre Array (SKA) (Sethusa, 2016). In 2016 the Karoo Biogaps project was undertaken to gain insight into the biodiversity standing of the region, which until then had never been fully assessed (Holness et al., 2016; Sethusa, 2016). This provided the opportunity to try to unravel the species diversity of *Stasimopus* in this unique ecosystem. This study is a stepping stone to increasing our understanding of the greater diversity of the entire genus, which given greater funding opportunities, should be addressed in the future. We hypothesise that a combination of species delimitation techniques will be able to identify species clusters which match morphological identifications and monophyletic groupings.

This paper therefore aims to investigate i) the phylogenetic relationship between the various *Stasimopus* specimens within the Karoo region of South Africa and ii) to determine the species boundaries and estimate the number of species present in the region using coalescent species delimitation and confirm their accuracy against both the presence of monophyletic genetic clusters and morphological identification.

2. Methods and materials

2.1. Taxon sampling

Stasimopus specimens were collected from part of the Great Karoo, South Africa. The area is within the demarcated area for potential shale gas fracking. The boundary of the area is set approximately by the following coordinates: −30.88688, 26.29295 and –33.03079, 20.01661 (Fig. 1). The sites were selected by the Karoo BioGaps team to cover the range of environmental conditions present in the region. A total of 79 sites were sampled, on average 50 km apart (Fig. 1). *Stasimopus* specimens were collected at 55 of these sites. Females and juveniles were collected by actively excavating their burrows, while males were found active while crossing isolated roads on rainy evenings. All specimens were euthanised in a solution of cold alcohol and kept in an ice box. Specimens were later preserved in 80 % ethanol in glass polytop vials for long term storage.

Additional recently collected museum specimens (*Stasimopus hewitti*) as well as individuals collected on other sampling trips from the Karoo area and Hluhluwe (KwaZulu-Natal) were included in the *Stasimopus* dataset. All locality information is available in Table S1. All newly collected material was deposited in the National Collection of Arachnida (NCA) of the Agricultural Research Council (ARC), Roodeplaat, South Africa.

2.2. Morphological classification

All adult specimens were examined against existing species descriptions and all type specimens of the genus. All adults were then assigned to morphospecies, based solely on the morphological species concept. Following this, species were assigned to existing species based on current species descriptions and types. All morphospecies which could not be classified into an existing species were classified as ‘undescribed’. The examined morphological characters are in a character matrix available in Appendix A.

2.3. DNA extraction, sequencing and alignment

Genomic DNA was extracted from the removed third right leg of each specimen. DNA extraction was performed using the Macherey-Nagel NucleoSpin® Tissue kit (Düren, Germany) following the manufacturer’s instructions.

Three genes were selected for sequencing to account for the differing mutational rate changes over time as well as different patterns of inheritance. These were ribosomal 16S, mitochondrial cytochrome *c* oxidase subunit 1 (CO1) and nuclear elongation factor 1 gamma (EF-1ɣ). The rationale for selecting these three gene regions was that they allow for the maximisation of phylogenetically informative data at a very fine genetic scale (species level variation). Histone H3 (H3) was also sequenced but was found to not be phylogenetically informative at this level. The phylogenetic topologies of the three genes will show if the gene trees reflect the species tree by congruence (Doyle, 1992;

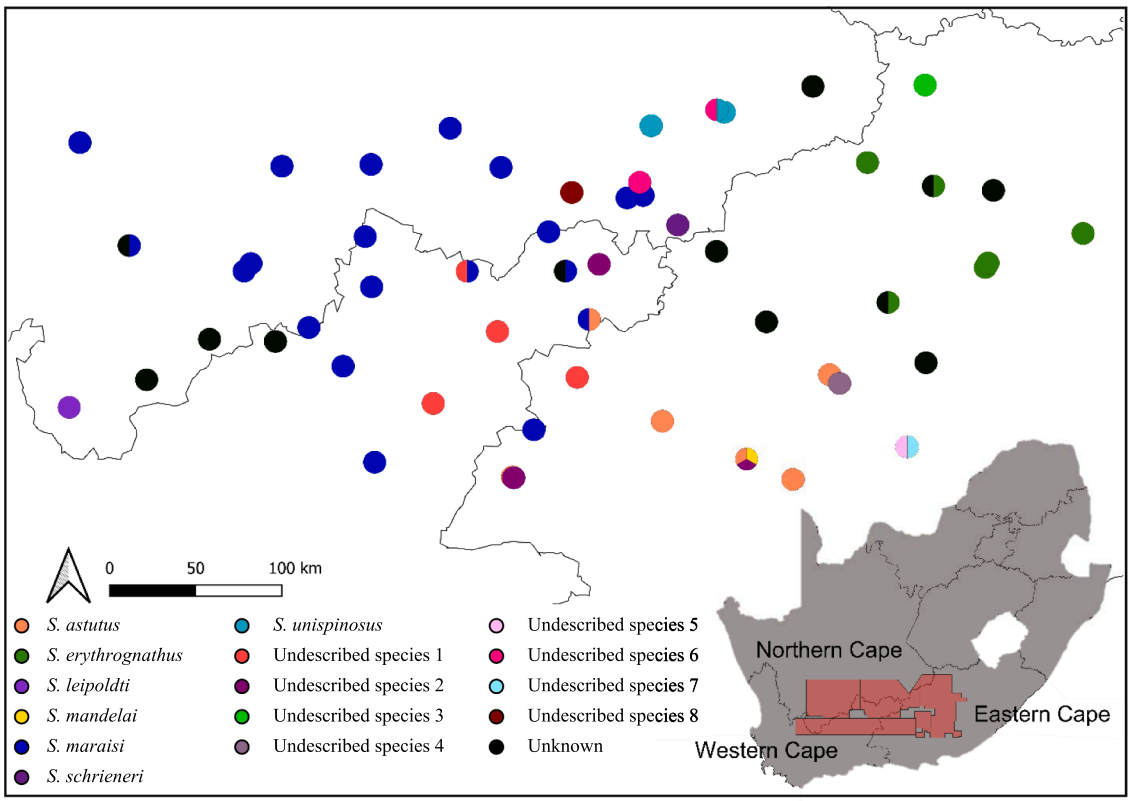

Fig. 1. Map of the study area in the greater Karoo, in the south-western part of South Africa. Markers indicate the 55 sites where *Stasimopus* specimens were found. Specimens are divided into proposed morphospecies. Red polygons indicate areas identified for shale gas exploration. Map created in QGIS version 3.4.8-Madeira (2019), . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
available at: http://qgis.osgeo.org

Maddison, 1997).

Genomic DNA was amplified by polymerase chain reaction (PCR) for the target genes using previously published primer sequences indicated in Table S2 (Ayoub et al., 2007; Cognato and Vogler, 2001; Folmer et al., 1994; Kornilios et al., 2016; Simon et al., 1994). Amplification mixtures were prepared to reach a final volume of 50 μL containing: 2.5 mM $MgCl_2$, 20 pmol of each primer, 10 mM dNTPs, 1 X PCR buffer, one unit of TaqDNApolymerase (Supertherm® DNA polymerase, Separation Scientific SA (Pty) ltd, South Africa) or Emerald Amp®MAX HS PCR Mastermix (TAKARA BIO Inc., Otsu, Shiga, Japan), for problematic samples as well as the EF-1ɣ gene region, in combination with 10–50 ng of extracted genomic DNA template. The PCR cycling parameters performed for the CO1 and 16S gene regions can be viewed in Table S3, and the parameters for EF-1ɣ were as stated in Table S4. Purification of the successful amplifications was done using the Macherey-Nagel NucleoSpin® Gel and PCR Clean-up kit (Düren, Germany) according to manufacturer's specifications. Samples which presented double bands were gel purified following the manufacturer's specifications. The BigDye® Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, Foster City, USA) was used for the cycle sequence reactions in both sequence directions. The cycle sequencing products were precipitated using the standard protocol of sodium acetate and ethanol. All sequences generated were assembled in CLC Bio Main Workbench Version 6.9 (http://www.clcbio.com). All four gene regions were submitted to GenBank and the accession numbers are recorded in Table S1. All the sequences generated from the barcoding CO1 region will be submitted to the Barcode of life database (BOLD).

The CO1, 16S and EF-1ɣ datasets were concatenated using FASconCAT v1.11 (Kück and Meusemann, 2010). The edited sequences for each gene were aligned using MAFFT online (Katoh, 2005; Katoh and Toh, 2008). The 'Auto' strategy for alignment was used in MAFFT followed by visual inspection.

2.4. Phylogenetic analysis

All analyses were applied to the individual gene regions and concatenated molecular datasets for the *Stasimopus* data. The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were both implemented within jModelTest v2.1.7 (Posada, 2008).

Maximum likelihood (RAxML v8.1.20; Stamatakis, 2014) analyses were performed in RAxML implementing a General Time Reversible (GTR) model, with coding genes further partitioned for codon position. A maximum likelihood search was done to find the best tree, followed by bootstrapping for support of 1000 pseudoreplicates (Felsenstein, 1985).

For the Bayesian inference (MrBayes v3.2.5; Ronquist and Huelsenbeck, 2003) analyses, the parameters for each partition were unlinked in order to obtain separate estimates for each gene, the rate-prior was set to variable and flat Dirichlet-priors used. The analysis was run by two simultaneous Monte-Carlo-Markov-Chains, each with three hot and one cold chain (Drummond and Rambaut, 2007). This ran for 10,000,000 iterations, sampling every 500th iteration. This produced 20,000 trees of which the first 25 % (5000) were discarded as burn-in.

Tracer v1.7.1 was used to confirm convergence (Rambaut et al., 2018). The individual runs per dataset were then compiled to ensure a normal distribution.

Outgroups were selected based on the most recent Mygalomorphae revision by Opatova et al. (2019). The sister clade to Stasimopidae in this study is the 'Venom clade', composed of Atracidae and Actinopodidae, represented by *Hadronyche* and *Actinopus* genera respectively. Additional outgroup representation was included with the nemesioids (genera: *Calisoga* (Nemesiidae), *Kiama* (Microstigmatidae), *Pionothele* (Pycnothelidae)) and an paratropidid (*Paratropis*) as this added

additional support.

2.4.1. Ultrametric tree preparation

Two of the species delimitation methods tested require an ultrametric tree as a prior. To comply with this, ultrametric trees were constructed for each gene (some of the species delimitation approaches test only single locus data) using BEAST v1.10.4 (Suchard et al., 2018). For CO1 and 16S, a combined tree was produced as there is mitochondrial linkage (Rubinoff and Holland, 2005). The priors were set in BEAUTi (Suchard et al., 2018) as follows: The nucleotide substitution was set according to the results of jModelTest for each gene and codon partitioning was implemented for CO1 and EF-1ɣ; rate of molecular evolution to a relaxed clock with a lognormal distribution (this allows mutational rates to vary over the tree) (Michonneau, 2017); the MCMC chains were set to a chain length of 100,000,000 with echo states and log parameters to 1000. The last prior altered, was the model describing the branching pattern of the tree. For this prior we aimed to test two models for the GMYC method: the Yule model (all branching is covered by a constant speciation rate) and the Coalescent model with a constant population size (usually used for datasets which have individuals from the same species) (Michonneau, 2017). We decided to test both of these priors as it would allow for numerous (Yule model) or a few species (Coalescent model) (Michonneau, 2017). All other priors were left on the default settings in BEAUTi. Both files (Yule and Coalescent model priors) were then executed in BEAST. This was repeated four times per dataset and model to ensure convergence.

Tracer v1.7.1 was again used to confirm convergence (Rambaut et al., 2018). Log Combiner v1.10.4 (Suchard et al., 2018) was used to combine the tree files into one file per model and per gene and was set with a burn-in of 25,000,000.

2.5. Species delimitation

A variety of different species delimitation approaches were tested to determine i) if there is congruence among methods and ii) to test if these approaches can accurately delimit *Stasimopus* species given the strong population structuring of mygalomorph spiders. All the species delimitation results were tested against the morphological identifications of the specimens and monophyletic clusters, for this reason only identified adult specimens or juveniles whose identity could be inferred genetically were included.

2.5.1. Automated Barcode Gap Discovery (ABDG)

Automated Barcode Gap Discovery (Puillandre et al., 2012) uses the aligned gene sequences to determine species delimitations based upon the degree difference between individuals. The programme identifies the different thresholds for determining species boundaries specific to a particular taxon. This is done instead of the traditional method of having a set cut off divergence value *a priori* to determine species delimitations (Puillandre et al., 2012). The method uses the first barcode gap estimated at a distance limit deemed large enough to delimit intraspecific variation (using the population mutation rate estimated from the dataset) (Puillandre et al., 2012). This is given as the incursive value for the various P – distances partitions. The recursive value adds an addition step to this and takes into account the variability in mutation rate across the taxa in the dataset as well as the overlap of the intraspecific and interspecific variations (Puillandre et al., 2012). The recursive partition is thus considered more reliable (Puillandre et al., 2012).

The online tool was used (available at: http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html) and the analysis parameters set to: pmin = 0.03; pmax = 0.6; x = 0.5; number of steps = 30, number of bins (Nb) = 50 and Kimura distance (TS/TV) left at the default setting (TS/TV = 2.0). The minimum and maximum p-value was determined iteratively by testing different combinations until more than one putative species was given as an output. For this reason, the range of intraspecific divergence values selected was between 0.0300 and 0.0618 as the species threshold. The species delimitations were then assessed congruency against the morphological identifications, genetic clustering and other delimitation methods.

2.5.2. General Mixed Yule coalescent (GMYC)

The GMYC model of species delimitation uses an algorithm which attempts to determine where a swap in branching pattern of a tree is due to a speciation event (which would place one lineage in a species) or due to coalescent events within a species (places multiple lineages in a species) (Michonneau, 2017; Pons et al., 2006).

The GMYC is applied to single locus data, for this reason EF-1ɣ and the combined CO1 / 16S were analysed separately. The ultrametric trees were then imported into R v3.4.3 (R Core Team, 2017) and analysed using the packages ape, paran and splits with the GMYC method (Dinno, 2014; Ezard et al., 2014; Paradis and Schliep, 2019). Both the Yule speciation and Coalescent speciation models were tested.

2.5.3. Bayesian Poisson tree processes (bPTP)

The Bayesian implementation of the Poisson Tree Processes (bPTP) model makes use of jumps in the branching patterns to determine differences between and within species. The model then links the number of nucleotide substitutions in a gene tree (the branch length) to model speciation events – assuming that there is a higher number of substitutions per site between species than within a species (Zhang et al., 2013).

To implement the PTP model with Bayesian support we used the bPTP server (Zhang et al., 2013). We performed the analyses EF-1ɣ and the combined CO1 / 16S separately. The RAxML tree with outgroup was uploaded and the rooted option was selected. All datasets were run for 500,000 MCMC generations. The burn-in was set to 0.25. All other parameters were left on default.

2.5.4. Brownie species delimitation

The heuristic method for species delimitation, produced by O'Meara (2009), makes use of the concept of gene tree congruence in a clade reflecting a cohesive species and incongruence of the gene trees for the clade reflecting variation within a species (O'Meara, 2010). Brownie takes gene trees as input files, performs a heuristic search comparing the topologies of the gene trees and locates areas where there is congruence in the gene trees (within a species) and areas of extensive incongruence (between species). This then sets species limits and estimates the species tree (O'Meara, 2010). The programme represents a species by a polytomy of the related samples. The programme does not inherently produce an estimation of uncertainty, but this can be estimated by bootstrapping (O'Meara, 2010).

Brownie v2.1.2 was run on a Mac OS (O'Meara, 2010). The ultrametric gene trees were imported to R v3.4.3 (R Core Team, 2017) and edited using the 'ape' package (Paradis and Schliep, 2019). This was done to remove tips for taxa which were not present across all three gene regions (14 taxa removed from CO1, 21 from 16S and 4 from EF-1ɣ). This was done as Brownie requires complete coverage of each gene region. The edited individual ultrametric gene trees were used as the input data in NEXUS format in a combined file. The heuristic settings were all left on default. These analyses were repeated ten times to ensure consistent results.

The best trees selected by the program from each run were then used to create a consensus tree, from which another overall consensus tree was constructed using PAUP v4.0a (Swofford, 2002).

3. Results

3.1. Taxon sampling

A total of 125 *Stasimopus* individuals were collected. An additional 4 *Stasimopus* individuals were genetically analysed from outside the study area. These additional samples were included to aid in the identification

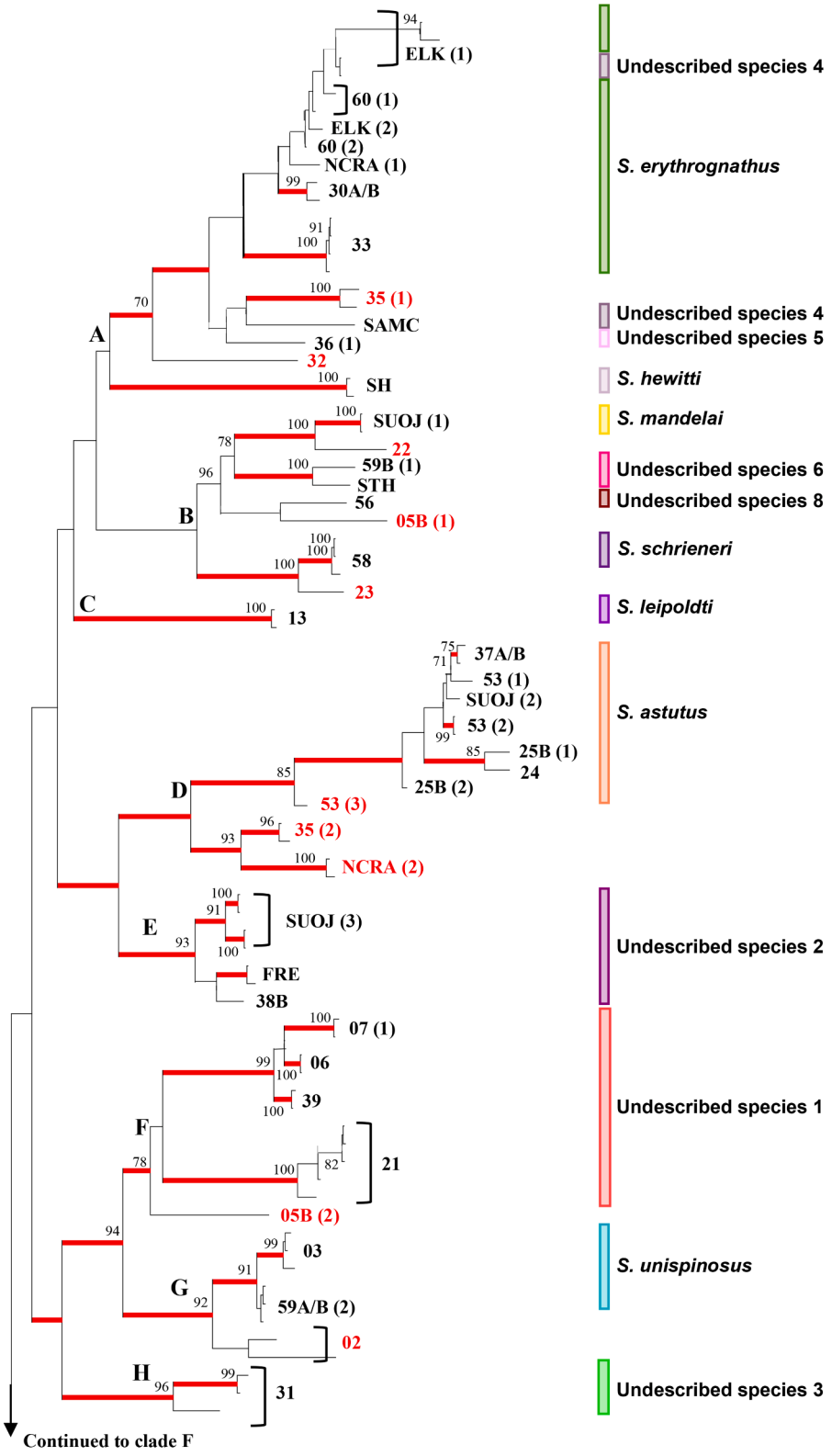

Fig. 2. (Part A). Majority-rule (MrBayes) consensus tree for the concatenated dataset of CO1, 16S and EF-1ɣ for the *Stasimopus* dataset from the Karoo. Posterior probabilities above 0.95 are shown by the red bars. Bootstrap support values above 70 are given on the corresponding node. The morphologically identified species are indicated along the right. Specimens in red indicate juveniles that could not be morphologically identified or attributed to species clades. Broader clades are indicated on the tree. Fig. 2 (Part B). Majority-rule (MrBayes) consensus tree for the concatenated dataset of CO1, 16S and EF-1ɣ for the *Stasimopus* dataset from the Karoo. Posterior probabilities above 0.95 are shown by the red bars. Bootstrap support values above 70 are given on the corresponding node. The morphologically identified species are indicated along the right. Specimens in red indicate juveniles that could not be morphologically identified or attributed to species clades. Broader clades are indicated on the tree. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

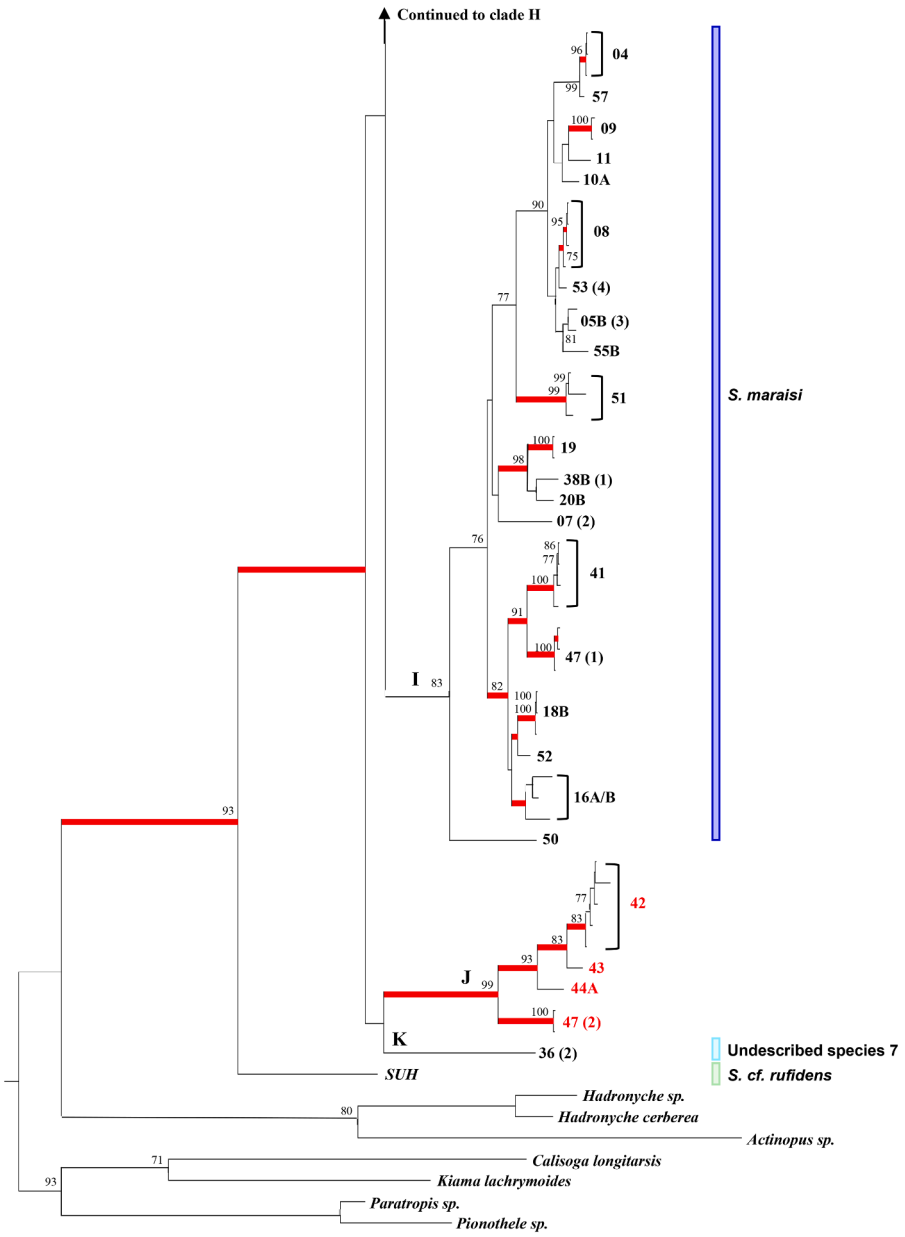

Fig. 2. (*continued*).

of the various lineages produced. At the majority of collecting sites only a few individuals were found during sampling. However, at a few sites, *Stasimopus* were especially abundant and easily found (due to high density or clear burrow entrances). The majority of specimens found were juveniles (61), followed by mature females (43) with few adult males collected (21). Seventeen morphospecies were morphologically assigned, nine of which could be assigned to existing *Stasimopus* species (*S. astutus*, *S. erythrognathus*, *S. hewitti*, *S. leipoldti*, *S. mandelai*, *S. maraisi*, *S. schrieneri*, *S. rufidens*, and *S. unispinosus*). While the other eight species did not match any existing species descriptions or type specimens and were denoted as 'undescribed'.

3.2. Phylogenetic analysis

The number of individuals and base pairs per gene region are summarised in Table S5. The table also indicates the nucleotide substitution model selected by jModelTest for the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

3.3. Stasimopus phylogenies and identities

Stasimopus is a well-supported monophyletic clade from the data (BS: 93; PP:0.97). The individual sample of *Stasimopus* cf. *rufidens* from KwaZulu-Natal (denoted as SUH) is sister to the rest of the *Stasimopus* samples (Fig. 2). The remainder of the phylogram is then divided into ten distinct genetic clades (Fig. 2, A-K).

Clade K is composed of one individual, this is however with weak support. This individual being in its own evolutionary lineage is supported by the results of the species delimitation analyses and morphology. This individual could not be matched to any described species ('Undescribed species 7′). Clade J could not be identified as it is comprised of only juveniles and is likely a separate species. Clade I is a large clade, morphologically identified to be *Stasimopus maraisi*. Individuals from site 47 occur in both clades I and J, indicating synoptic species occurring at this locality. Clade H (BS: 96; PP: $>$ 0.95) consists of individuals from only locality 31. This lineage is supported by the species delimitation and does not match any previously described species morphologically and was donated as 'Undescribed species 3′. Clade G

(BS: 92; PP: 0.99) has been morphologically identified as *S. unispinosus*, but the 02 samples could not be confirmed due to only being juveniles. Clade F (BS: 78; PP: 0.95) did not match any described species ('Undescribed species 1′). The terminal nodes of the clade have much higher support values than the internal nodes. Individuals from site 7 are placed in clade F and I, indicating another case of syntopic species. Clade E (BS: 93; PP: 0.97) did not match any described species ('Undescribed species 2′). Clade D has many samples which could not be identified (NCRA (2), 35 (2) and 53 (3)) as they are juveniles, whereas the rest were identified as *S. astutus*. Individuals from Site 38B and 35 occur in both clades D and I. Clade C is comprised solely of individuals from site 13, these samples have been identified as *S. leipoldti*. The most derived clades are A, and B. Clade B contains multiple morphologically distinct species, each represented by only a few individuals. These species are *S. mandelai*, *S. schrieneri*, 'Undescribed species 6′. and 'Undescribed species 8′. Individuals from site 05B are placed in clades C, F and H, individuals from site 59B are in both clades B and G, and individuals from SOUJ are in clades B, D and E. Clade A also comprises of multiple species. The two SH individuals are *Stasimopus hewitti.* One individual from locality 36 could also not be matched to any described species ('Undescribed species 5′). The most derived portion is an apparent species complex of *S. erythrognathus* and another species which did not match any existing species descriptions ('Undescribed species 4′). Individuals from site NCRA occur in clades A and E.

3.4. Species delimitation

3.4.1. Automated Barcode Gap Discovery species delimitation

ABDG produced various species delimitations based on the different prior maximal distance (P) values used in the input. The larger the P value given, the more split the data became (more species were delimited). When the P value was smaller than 0.0300, more than 60 species were delimited. When the P value was larger than 0.0623 only one species was delimited (Fig. 3).

The initial distance value is more conservative in the number of species delimited than the recursive. The smallest distance value (P = 0.0623) classifies all the samples as one species (Grey, Fig. 3). The incursive results for P 0.033 and 0.037 were identical, thus only 0.033 (I) is shown. The results of the 0.411 and 0.455 distance values were also identical and only 0.0455 is shown in Fig. 3. The incursive results for P 0.0455 and 0.0561 were identical, thus only 0.0561 (I) is shown. This was again true for the distance values of 0.505 and 0.561, and only 0.0561 is shown.

Certain samples are consistently recovered as separate species. *Stasimopus hewitti* (SH02; thistle), *S. leipoldti* (13STAS02; blue violet), *S. unispinosus* (03STAS01, 03STAS02; cyan) and 'Undescribed species 7′. (36STAS01; cadet blue), are all separate species on their own in all distance values except for P = 0.0623 (Fig. 3). A few species are recovered most of the time (except for the distance values of 0.0561 (I) and 0.0623). These include 'Undescribed species 4′. (SAMC7293), 'Undescribed species 5′, 'Undescribed species 8′ and *S. schrieneri* (Fig. 3). *Stasimopus astutus,* 'Undescribed species 1′ and *S. maraisi* were only recovered at the lower distance values, whereas for 'Undescribed species 6′ this is true of the middle-distance values (Fig. 3). *Stasimopus erythrognathus,* 'Undescribed species 1′ and 'Undescribed species 3′ were never fully recovered as they were always overspilt and in the case of *S. erythrognathus* also had additional species included according to genetic clustering and morphology.

The species complex described in the phylogeny (Fig. 2) is apparent here again between *S. erythrognathus* and 'Undescribed species 4′, whereas ELKSTAS06 (ELK (1) – Fig. 2) is never recovered as a separate species (Fig. 3).

3.4.2. General Mixed Yule coalescent (GMYC)

The General Mixed Yule Coalescent (GMYC) gave fairly similar results between the Yule and coalescent model of speciation for each gene region. The CO1_16S gene region presented with 2 putative species (coalescent Confidence interval (CI): 1–10–; yule CI: 1–4) and 2 distinct genetic clusters (coalescent CI: 1–7, yule CI: 1–4) which were identical for the Yule and Coalescent trees. The LR test was not significant for both. The results for the EF-1ɣ region were also not significant under the LR test and presented with more putative species. The coalescent model predicted 8 putative species (CI: 1–17) whereas the yule model predicted 18 putative species (CI: 1–24). The confidence ranges are exceptionally wide for this gene region. The putative species are given for each gene region in Figs. 4 and 5.

3.4.3. Bayesian Poisson tree Process (bPTP)

The Bayesian Poisson Tree Process (bPTP) method of species delimitation results vary between the different gene regions. The two mitochondrial genes (CO1_16S) have a lower number of putative species (36 or 53) than the nuclear gene (29 or 75) (Table 1). There is a wide margin of error as the estimated number of species is a large range, approximately 30 species between the minimum and maximum for each gene region. The acceptance rate is much higher for EF-1ɣ (0.889) than CO1_16S (0.417). The putative species are given for each gene region in Figs. 4 and 5. The number of species for each confidence interval is given in Table S6 as well as putative species acceptance rates in tables S7 – S8.

3.4.4. Summary of single locus coalescent methods

The species delimitation results for linked CO1 and 16S are not congruent between the GMYC and bPTP methodologies (Fig. 4). The bPTP method severely overspilt the dataset, whereas GMYC lumps the data when compared to the morphological identities of the specimens as well as the monophyletic clades. There are only individuals constituting a 'species' for the bPTP method while matching the morphology and genetic clusters: 'Undescribed species 4′ (SAMC7293), 'Undescribed species 5′, *S. hewitti, S. schrieneri, S. leipoldti, S. unispinosus* 'Undescribed species 8′ and 'Undescribed species 7′ (Fig. 4).

EF-1ɣ species delimitation has a different result to the CO1_16S. The GMYC method is more conservative than the bPTP method with 8 or 18 putative species as opposed to 29 or 75 putative species (Fig. 5). No species is congruent across methodologies. The phylogeny produced for the EF-1ɣ is however not in agreement with the morphological identifications.

Across the two separate gene regions' species delimitation results, no species are congruent and neither method could resolve the *S. erythrognathus* and 'Undescribed species 4′ complex.

3.4.5. Brownie species delimitation

The results of the heuristic search run in brownie are inconclusive. The analyses produced 633 tree topologies from the ten runs. However, there is no true congruence between these trees. Individuals constantly swap between the different putative species – having no true placement. A 50 % majority-rule consensus tree is shown as Figure S1. The tree has a hard polytomy, meaning that no real putative species could be delimited. When testing for accuracy within Brownie, either 51 or 53 species were constantly delimited (individuals in the species are not given).

4. Discussion

4.1. Stasimopus diversity

This study is the first detailed phylogenetic study on the group including the use of species delimitation methods. The study generated extensive sequence data for a severly understudied organism in a potentially threatened and sensitive ecosystem. These resources can be used in future studies to better understand the genus.The *Stasimopus* samples assessed in this study form a well-supported monophyletic clade. The phylogeny consists of 17 morphospecies and juvenile specimens which could not be identified, based on morphology. Fifteen species were only recovered from a single or few localities which were

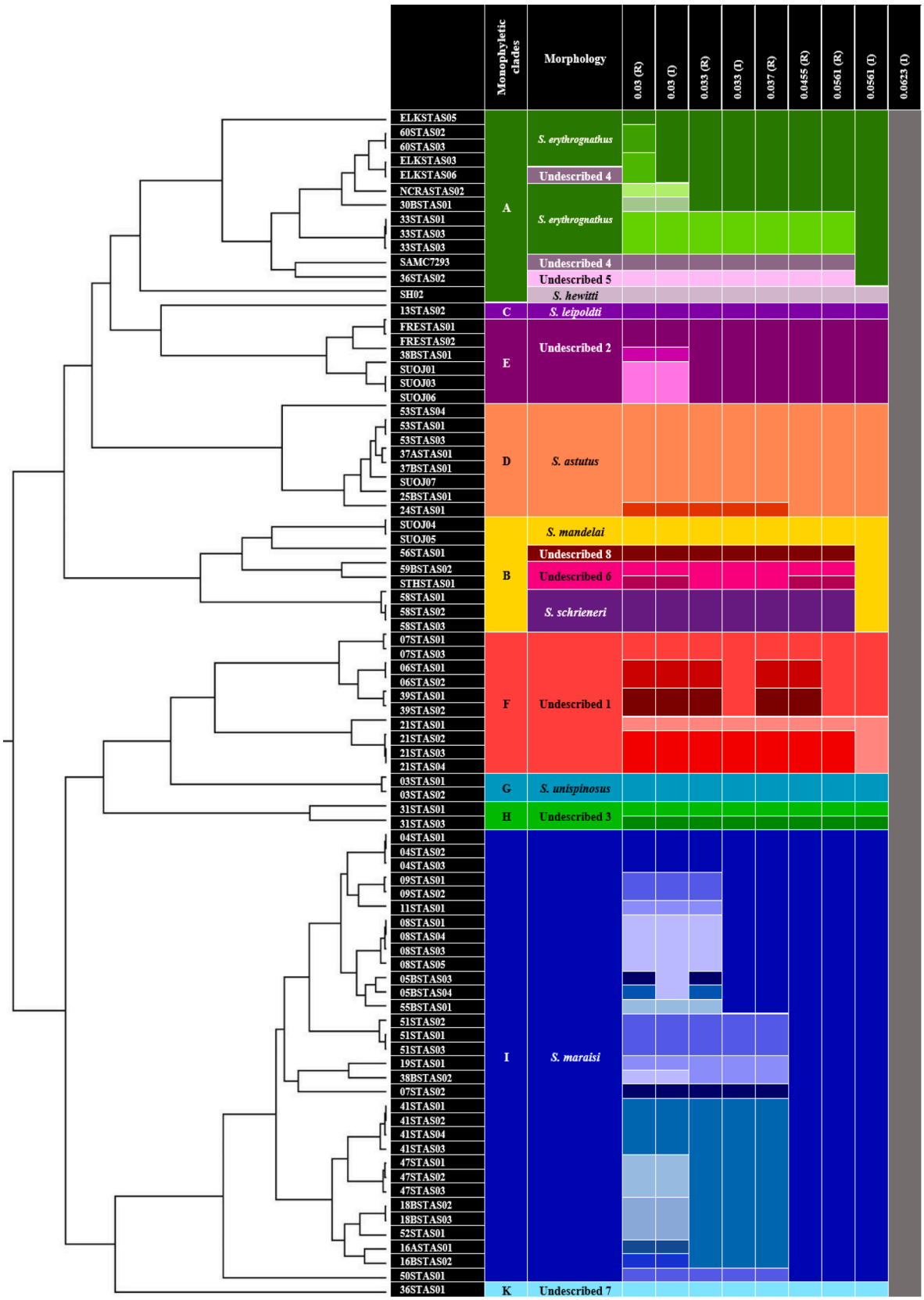

Fig. 3. The species delimitations of the Automated Barcode Gap Discovery tool distance values on the cladogram of the CO1 data. The various partitions for different values of the prior maximal distance (P) are given in the first row. The recursive and incursive partitions are denoted by (R) and (I) respectively. The samples are split into their proposed species and monophyletic clades by the different coloured blocks in a pantone corresponding the *Stasimopus* phylogram (Fig. 2).

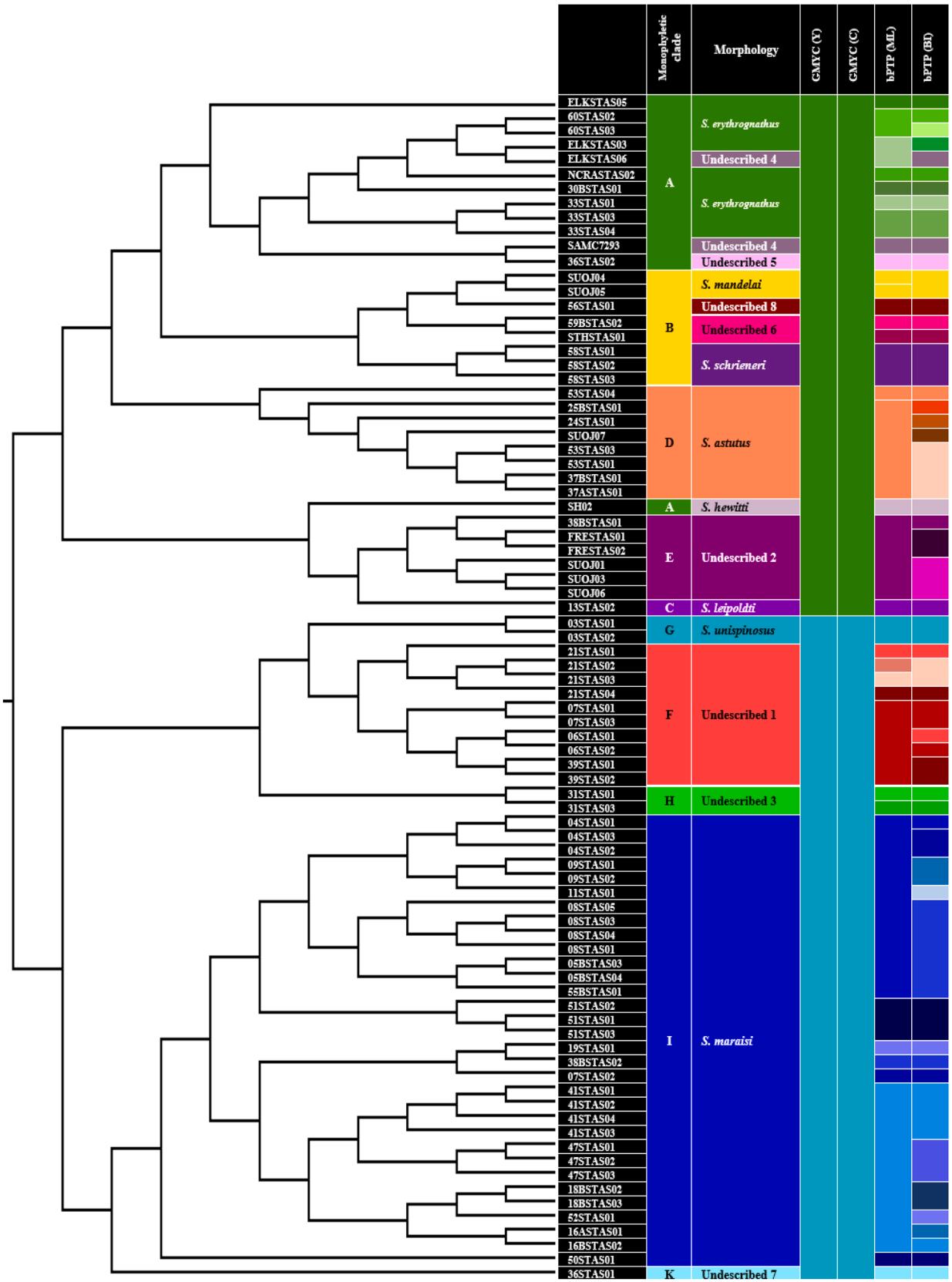

Fig. 4. Combined results for cytochrome oxidase 1 (CO1) and ribosomal 16S bPTP and GMYC species delimitation. The results are overlaid on the CO1 and 16S cladogram. The samples are split into their proposed species and monophyletic by the different coloured blocks in a pantone corresponding the *Stasimopus* phylogram (Fig. 2). The Yule and coalescent speciation models are indicated by (Y) and (C) respectively.

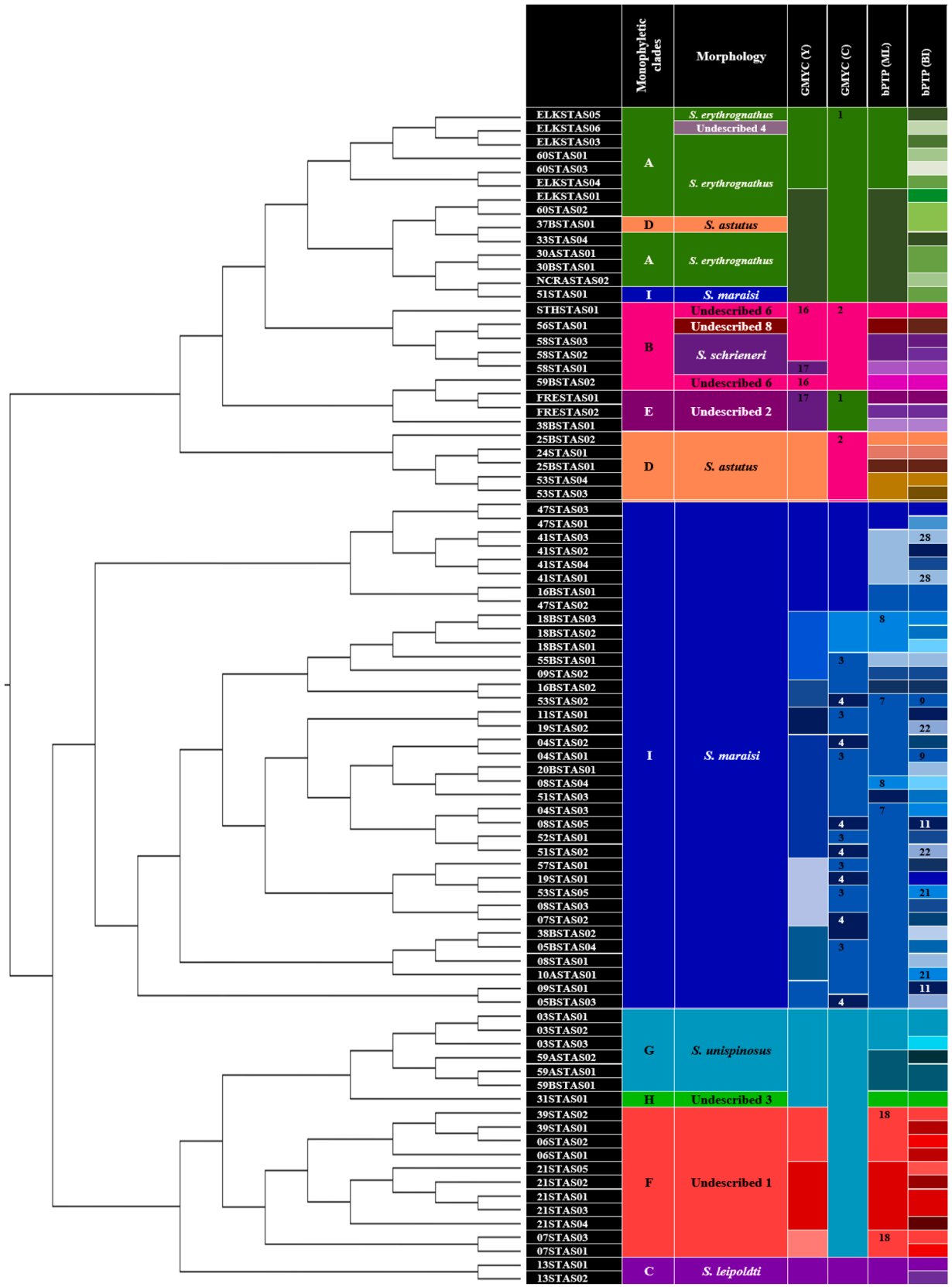

Fig. 5. The elongation factor 1 gamma (EF-1ɣ) species delimitation results for the bPTP and GMYC. The results are overlaid on the EF-1ɣ cladogram. The samples are split into their proposed species and monophyletic clades by the different coloured blocks in a pantone corresponding the *Stasimopus* phylogram (Fig. 2). The Yule and coalescent speciation models are indicated by (Y) and (C) respectively.

Table 1
The number of species delimited for each gene region (CO1_16S and EF-1ɣ) with the Bayesian Poisson Tree Process (bPTP) species delimitation methodology.

	CO1_16S	EF-1ɣ
Acceptance rate	0.417	0.889
Estimated number of species	26–68	29–76
Mean number of species	49.34	55.86
Bayesian Inference species estimation	53	75
Maximum likelihood species estimation	36	29

all in close proximity, indicating that these species are 'short-range endemics'. Short-range endemics are species which have a naturally narrow distribution (<10,000 km^2) (Harvey, 2002; Harvey et al., 2011; Mason et al., 2018). Short-range endemics often have life history traits such as poor dispersal ability, reliance on rare or fragmented habitats and low fecundity (Harvey et al., 2011). *Stasimopus*, like most of the Mygalomorphae, are known to exhibit sedentarily habits and have short dispersal distances (Leavitt et al., 2015) These factors would make a species which is a short-range endemic especially sensitive to climate change, degradation of the habitat or habitat loss (Harvey et al., 2011). This is vital information in light of the slew of land use change proposed for the Karoo region. The morphologically identified species *S. maraisi* and *S. erythrognathus* have wide distributions across the Karoo, and are not short-range endemics. Several species sampled appear to experience sympatry at multiple locations. This indicates that the species may have an unknown ecological niche separation or are syntopic in nature (Wilson et al. 2018).

The species *S. erythrognathus* distribution is not completely known but appears very wide (sampled in the eastern portion of the Karoo and the type locailty being Worcester which are 600 km apart) and likely extends to other unsampled parts of South Africa. There were no specimens found between these two localities during our sampling. This indicates that further sampling needs to be performed in order to determine if this whole range is still occupied or if these may be relict populations/ remnants or perhaps a species complex.

Stasimopus erythrognathus and 'Undescribed species 4′. appear to form a genetic species complex as one specimen is 'Undescribed species 4′ and falls within the *S. erythrognathus* clade. One specimen is placed sister to the *S. erythrognathus* clade as is also 'Undescribed species 4′. All of these specimens were collected in the eastern Karoo and within fairly close geographic proximity. These two species are morphologically distinct and more gene regions are required to resovle the genetic complex found here.

The placement of *Stasimopus* cf. *rufidens*, from Hluhluwe (KwaZulu-Natal) indicates that there is a great amount of genetic distance between the *Stasimopus* of the coastal area and that of the interior Karoo.

4.2. Uncertainty in the phylogenies

Low support values on internal branches couldbe due to factors such as homoplasy, low genetic signal or incomplete lineage sorting (Buckley et al., 2006).

The incongruence of the ML and BI Mygalomorphae and *Stasimopus* trees, as well as low support values may be the impact of missing data or an insufficient number of loci to resolve these family-level relationships. Methods such as concatenation assume that the different genes used have moderately similar evolutionary histories, thus not considering that different genes evolve at different rates (mitochondrial genes evolve more rapidly than nuclear genes due to the pattern of inheritance) (Xi et al., 2016). Thus, when comparing a species for which only mitochondrial information is available to one with only nuclear information in the same phylogeny, would put the individual with the mitochondrial information on a longer branch. This may be the case with clade D of the *Stasimopus* phylogeny. This would make it appear as though there are different rates of evolution between the species which is untrue. Concatenation methods do however fair well with missing data if the data is evenly distributed among the taxa and gene regions (Xi et al., 2016). Our data sets had fairly evenly distributed missing data.

It is also noted that the phylogenetic inference and species delimitation may be impacted by the scattered phylogenetic sampling. The Karoo region was throughly sampled, but this region has no natural boundaries limiting the distribution of taxa, and thus their full ranges were not captured by these data. This may affect the accuracy of the species delimition models (Papadopoulou et al. 2009). These methods can also be affected by using 'singleton' representatives of species, this is however largely unavoidable when working with unforthcoming taxa (Ahrens et al. 2016). Future studies should expand on this sampling effort to better encaptulate the distribution of the genus as well as attempt to include multiple representatives of each species.

4.3. Coalescent species delimitation

Stasimopus as a genus is challenging to identify using morphology alone at species level. For this reason, use of coalescent based species delimitation was investigated as a tool to determine how many species can be identified from a genetic dataset and validated using both morphology and genetic clustering. It is also important to note that the study used not only mitochondrial data, but a nuclear region too.The Automated Barcode Gap Discovery (ABDG) online tool delimited the most species in the largest partition (>0.03) which severely oversplit the CO1 dataset, placing each locality into its own species. This delimitation of 46 species is highly unlikely. This most likely reflects the genetic structure between sites, as a result of limited dispersal. A species delimitation study on the Nemesiidae returned a similar result where each locality was returned as a putative species, the over estimation is most likely due to genetic substructuring which occurs in mygalomorph spiders at small geographic scales (Leavitt et al., 2015). Another example of this was seen in Idiopidae studied in the Canary islands (Opatova and Arnedo, 2014). This appears to be the case with the 0.03 and 0.033 distance values. The largest distance value (0.0623) lumped all individuals into one species. The result of one species is inaccurate as from the males and females collected, 17 separate morphospecies were morphologically identified. This lumping is most likely because the barcode gap is overshot: (P) prior maximum divergence of the intraspecific diversity is too large and overshadows the interspecific diversity (Puillandre et al., 2012). The distance value with the closest results is 0.0561, where the initial distance value which suggests 13 species and recursive suggests 21. The recursive 0.0561 partition delimits most species correctly except the *S. erythrognathus* / 'Undescribed species 4′ complex, and oversplits 'Undescribed species 1′ and 'Undescribed species 6′. This partition being the most reliable is also supported by the fact that the recursive partition is usually considered more reliable than the incursive due to taking more factors in the dataset into account (Puillandre et al., 2012). This methodology appears to have delivered the most plausible results.

The Bayesian Poisson Tree Process (bPTP) results were the least conservative for the EF-1ɣ gene region, estimating 29 or 75 putative species. The number of putative species predicted for the CO1_16S region also appears to be unrealistically high too. The number of species predicted by this analysis would be lower (and possibly more accurate) if nodes with low support values (<0.80) are collapsed (Zhang et al., 2013). There is also incongruence between the gene tree produced for the EF-1ɣ dataset and the morphology. This is clear from the one *S. atutus* and *S. marasi* samples that are embedded in clade A. This pattern is however not seen in the mitochondrial gene tree, nor the combined tree. More nuclear genes need to be added to investigate this fully. This may in part be responsible for the poor performance of the species delimitation methods.

The General Mixed Yule Coalescent (GMYC) gave similar results between the coalescent and yule model of speciation for the CO1_16S dataset. The Yule model assumes a constant speciation rate across

branches and is often used in cases where the individuals are from numerous species, whereas the Coalescent model (constant growth rate) is often used in cases with a few species (Michonneau, 2017). It appears that the speciation model of the ultrametric tree only had a significant effect on the EF-1ɣ data set where the number of delimited species changed from 18 (Yule) to 8 (Coalescent), but this result was not significant under the LR test. The species were however incorrectly delimited according to morphology and monophyletic clusters. The discrepancy between the mitochondrial genes (CO1 and 16S) and the nuclear gene in the number of putative species produced may be attributed to the fact that these two parts of the genome have very different rates of evolution. For the class Arachnida, the rate of mitochondrial evolution is on average 3.1 times faster than nuclear evolution (Allio et al., 2017). The GMYC appears to have lumped the CO1_16S dataset into only two species, which is in reality not feasible based on the monophyletic clades and morphology. GMYC is commonly cited as oversplitting data for the mygalomorph group due to the strong population structuring observed in spiders with sedentary lifestyles, whereas the opposite was observed (Hamilton et al., 2014; Hedin, 2015; Opatova and Arnedo, 2014; Satler et al., 2013).

As a whole, the single-locus methods of species delimitation performed poorly at delimiting species for the *Stasimopus* genus when validated against the morphological identifications as well as genetic clustering. There was no full congruence across methods for any of the gene regions. The Automated Barcode Gap Discovery (ABDG) shows the most promising results for the CO1 region, where both bPTP and GMYC oversplit the data. Overall, CO1 and DNA barcoding may need to be interpreted with caution when used in isolation, as it may not capture all the variation present and may be lacking some resolving power (Galtier et al., 2009). Galtier et al. (2009) found that many of the underlying principles of mitochondrial DNA which supports barcoding may be false. Some of these characteristics include mitochondrial DNA undergoes recombination, positive selection as well as an evolutionary rate which changes over time (Galtier et al., 2009). This emphasises that species delimitation cannot be performed using only a single gene region, as no single gene region is impervious to these factors. This means that multiple gene regions should be considered when delimiting species or that the species delimitation tools must be developed so that the variation between different gene regions is taken into account (Hedin et al., 2015).

Brownie (multilocus species delimitation) provided inconclusive results as no congruence was found between the trees produced by the analysis. A study conducted on frogs in South America by Correa et al. (2017) had the same result from analyses run by Brownie. They recovered a consensus tree with a hard polytomy – resulting in no reliable species being delimited (Correa et al., 2017).

This study has disproven the hypothesis as none of the species delimition techniques were able to acurately produce species clusters when compared to each other, morphology or monophyletic clusters. From the results of the species delimitation analyses one of three scenarios are likely taking place. Firstly, there is the possibility that there are approximately 40 species of *Stasimopus* in this region of the Karoo. This result was recovered in part by all the methods, but is highly unlikely. Secondly, the *Stasimopus* of the Karoo are undergoing speciation. This may be a possibility as the life history traits of the genus makes them prime candidates to undergo allopatric speciation. The largely sedentary lifestyles of females and limited dispersal of spiderlings tends to lead to extensive population structuring which are referred to as 'microallopatric populations' within a species (Bond et al., 2001; Satler et al., 2013). These populations may then remain isolated for extensive periods of time leading to speciation events. This scenario is supported by the short terminal branches of the phylogeny, which is a common indicator of speciation events. The last scenario is simply that due to the extensive population structure (microallopatric populations), the species delimitation methods used are not suitable for resolving the relationships and are prone to oversplitting the dataset (except for GMYC) (Satler et al., 2013). This scenario is supported as coalescent species delimitation techniques appear to delimit genetic structure, not species (Sukumaran and Knowles, 2017).

5. Conclusion

The phylogenetic diversity of *Stasimopus* within the Karoo region of South Africa is higher than previously believed with eight undescribed species identified. *Stasimopus* and other trapdoor spiders have not been extensively researched, but are a significant part of the Karoo fauna. For this reason, they should be considered for inclusion in environmental impact assessments, evaluated for placement on the IUCN Red List and other conservation initiatives. It is suggested that species delimitation is not purely based on barcoding genes but a multi-locus approach should always be implemented. From the results of the species delimitation, we find that these methods are not suitable to delimit species for the genus as the population has extensive substructure at fine geographical scales, leading to 'microallopatric populations'. More research is needed to find gene regions which have greater resolution for this group of spiders. The way forward is to determine the evolutionary history and environmental pressures which may be causing the divergence in the populations by implementing a phylogeographic analysis of the data.

Funding

The National Research Foundation (NRF) are acknowledged for funding to the second author. The first author would also like to acknowledge the National Research Foundation for funding (Grant no: 120,190 and 139125). This work is based on the research supported in part by the Karoo BioGaps Project (NRF Grant No 98864), which was awarded through the Foundational Biodiversity Information Programme (FBIP), a joint initiative of the Department of Science and Innovation (DSI), the National Research Foundation (NRF) and the South African National Biodiversity Institute (SANBI).

CRediT authorship contribution statement

Shannon Brandt: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Catherine Sole:** Investigation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Robin Lyle:** Validation, Investigation, Resources, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to acknowledge Carmen Jacobs and Ishtiag Abdalla Mohammed for assistance with molecular work and phylogenetic analyses. Tshepiso Majelantle and Elmé Brand for their assistance on fieldwork. Equipment for field work was supplied by the Agricultural Research Council. We would like to thank the following institutions for loaning the specimens required for specimen identification: Albany Museum, Grahamstown, South Africa; The National Museum, Bloemfontein, South Africa; Ditsong Museum of Natural History, Pretoria, South Africa; Iziko Museum of Cape Town, South Africa; Museum für Naturkunde, Berlin, Germany; National Collection of Arachnida, Pretoria, South Africa.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.

org/10.1016/j.ympev.2023.107798.

References

Allio, R., Donega, S., Galtier, N., Nabholz, B., 2017. Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. Mol. Biol. Evol. 34, 2762–2772.

Ahrens, D., Fujisawa, T., Krammer, H.-J., Eberle, J., Fabrizi, S., Vogler, A.P., 2016. Rarity and incomplete sampling in DNA-based species delimitation. Syst. Biol. 65, 478–494.

Ayoub, N.A., Garb, J.E., Tinghitella, R.M., Collin, M.A., Hayashi, C.Y., 2007. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. PLoS One 2, e514.

Beavis, A.S., Sunnucks, P., Rowell, D.M., 2011. Microhabitat preferences drive phylogeographic disparities in two Australian funnel web spiders. Biol. J. Linn. Soc. 104, 805–819.

Bond, J.E., Hedin, M.C., Ramirez, M.G., Opell, B.D., 2001. Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. Mol. Ecol. 10, 899–910.

Buckley, T.R., Cordeiro, M., Marshall, D.C., Simon, C., 2006. Differentiating between Hypotheses of Lineage Sorting and Introgression in New Zealand Alpine Cicadas (Maoricicada Dugdale). Syst. Biol. 55, 411–425.

Cognato, A.I., Vogler, A.P., 2001. Exploring data interaction and nucleotide alignment in a multiple gene analysis of Ips (Coleoptera: Scolytinae). Syst. Biol. 50, 758–780.

Carstens, B.C., Pelletier, T.A., Reid, N.M., Satler, J.D., 2013. How to fail at species delimitation. Mol. Ecol. 22, 4369–4383.

Correa, C., Vasquez, D., Castro-Carrasco, C., Zuniga-Reinoso, A., Ortiz, J.C., Palma, R.E., 2017. Species delimitation in frogs from South American temperate forests: the case of *Eupsophus*, a taxonomically complex genus with high phenotypic variation. PLoS One 12, e0181026.

De Queiroz, K., 2007. Species concepts and species delimitation. Syst. Biol. 56, 879–886.

Derkarabetian, S., Hedin, M., 2014. Integrative taxonomy and species delimitation in harvestmen: A revision of the western North American genus *Sclerobunus* (Opiliones: Laniatores: Travunioidea). PLoS One 9.

Dinno, A., 2014. paran: Horn's test of Principle Components/ factors.

Dippenaar-Schoeman, A.S., 2002. Baboon and Trapdoor Spiders of Southern Africa: An Identification Manual, Handbook. ed. Plant production research institute. Agricultural Research Council, Pretoria., Pretoria.

Dippenaar-Schoeman, A.S., Haddad, C.R., Foord, S.H., Lyle, R., Lotz, L., Helberg, L., Mathebula, S., van der Berg, A., van Niekerk, E., Jocque, R., 2010. First atlas of the spiders of South Africa (Arachnida: Araneae), in: SANSA Technical Report. p. 1160.

Doyle, J.J., 1992. Gene trees and species trees: Molecular systematic one-character taxonomy. Syst. Bot. 17, 144–163.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 1–8.

Engelbrecht, I., Prendini, L., 2012. Cryptic diversity of South African trapdoor spiders: Three new species of *Stasimopus* Simon, 1892 (Mygalomorphae, Ctenizidae), and redescription of *Stasimopus robertsi* Hewitt, 1910. Am. Mus. Novit. 1–42.

Engelbrecht, I., Prendini, L., 2011. Assessing the taxonomic resolution of southern African trapdoor spiders (Araneae: Ctenizidae; Cyrtaucheniidae; Idiopidae) and implications for their conservation. Biodivers. Conserv. 20, 3101–3116.

Ezard, T., Fujisawa, T., Barraclough, T., 2014. splits: SPecies' LImits by Threshold Statistics.

Felsenstein, J., 1985. Confidence limits on phylogenetics: an approach using the bootstrap. Evol. 783–791.

Ferretti, N., González, A., Pérez-Miles, F., 2014. Identification of priority areas for conservation in Argentina: quantitative biogeography insights from mygalomorph spiders (Araneae: Mygalomorphae). J. Insect Conserv. 18, 1087–1096.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech, 3, 294–299.

Fujita, M.K., Leaché, A.D., 2011. A coalescent perspective on delimiting and naming species: a reply to Bauer *et al*. Proc. Royal Soc. B. 278, 493–495.

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.

Galtier, N., Nabholz, B., Glemin, S., Hurst, G.D.D., 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 18, 4541–4550.

Godwin, R.L., Opatova, V., Garrison, N.L., Hamilton, C.A., Bond, J.E., 2018. Phylogeny of a cosmopolitan family of morphologically conserved trapdoor spiders (Mygalomorphae, Ctenizidae) using Anchored Hybrid Enrichment, with a description of the family, Halonoproctidae Pocock 1901. Mol. Phylogenet. Evol. 126, 303–313.

Hamilton, C.A., Hendrixson, B.E., Brewer, M.S., Bond, J.E., 2014. An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: A case study of the North American tarantula genus *Aphonopelma* (Araneae, Mygalomorphae, Theraphosidae). Mol. Phylogenet. Evol. 71, 79–93.

Harvey, M.S., 2002. Short-range endemism among the Australian fauna: some examples from non-marine environments. Invertebr. Syst. 16, 555–570.

Harvey, M.S., Johnson, M.S., Humphreys, W.F., 2011. Protecting the innocent: Studying short-range endemic taxa enhances conservation outcomes. Invertebr. Syst. 25, 1–10.

Hedin, M., 2015. High-stakes species delimitation in eyeless cave spiders (Cicurina, Dictynidae, Araneae) from central Texas. Mol. Ecol. 24, 346–361.

Hedin, M., Carlson, D., Coyle, F., 2015. Sky island diversification meets the multispecies coalescent – divergence in the spruce-fir moss spider (*Microhexura montivaga*, Araneae, Mygalomorphae) on the highest peaks of southern Appalachia. Mol. Ecol. 3467–3484.

Hendrixson, B.E., Bond, J.E., 2004. A new species of *Stasimopus* from the Eastern Cape Province of South Africa (Araneae, Mygalomorphae, Ctenizidae), with notes on its natural history. Zootaxa 619, 1–14.

Holness, S., Driver, A., Todd, S., Snaddon, K., Hamer, M., Raimondo, D., Daniels, F., Alexander, G., Bazelet, C., Bills, R., Bragg, C., Branch, B., Bruyns, P., Chakona, A., Child, M., Clarke, R.V., Coetzer, A., Coetzer, W., Colville, J., Conradie, W., Dean, R., Eardley, C., Ebrahim, I., Edge, D., Gaynor, D., Gear, S., Herbert, D., Kgatla, M., Lamula, K., Leballo, G., Lyle, R., Malatji, N., Mansell, M., Mecenero, S., Midgley, J., Mlambo, M., Mtshali, H., Simaika, J., Skowno, A., Staude, H., Tolley, K.A., Underhill, L., van der Colff, D., van Noort, S., von Staden, L., 2016. Biodiversity and Ecological Impacts: Landscape Processes, Ecosystems and Species, in: Scholes, R., Lochner, P., Schreiner, G., Snyman-Van der Walt, L., de Jager, M. (Eds.), Shale Gas Development in the Central Karoo: A Scientific Assessment of the Opportunities and Risks. CSIR, Pretor, pp. 7–1 to 7–72.

Katoh, K., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33, 511–518.

Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinformatics. 9, 286–298.

Kornilios, P., Thanou, E., Kapli, P., Parmakelis, A., Chatzaki, M., 2016. Peeking through the trapdoor: historical biogeography of the Aegean endemic spider *Cyrtocarenum* Ausserer, 1871 with an estimation of mtDNA substitution rates for Mygalomorphae. Mol. Phylogenet. Evol. 98, 300–313.

Kremen, C., Cowell, R.K., Erwin, T.L., Murphy, D.D., Noss, R.F., Sanjayan, M.A., 1993. Terrestrial Arthropod Assemblages: Their Use in Conservation Planning. Conserv. Biol.

Kück, P., Meusemann, K., 2010. FASconCAT: Convenient handling of data matrices. Mol. Phylogenet. Evol. 56, 1115–1118.

Leavitt, D.H., Starrett, J., Westphal, M.F., Hedin, M., 2015. Multilocus sequence data reveal dozens of putative cryptic species in a radiation of endemic Californian mygalomorph spiders (Araneae, Mygalomorphae, Nemesiidae). Mol. Phylogenet. Evol. 91, 56–67.

Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.

Mason, L.D., Wardell-Johnson, G., Main, B.Y., 2018. The longest-lived spider: mygalomorphs dig deep, and persevere. Pac. Conserv. Biol. 24, 203.

Michonneau, F., 2017. Using GMYC for species delimitation. Zenodo.

Newton, L.G., Starrett, J., Hendrixson, B.E., Derkarabetian, S., Bond, J.E., 2020. Integrative species delimitation reveals cryptic diversity in the southern Appalachian *Antrodiaetus unicolor* (Araneae: Antrodiaetidae) species complex. Mol. Ecol. 29, 2269–2287.

O'Meara, B.C., 2010. New heuristic methods for joint species delimitation and species tree inference. Syst. Biol. 59, 59–73.

Opatova, V., Arnedo, M.A., 2014. Spiders on a hot volcanic roof: Colonisation pathways and phylogeography of the Canary Islands endemic trap-door spider *Titanidiops canariensis* (Araneae, Idiopidae). PLoS One 9, e115078.

Opatova, V., Hamilton, C.A., Hedin, M., De Oca, L.M., Král, J., Bond, J.E., 2019. Phylogenetic systematics and evolution of the spider infraorder mygalomorphae using genomic scale data. Syst. Biol. syz064.

Papadopoulou, A., Monaghan, M.T., Barraclough, T.G., Vogler, A.P., 2009. Sampling error does not invalidate the yule-coalescent model for species delimitation. A response to lohse (2009). Syst. Biol. 58, 442–444.

Paradis, E., Schliep, K., 2019. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528.

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55, 595–609.

Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.

Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, automatic barcode gap discovery for primary species delimitation. Mol. Ecol. 21, 1864–1877.

R Core Team, 2017. R: A language and environment for statistical computing.

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.

Rix, M.G., Raven, R.J., Harvey, M.S., 2018. Systematics of the giant spiny trapdoor spiders of the genus *Gaius* Rainbow (Mygalomorphae: Idiopidae: Aganippini): documenting an iconic lineage of the Western Australian inland arid zone. J. Arachnol. 46, 438–472.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Rubinoff, D., Holland, B.S., 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst. Biol. 54, 952–961.

Satler, J.D., Carstens, B.C., Hedin, M., 2013. Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*). Syst. Biol. 62, 805–823.

Satler, J.D., Starrett, J., Hayashi, C.Y., Hedin, M., 2011. Inferring species trees from gene trees in a radiation of Californian trapdoor spiders (Araneae, Antrodiaetidae, *Aliatypus*). PLoS One 9, E25355.

Sethusa, T., 2016. Karoo BioGaps Project [WWW Document]. SANBI [ONLINE]. URL https://www.sanbi.org/karoo-biogaps-project/ (accessed 2.7.18).

Simon, E., 1892. Histoire naturelle des araignées. Deuxième Édition, Tome Premier, pp. 1–256.

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, 651–702.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 4.

Sukumaran, J., Knowles, L.L., 2017. Multispecies coalescent delimits structure, not species. PNAS 114, 1607–1612.

Swofford, D.L., 2002. PAUP* Phylogenetic Analysis Using Parsimony (*and other methods), Version 4. ed. Sinauer Associates, Sutherland, Massachusetts.

Wilson, J.D., Hughes, J.M., Raven, R.J., Rix, M.G., Schmidt, D.J., 2018. Spiny trapdoor spiders (*Euoplos*) of eastern Australia: Broadly sympatric clades are differentiated by burrow architecture and male morphology. Mol. Phylogenet. Evol. 122, 157–165.

World Spider Catalog, 2022. World Spider Catalog Version 23.5 [WWW Document]. Natural History Museum Bern. https://doi.org/10.24436/2. (Accessed 5.6.22).

Xi, Z., Liu, L., Davis, C.C., 2016. The impact of missing data on species tree estimation. Mol. Biol. Evol. 33, 838–860.

Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A., 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29, 2869–2876.