Systematics of the dung beetle tribe Sisyphini Mulsant (Scarabaeidae: Scarabaeinae) inferred from a molecular phylogeny and biogeography of southern African species

GIMO M. DANIEL[1*], CATHERINE L. SOLE[1**], ADRIAN L.V. DAVIS[1], WERNER P. STRÜMPHER[2] & CLARKE H. SCHOLTZ[1]

[1]*Department of Zoology & Entomology, University of Pretoria, Private Bag X20, Hatfield, 0028 South Africa. Corresponding authors: Email:* [*]*gimodaniel@gmail.com*; [**]*catherine.sole@up.ac.za*

[2]*Agricultural Research Council (ARC), Plant Health and Protection, Private Bag X134, Queenswood, 0121 South Africa.*

Running title: Sisyphini: assessment of evolution and systematics

Abstract. The tribe Sisyphini Mulsant was recently redefined following the transfer of the endemic southern African genus *Epirinus* Dejean from the polyphyletic tribe Deltochilini Lacordaire. A molecular phylogeny of the southern African members of Sisyphini supports *Epirinus* as sister to *Sisyphus* Latreille and recovered three major clades in *Sisyphus* classified here as subgenera *Sisyphus* (*Neosisyphus* Müller) **stat. rev**., *Sisyphus* (*Parasisyphus* Barbero, Palestrini & Zunino) **stat. n**. and *Sisyphus* (*Sisyphus*) **stat. n**. A molecular clock analysis suggests that *Sisyphus* and *Epirinus* diverged from their last common ancestor during the Lower to Middle Oligocene (ca. 29.37 Ma). Biogeographical analysis indicated that southern African *Sisyphus* species are centred in the east and northeast in Highveld grassland and warmer savannah regions. By contrast, *Epirinus* species are largely restricted to the southwest and southeast in the cooler winter and bimodal rainfall regions plus arid highland Karoo and Highveld grasslands. Based on morphological and biogeographical differences between *Epirinus* and *Sisyphus,* we propose that the monogeneric *Epirinus* be placed in its own tribe, Epirinini van Lansberge **stat. rev.**

Key-words: Distribution pattern – *Epirinus* – molecular systematics –– *Sisyphus*

Introduction

Dung beetles in the tribe Sisyphini Mulsant represent some of the most charismatic members of the subfamily Scarabaeinae (Coleoptera: Scarabaeidae) in southern Africa, notable for their distinctive dung rolling behaviour. The first phylogeny of the subfamily was proposed by Zunino (1983) based on morphology of the male and female genitalia. He hypothesized a clade of ball-rolling tribes in which Sisyphini was placed as sister to Canthonini van Lansberge (=Deltochilini Lacordaire); this was in turn sister to Gymnopleurini Lacordaire and Scarabaeini Latreille. These relationships were corroborated in the subsequent study by Luzzato (1994). Villaba *et al.* (2002) proposed the first molecular phylogeny of Scarabaeinae and recovered a clade in which Sisyphini was sister to Coprini Leach, Gymnopleurini and Scarabaeini. A more comprehensive morphological phylogeny by Philips *et al*. (2004) included both *Sisyphus* Latreille and *Neosisyphus* Müller and placed them as sister-genera in a clade also including Eurysternini Vulcano, Martínez & Pereira, Onitini Castelnau, Onthophagini Burmeister and Oniticellini Kolbe. Although most recent studies show a degree of consistency in proposed relationships within the clade, several others differ quite significantly. The large-scale molecular phylogeny of Monaghan *et al.* (2007) recovered Sisyphini as the sister clade to *Epirinus* Dejean (Deltochilini), although Sisyphini were also placed as sister to Onitini under different analytical parameters. Based on wing shape of Chinese dung beetles, Bai *et al.* (2011) proposed a phylogeny in which Sisyphini was the sister clade to Oniticellini. Still, most recent phylogenies using either morphological or molecular data have consistently recovered a sister relationship between Sisyphini and *Epirinus* (i.e., Philips *et al*., 2004; Monaghan *et al.,* 2007; Mlambo *et al*., 2015; Tarasov & Génier, 2015; Tarasov & Dimitrov, 2016). Indeed, in their large-scale molecular phylogenetic analysis of Scarabaeinae, Tarasov & Dimitrov (2016) expanded the concept of Sisyphini to also include *Epirinus*. As a result, present membership of Sisyphini comprises six genera: *Epirinus*, *Nesosisyphus* Vinson, *Neosisyphus*, *Sisyphus, Parasisyphus* Barbero, Palestrini & Zunino and *Indosisyphus* Barbero, Palestrini & Zunino.

Sisyphini is distributed largely throughout the Afrotropical region, with some species in the Palaearctic and Oriental regions (Davis *et al*., 2016b). However, some genera in Sisyphini show contrasting geographical distribution patterns. Whereas

Neosisyphus (Afrotropical and Oriental) and *Sisyphus* (Afrotropical, Palaearctic, Oriental, and central America) are widespread; *Epirinus* (southern Africa), *Indosisyphus* (Oriental) *Nesosisyphus* (Mauritius) and *Parasisyphus* (Afrotropical) show more restricted distributions (Barbero *et al.*,1991; Davis *et al.*, 2016b; Tarasov & Dimitrov, 2016). Such radically differing geographical patterns could imply quite different histories of adaptive radiation or extinction over time, particularly for *Epirinus* versus Sisyphini. The present paper examines the phylogeny, divergence times, biogeography of Sisyphini based on a molecular phylogeny of the southern African species and proposes a revised classification of the clade to comprise two tribes (Sisyphini and Epirinini **stat. rev.**) rather than a single tribe, Sisyphini. Moreover, based on these results, the genera *Neosisyphus* and *Parasisyphus* are reduced to subgenera within *Sisyphus*.

Materials and Methods

Sampling, amplification, processing of sequences and alignment

Representatives of three of the six extant genera of Sisyphini were obtained for phylogenetic analysis, including *Sisyphus*, *Neosisyphus* and *Parasisyphus* collected in South Africa and Mozambique. Additionally, DNA sequences of 16 species of *Epirinus* were sourced from Mlambo *et al.* (2011) and downloaded from Genbank (accession number: GQ289704.1–HQ290004.1). Collected individuals of Sisyphini were identified by two of the authors (GMD & ALVD) and voucher specimens are kept in an alcohol reference collection at the Department of Zoology & Entomology, University of Pretoria, South Africa (UPSA). *Nesosisyphus* and *Indosisyphus* are rarely collected genera with restricted distributions and were therefore not included in the study since we were unable to obtain suitable material for DNA extraction.

Four recent phylogenetic studies of the subfamily Scarabaeinae place the tribe Sisyphini as sister to *Epirinus.* In addition, they variously infer sister relationships between Sisyphini and several other tribes, including Coprini, Gymnopleurini, and Eurysternini (Monaghan *et al.*, 2007; Mlambo *et al.*, 2015; Tarasov & Génier, 2015; Tarasov & Dimitrov, 2016). Based on these findings and available molecular sequences in GenBank, we used outgroup taxa composed of *Eurysternus caribaeus*

(Herbst) (accession number: AY131893.1, AY131536.1, AY131725.1); *Eurysternus inflexus* Germar (accession number: AY131726.1, AY131538.1, AY131895.1); *Gymnopleurus virens* Erichson (accession number: AY131731.1, AY131543.1, AY131900.1); and *Heliocopris hamadryas* (Fabricius) (accession number: GQ289971.1, AY131878.1, AY131519.1, AY131708.1). Total genomic DNA was extracted from all Sisyphini samples using the Roche High Pure PCR Template Preparation Kit (Roche Diagnostics, Penzeberg, Germany). We amplified four gene regions; these comprised two nuclear genes: CAD (carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase) and 28S rDNA (28S rDNA domain 2) and two mitochondrial genes: 16S (16S rDNA) and COI (cytochrome c oxidase subunit I). Primers used for PCR amplification are listed in Table S1. All PCR amplifications of genes were performed in a total volume of 25 µl. Amplification mixtures contained Emerald Amp®MAX HS PCR Mastermix (TAKARA BIO INC., Otsu, Shiga, Japan), 10 pmol of each primer (forward and reverse) and 50-100ng of extracted DNA template. Distilled water was used to ensure the mixture reached a total volume of 25µl. Successful amplifications were purified using the Roche High Pure Product Purification Kit (Roche Diagnostics, Penzeberg, Germany) following the manufacturer's specifications. To obtain DNA sequences, the cycle sequencing reactions were carried out in both directions using the BigDye Terminator v 3.1 Cycle Sequencing kit (Applied Biosystems, Foster City, CA, USA). Cycle sequencing products were precipitated using a standard sodium acetate/ethanol precipitation. All generated sequences were viewed, assembled and edited in CLC Main workbench version 7.0 (developed by CLC Bio, http://www.clcbio.com). New sequences were submitted to GenBank (Table S2). The sequences were aligned using default settings of the online platform, MAFFT (Katoh & Toh, 2008).

Additional examined material

Additional dry specimens used in this study were loaned from the following institutions: South African National Collection of Insects, Plant Protection Research Institute, Pretoria, South Africa (SANC); Ditsong Museum of Natural History, Pretoria, South Africa (previously the Transvaal Museum) (TMSA) and Iziko South African Museum, Cape Town, South Africa (SAMC).

Partitioning

We partitioned our data sets using PartitionFinder software v 2.1 (Lanfear *et al.*, 2016). The appropriate model selection and partitioning (Table 1) was determined under the corrected Akaike Information Criterion (AICc). According to Kainer & Lanfear (2015) partitioning involves two steps: (1) defining groups of sites that are assumed to have evolved in similar ways; and (2) choosing an appropriate model of molecular evolution for each group of sites. The first step in partitioning involves the assignment of each site in an alignment to a data block. Data blocks are user-defined sets of sites, typically encompassing distinct DNA features such as genes, introns, exons, and codon positions. We identified codon positions 1st, 2nd and 3rd for the two protein coding genes (COI and CAD). The non-coding genes (16S and 28S) are regarded as a single data block (Kainer & Lanfear 2015).

Table 1. Data from PartitionFinder v 2.1: Subset partitions and best model used for Bayesian inference analysis. The subset partitions were used for Maximum Likelihood and estimate time divergence analyses.

Subset partition definitions	Partitions name	Best model	Sites
Subset 1 = 1-652\3;	CAD_pos1	GTR+I+G	218
Subset 2 = 2-652\3;	CAD_pos2	GTR+I+G	217
Subset 3 = 3-652\3;	CAD_pos3	GTR+G	217
Subset 4 = 653-1355\3;	COI_pos1	GTR+I+G	235
Subset 5 = 654-1355\3;	COI_pos2	GTR+I+G	234
Subset 6 = 655-1355\3;	COI_pos3	GTR+I+G	234
Subset 7 = 1356-1668;	16S	GTR+G	313
Subset 8 = 1669-2124;	28S	GTR+G	456

Maximum likelihood

Maximum likelihood analyses were performed on each gene individually and on the total concatenated data set. All ML analyses were implemented in RAxML v 8.2.4 (Stamatakis, 2014). Since RAxML allows only a single model of rate heterogeneity in partition analyses, we implement the General Time Reversible (GTR) model of nucleotide substitution under the GAMMA model of rate heterogeneity (Stamatakis, 2014). Nodal support confidence of the majority-rule consensus tree topology was

estimated from 1000, non-parametric, bootstrap replicates of likelihood (Felsenstein, 1981, 1985).

Bayesian inference

Individual gene and concatenated phylogenies were also estimated via Bayesian Inference in MrBayes v 3.2.6 (Ronquist & Huelsenbeck, 2003). To find the best-fit partitioning schemes and models of evolution, we used PartitionFinder (Table 1). Flat Dirichlet priors were used in all analyses. Bayesian analyses were conducted by simultaneously running two Metropolis-coupled Monte Carlo (MCMC) Markov chains for 30 million iterations. Trees were sampled every 200th iteration. The first 25% of sampled trees were discarded as burn-in to ensure that the analysis had converged properly, which was determined by TRACER v 1.6 (Rambaut & Drummond, 2014). The credibility of clade support was provided by posterior probability estimation. The tree topology in both analyses (ML and BI) were visualised in FigTree v 1.4.3 (Rambaut, 2009).

Diversification time estimation

The node ages for the major lineage-splitting events within Sisyphini were estimated using BEAST v 2.4.5 (Bouckaert *et al.,* 2014). As there is no fossil record for the tribe, the absolute divergence time is difficult to calibrate (Schenk, 2016). In such cases, the solution is (1) to infer divergence times by applying a substitution rate estimated from studies of close relatives (Ho, 2007; Weir & Schluter, 2008), or (2) to use secondary calibrations based on previous molecular-dating of fossil relatives, or, the age estimated for the earliest reliable fossil of the closest relative (Shaul & Graur, 2002). Therefore, in the present study we estimated node ages using the oldest valid scarabaeine fossil (tribe Ateuchini Perty: *Lobateuchus parisii* Montreuil, Génier & Nel), in which the estimated age is 53 Ma (Tarasov *et al.*, 2016).

As suggested for secondary calibrations (Heath, 2012), we used a Bayesian strict molecular clock analysis under a normal distribution model approach. The Yule speciation process was applied for all combined data. The combined data set was partitioned using PartitionFinder (see Table 1). However, we removed those coding partitions, which refer to the same nucleotides, such as COI_pos2 and CAD_pos2.

We implemented a separate GTR + G substitution for each partition, following the BEAST v 2.4.5 default settings (Bouckaert *et al.*, 2014). Two independent MCMC analyses were run for 30 million generations with parameters sampled in each 200th generation. The first 25% of trees sampled in each run were discarded as burn-in. The program, LogCombiner (BEAST 2 package), was applied to combine the log and tree output files from the two independent runs. TRACER v 1.6 was used to assess the convergence between runs. The program, TreeAnnotator (BEAST 2 package), was used to generate the consensus tree and determine the mean ages under 95% highest posterior density (HPD). The tree topology was visualized in FigTree v 1.4.3.

Geographical distribution

The geographical distributions of the genera, *Epirinus* and *Sisyphus* were plotted on a map panel of southern Africa (17ºS x 33ºE) according to presence records in ~15 x 15 km polygons (=1/16th degree squares). Further differences in generic/subgeneric and species distribution are explored in the discussion as regards both global and southern African patterns. This discussion is based on published literature and a distributional database for all species in southern Africa available on the web (http://vmus.adu.org.za/).

Biogeographical analysis

Biogeographical analysis was based on five climatic regions defined for dung beetle distribution in southern Africa (Davis 1997); (A) Highveld; (B) Kalahari; (C) Northeast Savannah; (D) South-West Arid; (E) Winter/Bimodal Rainfall. Part of the outgroup taxa (*Eurysternus* spp.) occur in the Neotropical region (F) (Morrone 2014). Each species was assigned to an area/areas according to its current known distribution range. Biogeographical inferences were obtained by statistical dispersal–vicariance analysis (S-DIVA) (Yu *et al*., 2015) implemented in RASP (version 3.0) at the default settings (Yu *et al*., 2010). Ancestral area probabilities were estimated to test whether vicariance or dispersal was dominant, and which ancestral area per node was most likely.

Results

The combined datasets for four gene regions comprised a total of 2124 bp (base pairs); COI = 703 bp; CAD = 652 bp; 28S (D2) = 456 bp; 16S = 313 bp.

Phylogenetic relationships

Sisyphini formed a strongly supported monophyletic group (Bayesian posterior probability (PP) 0.97 and Maximum Likelihood bootstrap (MLB) 89%). It comprised two main clades labelled I (*Epirinus*) (0.98 PP; 90% MLB) and II (remaining genera) (1.00 PP; 100% MLB). Clade II was subdivided into a further two clades: clade A containing *Parasisyphus* (1.00 PP; 95% MLB) and clade B comprising *Sisyphus* and *Neosisyphus* (1.00 PP; 96% MLB). It is clear here that *Sisyphus* was recovered as paraphyletic with respect to *Parasisyphus* and *Neosisyphus*. Therefore, we propose a revised classification of *Sisyphus* with *Neosisyphus* **stat. rev**. and *Parasisyphus* **stat. n.** regarded as subgenera within *Sisyphus* (Fig. 1).

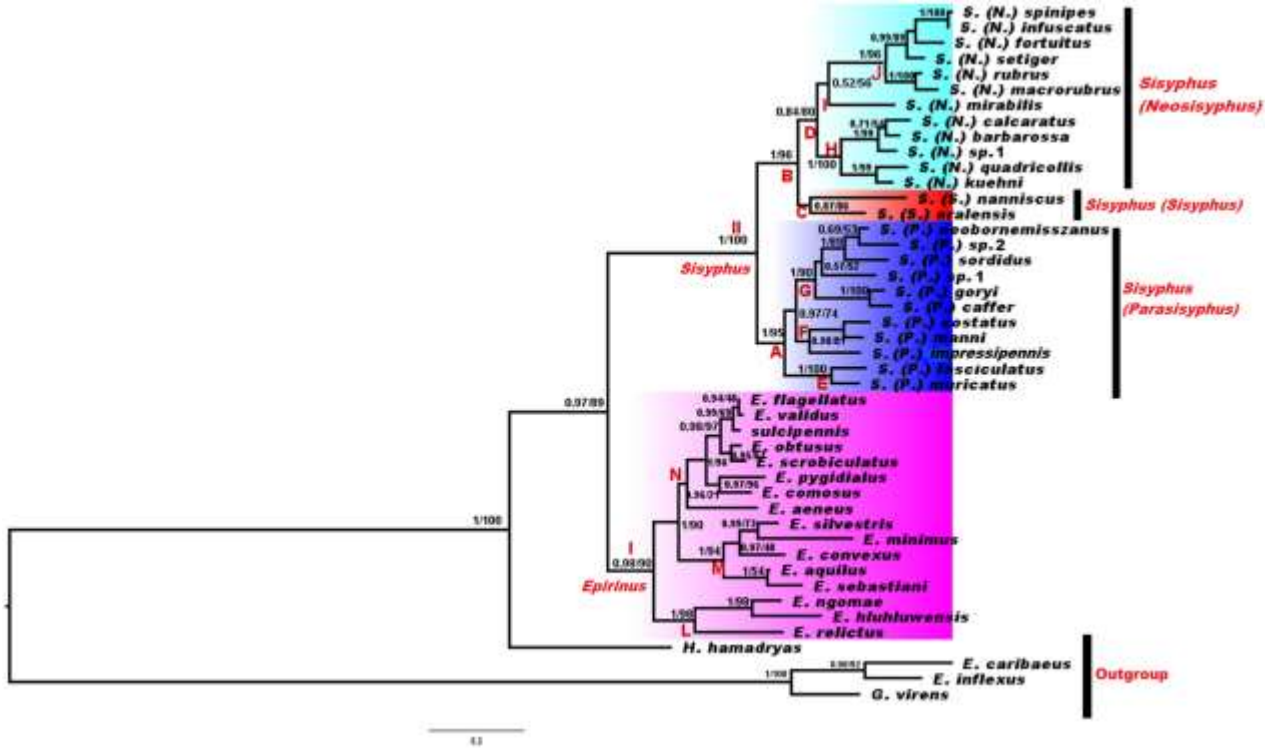


Fig. 1. Phylogram of combined data set analysis (COI, 16S, CAD and 28S domain 2) for Sisyphini. Maximum likelihood bootstrap (MLB) and Bayesian posterior probabilities (PP) are presented for lineages (I – *Epirinus*) and (II – *Sisyphus*). See supplementary material for separate Bayesian (Figure S1) and Maximum likelihood (Figure S2) phylograms.

The genus *Epirinus* was recovered as a well-supported monophyletic taxon (0.98 PP; 90% MLB) with three major subclades (Fig. 1). Subclade L (1.00 PP; 98% MLB) was recovered as sister to the rest of the species in the genus; containing a volant species (*E. relictus* Scholtz & Howden) and two non-volant species (*E. ngomae* Medina & Scholtz and *E. hluhluwensis* Medina & Scholtz). Subclade M (1.0 PP; 94% MLB) comprised non-volant species: *E. aquilus* Scholtz & Howden, *E. sebastiani* Scholtz & Howden, *E minimus* Scholtz & Howden, *E. silvestris* Cambefort and *E. convexus* Scholtz & Howden. While the poorly supported subclade N (0.96 PP; 21% MLB) was composed of volant species including *E. aeneus* Wiedemann, *E. pygidialus* Scholtz & Howden, *E. comosus* Péringuey, *E. obtusus* Boheman, *E. scrobiculatus* Harold, *E. sulcipennis* Boheman, *E. flagellatus* (Fabricius) and *E. validus* Péringuey.

Results for both Bayesian inference and Maximum likelihood indicated that clade A, represented by *Sisyphus* (*Parasisyphus*) **stat. n.** may be divided into three well-supported subclades. Firstly, *Sisyphus* (*Parasisyphus*) *muricatus* (Olivier) and *S.* (*P.*) *fasciculatus* Boheman (subclade E) were recovered as sister to each other and the rest of the clade (1.00 PP; 100% MLB). Subclade F comprised three species, *S.* (*P.*) *impressipennis* van Lansberge, *S.* (*P.*) *manni* Montreuil and *S.* (*P.*) *costatus* (Thunberg) (0.98 PP; 81% MLB) and the third subclade (G) comprised five species: (*S.* (*P.*) *caffer* Boheman, *S.* (*P.*) *goryi* Harold, *Sisyphus* (*P.*) sp.1, *S.* (*P.*) *neobornemisszanus* Daniel & Davis, *S.* (*P.*) *sordidus* Boheman and *Sisyphus* (*P.*) sp.2 (1.00 PP; 90% MLB) (Fig. 1).

Clade B had a distinct basal dichotomy represented by two subgenera *Sisyphus* (*Sisyphus*) and *Sisyphus* (*Neosisyphus*) **stat. rev.** Clade C was composed of two species of *S*. (*Sisyphus*): *Sisyphus* (*Sisyphus*) *oralensis* Daniel & Davis, *S.* (*S.*) *nanniscus* Péringuey (0.87 PP; 86% MLB) and was sister to the rest of the group. Clade D comprised species of *(S.) Neosisyphus* **stat. rev.** and may be divided into three subclades (labelled H–J). Subclade H comprised five species (1.00 PP; 100% MLB): (*Sisyphus* (*Neosisyphus*) *kuehni* Haaf, *S.* (*N.*) *quadricollis* Gory, *S.* (*N.*) sp.1, *S.* (*N.*) *calcaratus* (Klug) and *S.* (*N.*) *barbarossa* Wiedemann). A single species *S.* (*N.*) *mirabilis* Arrow (0.52 PP; 56% MLB) comprised subclade I. Subclade J

contained six species (1.00 PP; 96% MLB): (*S.* (*N.*) *rubrus* Paschalidis, *S.* (*N.*) *macrorubrus* Paschalidis, *S.* (*N.*) *fortuitus* Péringuey, *S.* (*N.*) *setiger* Roth, *S.* (*N.*) *infuscatus* Klug and *S.* (*N.*) *spinipes* (Thunberg)).

Divergence time estimates

Relative to the selected outgroup genera, *Sisyphus* and *Epirinus* diverged from their last common ancestor during the Lower to middle Oligocene (29.37 Ma; 95% HPD interval: 33.9 to 23.0 Ma). Diversification of extant species in the genus *Epirinus* (I) occurred in the Early Miocene (21.83 Ma; 95% HPD interval: 29.0 to 18.0 Ma) whereas that of *Sisyphus* (II) occurred in the Lower to Middle Miocene (16.86 Ma; 95% HPD interval: 21.0 to 10 Ma). Divergence within the genus *Sisyphus* (clades A-C) are estimated to have originated during the Middle to Upper Miocene with the subgenera *Parasisyphus* ≈ 12.57 Ma, *Sisyphus.* ≈ 10.91 Ma and *Neosisyphus* ≈ 11.5 Ma (95% HPD interval: 13.0 to 8.0 Ma) (Fig. 2).

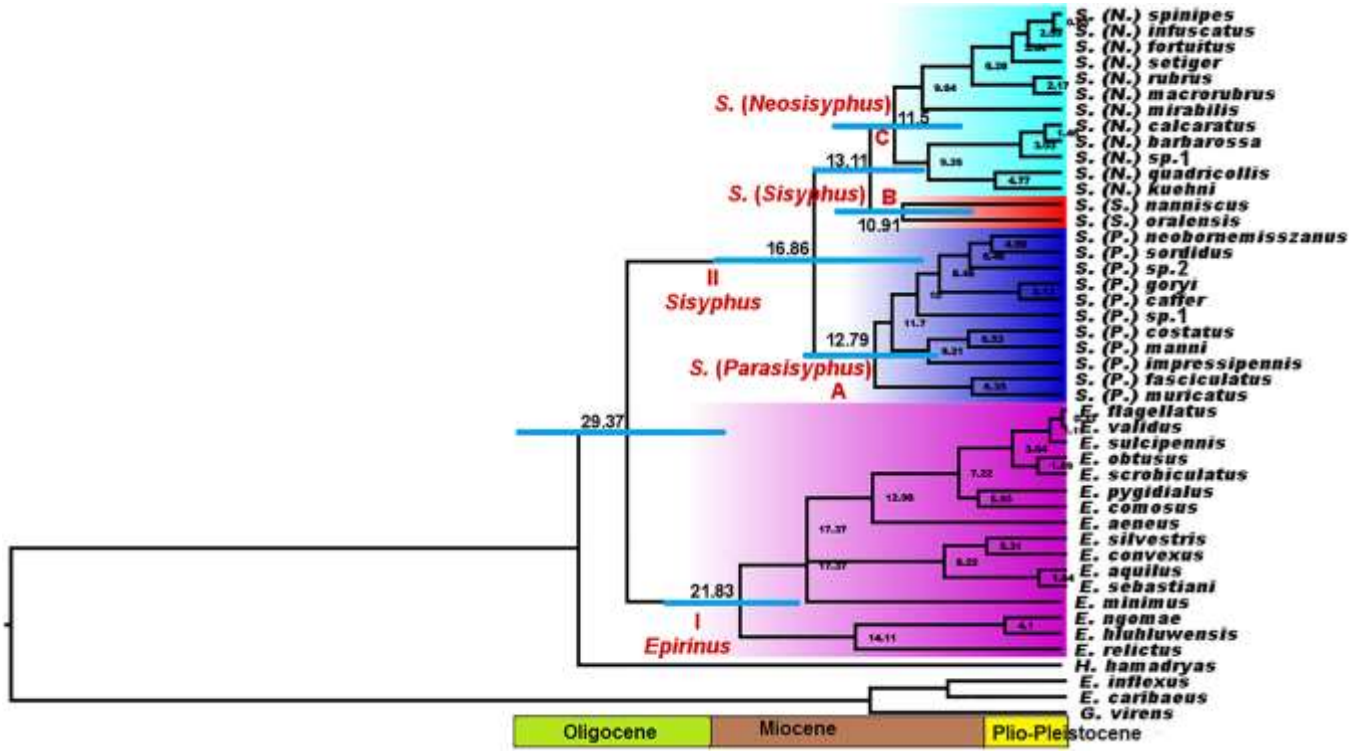


Fig. 2. Relative estimated time of divergence for the major lineages of Sisyphini. The blue bars in the main nodes represent the time intervals for the 95% probability of actual age. Values at nodes represent mean estimated ages of divergence. The geological time scale represents millions of years ago (Ma).

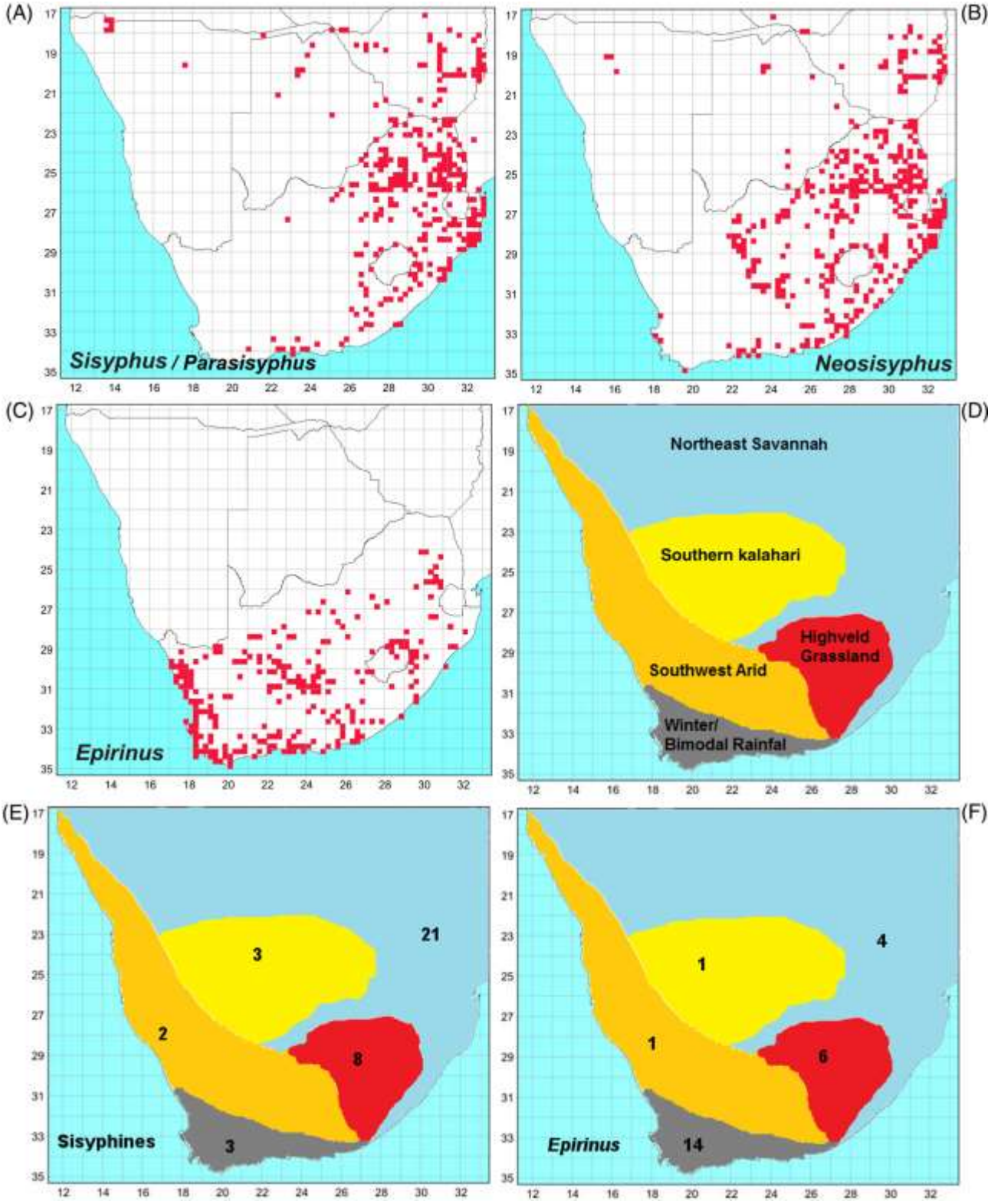


Fig. 3. Southern African species distribution: (A) *Sisyphus/Parasisyphus*, (B) *Neosisyphus,* (C) *Epirinus* and (D) Main climatic regions of southern Africa, representing five climatic biogeographical regions for dung beetles, modified from Davis (1997). Number of species in different climatic regions showing biases in occurrence to the northeast (sisyphines) versus southwest and southeast (*Epirinus*): (E) sisyphines and (F) *Epirinus*

Current geographical distribution

S-DIVA analysis suggests a complex distributional history for both genera; mostly dominated by dispersal as the driver of these patterns. Quite different patterns of geographical distribution are shown in southern Africa by *Epirinus* and the three subgenera of *Sisyphus* (Figs 3A–D, 5). *Epirinus* data are restricted to South Africa where they largely coincide with winter and bimodal rainfall, southern southwest arid and Highveld regions (Fig. 3F). Within southern Africa, data for *Sisyphus* species show a strong eastern bias in distribution which largely coincides with Highveld and savannah regions (Fig. 3E) (see also Table S3, 4,5).

S-DIVA analysis indicated several distribution patterns for *Epirinus.* Dominant patterns were shown by species centred on the winter and bimodal rainfall regions (10 spp.) including non-volant taxa (*E. aquilus*, *E. sebastiani*, *E. silvestris*, *E. convexus* and *E. minimus*) associated with southeastern Afrotemperate forests with other volant taxa in shrublands or grasslands of the southwest (*E. aeneus* and *E. flagellatus*), south (*E. comosus* and *E. sulcipennis*) or west coast sands (*E. scrobiculatus*). Four volant species were associated, especially, with Highveld grasslands from east to northeast (*E. relictus* and *E. validus*), southeast grasslands (*E. obtusus*) or northeast forest (*E. pygidialus*). Two other non-volant species were associated with lower-laying northeast forest (*E. ngomae* and *E. hluhluwensis*) (Fig. 4).

Distribution of *Sisyphus* in southern Africa was dominated by northeast bias in grassland to woodland savannah or forest (24 out of 25 spp.). Of the named species, seven showed a northeast savannah distribution (*S.* (*N.*) *calcaratus*, *S.* (*N.*) *fortuitus* – shade*, S.* (*P.*) *goryi*, and *S.* (*P.*) *impressipennis* – shade) or northeast to southeast savannah bias (*S.* (*N.*) *infuscatus*, *S.* (*N.*) *rubrus* and *S.* (*N.*) *spinipes*). Seven species showed a primarily northeast coast or northeast lowland distribution in shaded savannah (*S.* (*S.*) *nanniscus*), open savannah (*S.* (*P.*) *sordidus*) or forest (*S.* (*P.*) *fasciculatus*, *S.* (*P.*) *neobornemisszanus*, *S.* (*S.*) *oralensis*, *S* and *S*. (*N.*) *mirabilis* - also southeast coast). Eight species are centred on the Highveld, four across the moist northeast grasslands (*S.* (*P.*) *caffer*, *S.* (*P.*) *costatus, S.* (*P.*) *manni* and *S.* (*N.*) *setiger* – also coastal), three along the east escarpment to south coast

(*S*. (*N.*) *barbarossa*, *S*. (*N.*) *kuehni* and *S.* (*P.*) *muricatus*) and one on the arid southwest Highveld with outlier occurrence on the northern plateau of Namibia (*S*. (*N.*) *macrorubrus*). Only one of the studied species was restricted to the south coast of South Africa (*S*. (*N.*) *quadricollis*) (Fig. 4).

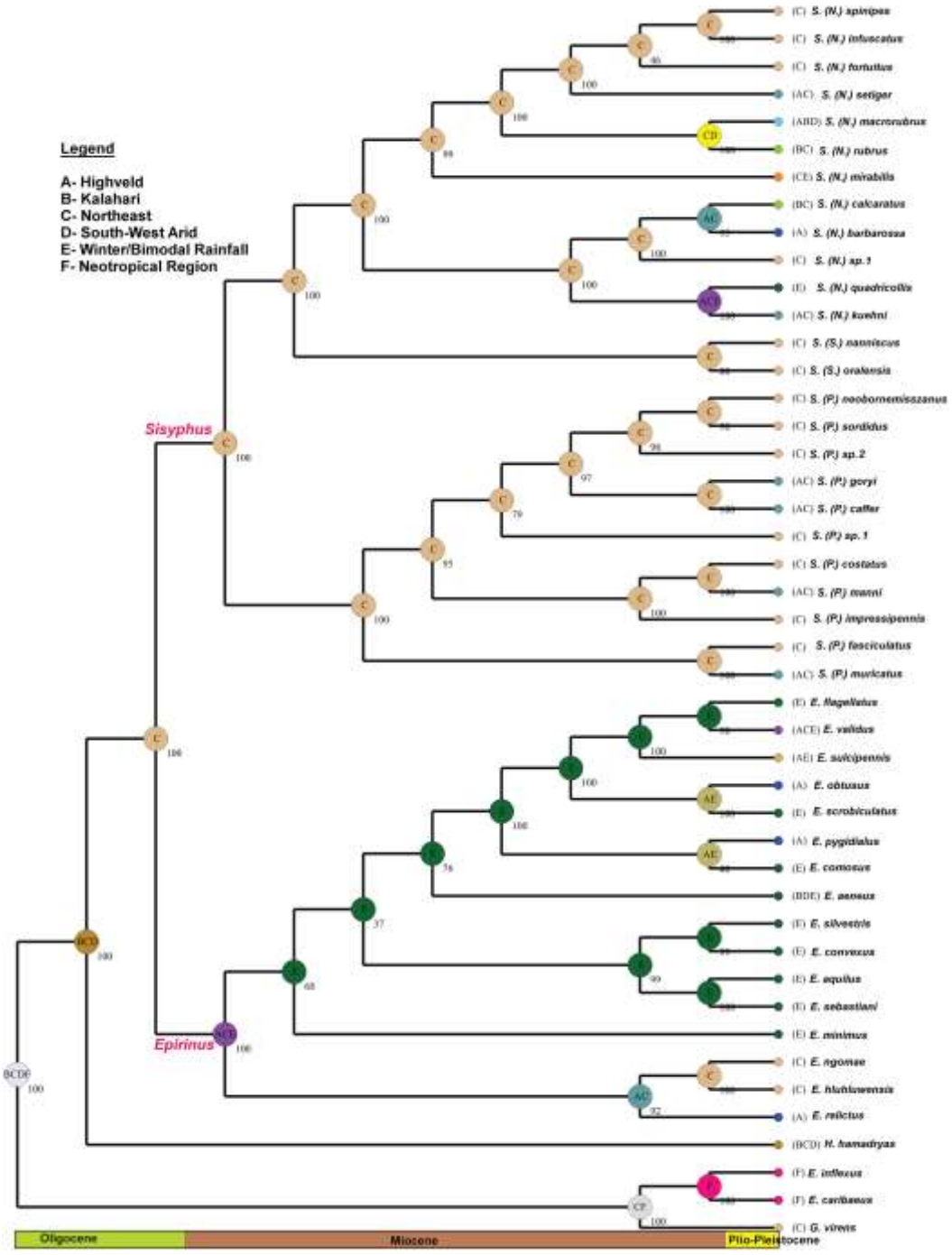


Fig. 4. Historical biogeography of the genera *Sisyphus* and *Epirinus* using S-DIVA. Coloured boxes identify biogeographical regions: A – Highveld; B – Kalahari; C – Northeast; D – South-West Arid; E – Winter/Bimodal and F – Neotropical. Pie charts show relative probabilities for ancestral area by colour.

Discussion

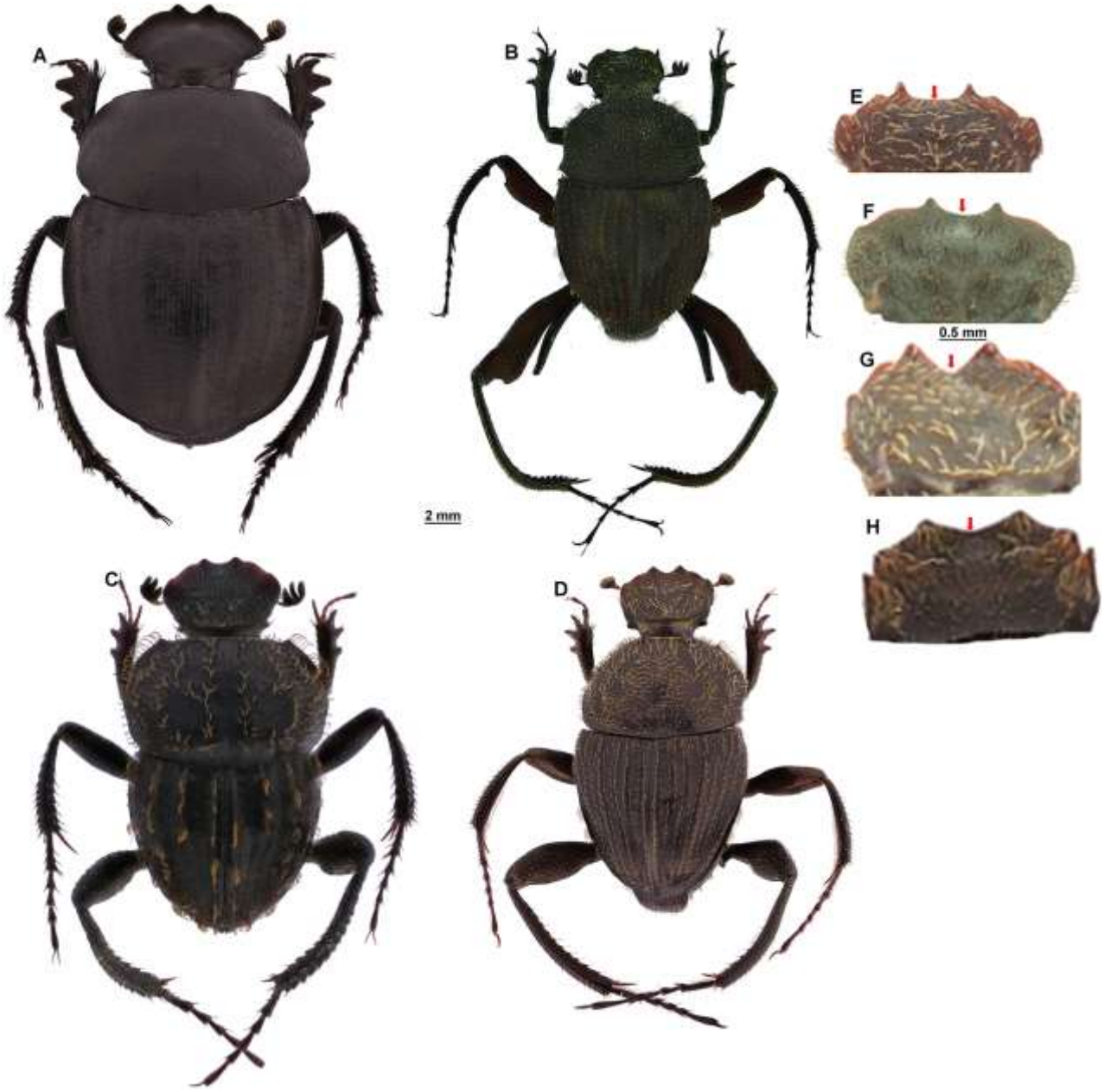
Systematics and molecular phylogeny of Sisyphini

Our molecular phylogeny of Sisyphini yields two major findings. Firstly, there is a distinct basal dichotomy between *Epirinus* and *Sisyphus.* Secondly, the subdivision of Sisyphini into clades A and B suggests a degree of paraphyly in species described within the genus *Sisyphus*. Furthermore, clade A topology has implications for characters used to define species groups within *Sisyphus*. These findings have important implications for the classification of the group as discussed below.

The present molecular analysis supports a strong separation between the sister clades of *Epirinus* and Sisyphini supporting previous morphological (Medina & Scholtz, 2005) and molecular studies (Mlambo *et al*., 2011), which suggest monophyly of *Epirinus*. This is based on the following synapomorphies: shape of the internal border of the eye oblique, with a carina running posteriorly; central plate of male genitalia with short projections; and presence of a ring shaped sclerite X (equivalent to superior right peripheral sclerite) in the internal sac (Medina & Scholtz, 2005). In addition, our data support previous findings that suggest the non-volant *Epirinus* species do not represent a monophyletic clade. As flightless *Epirinus* are primarily found in forests, such adaptive convergence appears to have evolved independently more than once in response to past environmental changes (Mlambo *et al.,* 2011).

Our phylogeny indicates that clade B (Fig. 1) comprises both *Sisyphus* (*Sisyphus*) and *S.* (*Neosisyphus*) species. Historically, *Sisyphus* was divided into two subgenera based on the presence of a complete *Sisyphus* (*Sisyphus*) or incomplete *Sisyphus* (*Neosisyphus*) lateral pronotal ridge (Müller, 1942). However, recent taxonomic and phylogenetic studies have raised *Neosisyphus* to generic rank (Daniel *et al*., 2016, 2018; Tarasov & Dimitrov, 2016; Tarasov & Génier, 2015; Montreuil, 2015a); results that are not supported here. The placement of clade C as sister to clade D suggests that the classic diagnostic character for *S*. (*Neosisyphus*), the incomplete lateral pronotal ridge (Daniel *et al*., 2016, 2018; Montreuil, 2015a; Müller, 1942) is homoplastic. Support for inconsistency in this character is provided by the holotype specimen of *S*. (*Neosisyphus*) *youngai* (Endrödi), in which the lateral pronotal edge

is complete, as in *S*. (*Sisyphus*). Such inconsistency in a principal character used to separate the two genera as defined by Müller (1942) suggests that the current classification within Sisyphini requires revision in a phylogenetic context.



.

Fig. 5. Morphological differences in habitus between genera: (A) *Epirinus validus* Péringuey, (B) *Sisyphus* (*Neosisyphus*) *kuehni*, (C) *S.* (*Parasisyphus*) *swazi* Daniel & Davis and (D) *S.* (*Sisyphus*) *umbraphilus* Daniel & Davis. Edge between medial clypeal teeth in three subgenera of *Sisyphus:* (E, F) *Parasisyphus*, (G) *Sisyphus*) and (H) *Neosisyphus*.

Although our molecular phylogeny is based only on southern African species, it does represent a reliable, albeit, partial evolutionary hypothesis for sisyphines. In order to maintain a natural classification, we propose that the group should be represented by a single genus, *Sisyphus*, subdivided into three subgenera *Parasisyphus* **stat. n**., *Sisyphus* **stat. rev**. and *Neosisyphus* **stat. rev**. as supported by clades; A, C, D (Fig. 1). The subgeneric categories herein proposed can be morphologically delimited by the clypeal shape as follows: *S*. (*Parasisyphus*) bears a straight or slightly concave edge between medial clypeal teeth (Fig. 5E-F); whereas, that of *S*. (*Sisyphus*) is strongly v-shaped (Fig. 5G). On the other hand, *S*. (*Neosisyphus*), the margin between the medial clypeal teeth comprises an obtuse angle (Fig. 5H).

In some species of *S*. (*Parasisyphus*) tufts of setae are borne on the elytra. In our phylogeny, these tufted species (i.e., *S. (P.) muricatus, S. (P.) fasciculatus*, *S. (P.) manni* and *S. (P.) neobornemisszanus*) are interspersed across clade A with non-tufted species (Fig. 1). This suggests that tufts of setae may have evolved or been lost, independently, several times during the evolutionary history of *Sisyphus.* Therefore, contrary to proposals by Daniel *et al*. (2016, 2018) and Montreuil (2015b), tufts of setae do not appear to be suitable as a key diagnostic character for any species-group.

Systematics of Sisyphini

The most recent and comprehensive molecular phylogeny of Scarabaeinae by Tarasov & Dimitrov (2016) demonstrated a monophyletic clade comprising *Sisyphus*, *Neosisyphus* and *Epirinus*. Based on these findings, the authors expanded the limits of Sisyphini to also include *Epirinus*. According to Tarasov & Dimitrov (2016), morphological justification for this expanded concept of Sisyphini was based on two unique non-homoplastic synapomorphies and a third homoplastic synapomorphy, i.e., (1) SRP sclerite is characterised by a flat lamella located along the right side of the aedeagal sack and by a small ring structure apically; (2) elytra with last striae (9^{th}, 8^{th}) visible at least pre-apically; and (3) internal surface of basal margin of pronotum with medial carina. Although they also proposed hind wing venation as a diagnostic character, i.e.: RP1 bears a wide posterior sclerite, the same character is

reported for both Onthophagini and Oniticellini (Philips *et al.*, 2004, Tarasov & Dimitrov, 2016).

Available morphological, molecular, chronological and distributional evidence leads us to question the validity of Sisyphini as defined by Tarasov & Dimitrov (2016). Although the sister relationship is clearly demonstrated for *Epirinus* and *Sisyphus* by various studies, we consider the morphological support for their inclusion in the same tribe to be limited. Of the three synapomorphies provided to support Sisyphini, one is ambiguous (Tarasov & Dimitrov, 2016). In addition, the diagnostic wing venation character for the tribe is not unique as it is shared by other tribes such as Onthophagini and Oniticelini (Bai *et al.,* 2011; Philips *et al.,* 2004; Philips, 2016; Tarasov & Dimitrov, 2016). On the other hand, a more restricted definition of Sisyphini (excluding *Epirinus*) is supported by the combination of the following features: antennae with eight articles; pronotal and elytral setose; meta- and meso-legs distinctly long; and external margin of metatibiae weakly serrated. Lastly, the habitus of *Epirinus* is quite different to that of *Sisyphus* (Fig. 5A–D).

Divergence times suggest that *Epirinus* split from *Sisyphus* during the Oligocene. Whilst available geographical data indicate that *Epirinus* is endemic to southern Africa; *Sisyphus* has undergone a wide radiation throughout Africa and beyond, dating from the Miocene. Furthermore, *Epirinus* is centred to the south in more temperate winter rainfall and highland regions, whereas *Sisyphus* is primarily centred in tropical regions.

The systematic, morphological and biogeographical evidence highlights the differences between the two sister clades and justifies their separation into two tribes. Previously, van Lansberge (1874) regarded *Epirinus* as sufficiently distinct to separate it from all other Scarabaeinae within the tribe Epirinides. Here, we propose that it should be reinstated as the tribe Epirinini **stat. rev**.

Taxonomy

Epirinini van Lansberge, 1874 **stat. rev.**

Type genus: *Epirinus* Dejean 1833, designated by van Lansberge, 1874: 189

Diagnosis: The tribe has the following diagnostic characters: (1) small to moderate body size and oval, weakly convex body shape; (2) the anterior margin of the clypeus is bidentate; (3) the pronotal disc punctuate or granular and setose in some species, the median longitudinal line visible or obsolete or fovea present (4) elytral striae weakly or deeply impressed; (5) the sublateral margin of each elytron has an acute, and occasionally right-angled pseudoepipleural carina with a narrow or wide lateral pseudoepipleuron obscuring the lateral margins of the abdomen; (6) the pygidium punctate, setose, granulate, occasionally tuberculate or costate.

Remark: The tribe is monogeneric. It should be noted that Epirinides was treated as an unavailable family-group name by Smith (2006). However, based on article 11.7.2 of the ICZN, Epirinides is available (Bouchard *et al.*, 2011) as the original French vernacular name was subsequently Latinised to Epirini by Bertkau (1875).

Current geographical patterns and possible historical drivers

The present range of dates for divergence between *Epirinus* and *Sisyphus* fall within the Middle Eocene to Upper Oligocene age (40 to 25 Ma) as estimated by Mlambo *et al.* (2011) from substitution rates. The current distribution patterns of *Epirinus* and *Sisyphus* in southern Africa may have been driven by three principal late Miocene to Pliocene geological or climatic trends. These factors have been reported as drivers of species diversification within Scarabaeinae in this region (Davis, 1993; 1997; Davis *et al.*, 2008; Davis *et al.*, 2016a). Climatic trends comprise the inception of winter rainfall in southwest coastal regions and aridification in the southwest interior (Deacon, 1983). Geological trends comprise an uplift in the southeast and erosion into the central Kalahari Basin (Haddon & McCarthy 2005; Dauteuil *et al.*, 2015). These events resulted in a southwest to northeast climatic trend modified by uplift and erosion as in the regions shown in (Figs 3D, 5), which were defined by Davis (1997) based on the climatic classification of Walter & Lieth (1964). These regions are: (1) winter and bimodal spring/autumn rainfall in the southwest; (2) arid climate of the southwest interior with high altitude Upper Karoo in the south; (3) Highveld and Drakensberg grassland in the southeast; mostly xeric savannah on deep sands of

the southern Kalahari Basin; and (4) mostly mesic savannah in the north and northeast (Davis, 1997).

Epirinus is distributed throughout southern Africa primarily in winter and bimodal rainfall climates of the southwest as well as southwest or northeast forests and highlands. Our S-DIVA analysis (Fig. 5) suggests that patterns of radiation may have been primarily from southwest to northeast in *Epirinus*. However, dispersal of species may have been limited by increasing temperatures to the northeast and by the arid barrier provided by the lower Orange River valley in the west during the Mio-Pliocene (Davis *et al*., 2008; Davis *et al.*, 2016a). This suggests that the ancestral of *Epirinus* was cooler pre-adapted.

Sisyphini, on the other hand, have radiated widely in Africa and beyond. Within southern Africa, radiations are shown by *Sisyphus* species, primarily in the warmer east and northeast (Figs 3E, 5). These radiations contrast with those of *Epirinus* to the southwest and southeast of southern Africa (Figs 3E, 5). However, there is some overlap on the Highveld plus southeast coast and in eastern forests with one sisyphine species penetrating to the southwest coast (*S.* (*N.*) *quadricollis*).

Although sisyphines are widely distributed in tropical Africa, there is currently no evidence to support the direction of radiation. However, our biogeographical analysis strongly suggests that dispersal of most southern African species is centred in warmer northeast and eastern regions. Furthermore, the phylogeny suggests that species found in cooler northern and southeast highland areas (*S. (P.) caffer; S. (P.) costatus S. (P.) muricatus*, *S. (N.) barbarossa*, and *S. (N.) kuehni;*) have diversified fairly recently. These species have, presumably, been driven by uplift of the Drakensberg and Highveld in the late Mio-Pliocene (King, 1944; Moore & Blenkinsop, 2006; Dauteuil *et al*., 2015) since each species diverged from an east savannah or dry area sister species within the last ~6 to 2 Ma. On the other hand, dispersal of these species to the southern temperate region of South Africa may have been assisted by climatic fluctuations, from warmer, wetter to relatively cooler, drier conditions in the Plio-Pleistocene (Demenocal 1995, 2004).

Acknowledgements

This study was supported by an NRF (National Research Foundation of South Africa) grant to CLS and CHS. GMD thanks the NRF for a Doctoral Innovation Scholarship (process number: 109628), and the University of Pretoria for the postgraduate financial support. GMD also thanks the Stanley W. Watson Foundation Education (Falmouth, MA, USA) who awarded a trip grant to attend the Workshop in Molecular Evolution-2017, at the Marine Biological Laboratory, University of Chicago, Woods Hole, MA, USA. We would like to express our gratitude to the curators who provided material for this study: Ruth Müller (TMSA); Riaan Stals (SANC) and Simon van Noort (SAM). Christian Deschodt is thanked for providing one of the habitus photographs. We are grateful to Bruno de Medeiros for providing some specimens suitable for a DNA extraction. The two anonymous reviewers are thanked for thoughtful comments in earlier drafts of this paper.

The authors declare no conflict of interest.

References

Bai, M., McCullough, E., Song, K.Q., Liu, W.G., & Yang, X.K. (2011) Evolutionary constraints in hind wing shape in Chinese dung beetles (Coleoptera: Scarabaeinae). *PloS one*, **6**, e21600.

Barbero, E., Palestrini, C., & Zunino, M. (1991) Filogenesi, biogeografia e sistematica generica dei Sisyphini (Coleoptera: Scarabaeidae). *16° Congresso Nazionale Italiano di Entomologia*, **1**, 837-844.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology*, **10**, e1003537.

Bouchard, P., Bousquet, Y., Davies, A.E., Alonso-Zarazaga, M.A., Lawrence, J.F., Lyal, C.H., Lyal CH, Newton, A.F, Reid, C.A, Schmitt, M., Ślipiński, S.A, Smith, A.B & Smith, A. B. (2011). Family-group names in Coleoptera (Insecta). *ZooKeys*, **88**, 1-972.

Bertkau P. (1875) Bericht über die wissenschaftlichen Leistungen auf dem Gebiete der Entomologie während der Jahre 1873 und 1874. Archiv für Naturgeschichte, **40**, 253-400.

Daniel, G.M., Davis, A.L.V., & Scholtz, C.H. (2016) Three new *Sisyphus* species (Coleoptera: Scarabaeidae: Scarabaeinae) from shaded vegetation in southern Africa. *Zootaxa*, **4147**, 67-74.

Daniel, G.M., Davis, A.L.V, Sole C.L., & Scholtz, C.H. (2018) Taxonomic review of the tribe Sisyphini sensu stricto Mulsant, 1842 (Coleoptera: Scarabaeidae: Scarabaeinae) in southern Africa, including new species descriptions. *Insect Systematics & Evolution,* (online edition), 1-61.

Dauteuil, O., Bessin, P., & Guillocheau, F. (2015) Topographic growth around the Orange River valley, southern Africa: A Cenozoic record of crustal deformation and climatic change. *Geomorphology*, **233**, 5-19.

Davis, A.L.V. (1993) Biogeographical groups in a southern African, winter rainfall, dung beetle assemblage (Coleoptera: Scarabaeidae)–consequences of climatic history and habitat fragmentation. *African Journal of Ecology*, **31**, 306-327.

Davis, A.L.V. (1997) Climatic and biogeographical associations of southern African dung beetles (Coleoptera: Scarabaeidae s. str.). *African Journal of Ecology*, **35**, 10-38.

Davis, A.L.V., Scholtz, C.H., & Chown, S.L. (1999) Species turnover, community boundaries and biogeographical composition of dung beetle assemblages across an altitudinal gradient in South Africa. *Journal of Biogeography*, **26**, 1039-1055.

Davis, A.L.V., Scholtz, C.H., & Deschodt, C. (2008) Multi-scale determinants of dung beetle assemblage structure across abiotic gradients of the Kalahari–Nama Karoo ecotone, South Africa. *Journal of Biogeography*, **35**, 1465-1480.

Davis, A.L.V., Scholtz, C.H., Deschodt, C., & Strümpher, W.P. (2016a) Edaphic and climatic history has driven current dung beetle species pool and assemblage structure across a transition zone in central South Africa. *Biological Journal of the Linnean Society*, **119**, 329-347.

Davis, A.L.V., Scholtz, C.H., & Sole, C.L. (2016b) Biogeographical and co-evolutionary origins of scarabaeine dung beetles: Mesozoic vicariance versus Cenozoic dispersal and dinosaur versus mammal dung. *Biological Journal of the Linnean Society*, **120**, 258-273.

Deacon, H.J. (1983) An introduction to the fynbos region, time scales and palaeoenvironments. *Fynbos palaeoecology: a preliminary synthesis*, 1-20.

Demenocal, P.B. (1995) Plio-pleistocene African climate. *Science*, **270**, 53-59.

Demenocal, P.B. (2004) African climate change and faunal evolution during the Pliocene–Pleistocene. *Earth and Planetary Science Letters*, **220**, 3-24.

Felsenstein, J. (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368-376.

Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783-791.

Haddon I.G. & McCarthy T.S. (2005) The Mesozoic-Cenozoic interior sag basins of Central Africa: the Late-Cretaceous–Cenozoic Kalahari and Okavango basins. *Journal of African Earth Sciences,* **43**, 316-333.

Heath, T.A. (2012) A hierarchical Bayesian model for calibrating estimates of species divergence times. *Systematic Biology*, **61**, 793-809.

Ho, S. Y. (2007) Calibrating molecular estimates of substitution rates and divergence times in birds. *Journal of Avian Biology*, **38**, 409-414.

Kainer, D., & Lanfear, R. (2015) The effects of partitioning on phylogenetic inference. *Molecular Biology and Evolution*, **32**, 1611-1627.

King, L.C. (1944) Geomorphology of the Natal Drakensberg. *South African Journal of Geology*, **47**, 255-282.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., & Calcott, B. (2016) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, **34**, 772-773.

Luzzatto, M. (1994) Tattiche comportamentalli nella reproduzione degli Scarabaidi degradatori (Coleoptera: Scarabaeoidea). Un approccio evolutivo e filogenetico, Tesi de Dottorato. *Università degli Studi di Pavia/Università degli Studi di Torino: Torino, Italy*.

Medina, C.A., & Scholtz, C.H. (2005) Systematics of the southern African genus *Epirinus* Reiche (Coleoptera: Scarabaeinae: Canthonini): descriptions of new species and phylogeny. *Insect Systematics & Evolution*, **36**, 145-160.

Mlambo, S., Sole, C.L., & Scholtz, C.H. (2011) Phylogeny of the African ball-rolling dung beetle genus *Epirinus* Reiche (Coleoptera: Scarabaeidae: Scarabaeinae). *Invertebrate Systematics*, **25**, 197-207.

Mlambo, S., Sole, C.L., & Scholtz, C.H. (2015). A molecular phylogeny of the African Scarabaeinae (Coleoptera: Scarabaeidae). *Arthropod Systematics and Phylogeny*, **72,** 303-321.

Monaghan, M.T., Inward, D.J., Hunt, T., & Vogler, A.P. (2007). A molecular phylogenetic analysis of the Scarabaeinae (dung beetles). *Molecular Phylogenetics and Evolution*, **45**, 674-692.

Montreuil O. (2015a) Le genre *Neosisyphus* Muller em Afrique (Coleoptera, Scarabaeidae, Sisyphini). *Catharsius La Revue,* **12,** 1-36.

Montreuil O. (2015b) Premiers cas de brachypterisme dans le genre *Sisyphus* Latreille (Coleoptera, Scarabaeidae, Sisyphini). *Annales de la Société Entomologique de France,* **51**, 281-293.

Morrone, J.J. (2014) Biogeographical regionalisation of the Neotropical region. *Zootaxa*, **3782**, 1-110.

Moore, A. & Blenkinsop, T. (2006) Scarp retreat versus pinned drainage divide in the formation of the Drakensberg escarpment, southern Africa. *South African Journal of Geology,* **109**, 599-610.

Müller, G. (1941). Nuovi coleotteri dell'Africa orientale. *Atti del Museo Civico di Storia Naturale di Trieste*, **14**, 319-352.

Mulsant M.E. (1842) Histoire naturelle des coléoptères de France, Paris. *Maison, Labraire*, **29**, 1-623.

Philips, T.K. (2016) Phylogeny of the Oniticellini and Onthophagini dung beetles (Scarabaeidae, Scarabaeinae) from morphological evidence. *ZooKeys*, **579**, 9-57.

Philips, T.K., Pretorius, E., & Scholtz, C.H. (2004) A phylogenetic analysis of dung beetles (Scarabaeinae: Scarabaeidae): unrolling an evolutionary history. *Invertebrate Systematics*, **18**, 53-88.

Rambaut, A. (2009) FigTree. 1.1. 2008; 19 Available: http://tree. bio. ed. ac. uk/software/figtree.

Rambaut, A., Suchard, M., Xie, W., & Drummond, A. (2014) Tracer v. 1.6. Institute of Evolutionary Biology, University of Edinburgh.

Ronquist, F., & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572-1574.

Schenk, J.J. (2016) Consequences of secondary calibrations on divergence time estimates. *PLoS One*, **11**, e0148228.

Shaul, S., & Graur, D. (2002) Playing chicken (*Gallusgallus*): methodological inconsistencies of molecular divergence date estimates due to secondary calibration points. *Gene*, **300**, 59-61.

Smith, A.B. (2006). A review of the family-group names for the superfamily Scarabaeoidea (Coleoptera) with corrections to nomenclature and a current classification. *Coleopterists Society*, **5**, 144-204.

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

Tarasov, S., & Dimitrov, D. (2016) Multigene phylogenetic analysis redefines dung beetles relationships and classification (Coleoptera: Scarabaeidae: Scarabaeinae). *BMC Evolutionary Biology*, **16**, 257.

Tarasov, S., & Génier, F. (2015) Innovative Bayesian and parsimony phylogeny of dung beetles (Coleoptera, Scarabaeidae, Scarabaeinae) enhanced by ontology-based partitioning of morphological characters. *PLoS One*, **10**, e0116671.

Tarasov, S., Vaz-de-Mello, F.Z., Krell, F.T., & Dimitrov, D. (2016) A review and phylogeny of Scarabaeine dung beetle fossils (Coleoptera: Scarabaeidae: Scarabaeinae), with the description of two *Canthochilum* species from Dominican amber. *PeerJ*, **4**, e1988.

van Lansberge, J.W. (1874) Observations sur la classification des Lamellicornes coprophages: *Annales de la Société Entomologique de Belgique*, **17**, 177-193.

Villalba, S., Lobo, J.M., Martín-Piera, F., & Zardoya, R. (2002) Phylogenetic relationships of Iberian dung beetles (Coleoptera: Scarabaeinae): insights on the evolution of nesting behavior. *Journal of Molecular Evolution*, **55**, 116-126.

Walter, H. & Lieth, H.(1964) Klimadiagramm Weltatlas. *Gustav Fischer, Jena*.

Weir, J.T., & Schluter, D. (2008) Calibrating the avian molecular clock. *Molecular Ecology*, **17**, 2321-2328.

Yu, Y., Harris, A.J., Blair, C., & He, X. (2015) RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution*, **87**, 46-49.

Yu, Y., Harris, A. J., & He, X. (2010) S-DIVA (Statistical Dispersal-Vicariance Analysis): a tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution*, **56**, 848-850.

Zunino, M. (1983) Essai préliminaire sur l'évolution des armures génitales des Scarabaeinae, par rapport à la taxonomie du groupe et à l'évolution du comportement de nidification. *Bulletin de la Societé Entomologique de France*, **88,** 531-543