Molecular Phylogenetics and Evolution (2004), submitted

Chapter 3

Phylogenetic Patterns in Multiple Data Sets used for Inferring Relationships among Genera of ball-rolling Scarabaeini (Coleoptera: Scarabaeidae).

Shaun A. Forgie[A], Paulette Bloomer[B] and Clarke H. Scholtz[A]

[A] Department of Zoology and Entomology, [B] Department of Genetics,
School of Biological Sciences, University of Pretoria, Pretoria 0002, Republic of South Africa

Running title: Phylogeny of the Scarabaeini

Key words: Combined Analysis, Dung beetle, Evolution, CO1, 16S, Morphology, Phylogeny

Abstract

The Scarabaeini is an old world tribe of ball-rolling dung beetles that have origins dating back to at least the mid-upper Miocene (8-18 Million years ago). The oldest classified and most revered of these beetles is the sacred scarab *Scarabaeus sacer* Linneaus 1758 once worshiped by ancient Egyptian society in the form of the solar deity, *Khepera*, who controlled the Sun's daily path across the sky. Yet, despite the notoriety of its members in societies old and new, the tribe has received little to no attention in morphological or molecular phylogenetics. We obtained sequence data from the mitochondrial Cytochrome Oxidase subunit I (1197 bp) and 16S ribosomal RNA (461bp) genes for 25 species of the Scarabaeini in an attempt to further resolve broad phylogenetic relationships inferred from morphology-based hypotheses of the tribe's evolution. Sequence data from both markers along with 216 morphological and 3 biological characters were analysed separately then combined and analysed simultaneously. Results show poorly resolved trees with many of the intermediate and basal nodes forming the backbone of each topology collapsed by low bootstrap values. In concordance with many insect studies involving mitochondrial DNA, many sites in both genes exhibited strong A+T nucleotide bias and high interlineage divergences with transition: transversion ratios reaching saturation. Morphological characters therefore appeared to carry more weight than the molecular data in combined analyses thus increasing their influence on tree reconstructions. Despite extensive topological incongruence, phylogenetic signal was present, however, in a number of well-supported relationships that were congruent between the molecular and morphological data. We investigated conflict and congruence in the data to evaluate if the combined analysis can be considered the most accurate estimate of the tribe's phylogeny.

Introduction

The Scarabaeini comprise a behaviourally advanced guild of Old World beetles within the Scarabaeinae best known for rolling balls of dung. The tribe includes approximately 146 species belonging to the genera *Scarabaeus* L., and *Pachylomerus* Bertoloni, *Kheper* Janssens, *Sceliages* Westwood and the *Scarabaeus* subgenera *Scarabaeus* S. Str., *Scarabaeolus* Balthasar and *Pachysoma* M'Leay. The tribe's origins are thought to date back to the Cenozoic stemming from ancestral lineages that may have appeared in the lower Jurassic ca.180-200 Mya (Million years ago) (Crowson, 1981; Cambefort, 1991a; Scholtz and Chown, 1995). Diversification of these scarabaeoids was thought to coincide with the radiation of both angiosperms (Eocene: ca.50 Mya) and, particularly artiodactyls (lower Oligocene: ca.35 Mya), with a shift from saprophagy and mycetophagy to coprophagy by adults and larvae (Cambefort, 1991b; Scholtz and Chown, 1995. But see Chin and Gill, 1996). While the majority of the Scarabaeini consequently evolved as dung specialists, many of its members also became opportunists in exploiting many types of dung or carrion and some even becoming obligate necrophages. Moreover, the Scarabaeini contain species that are non-rollers (see Halffter and Halffter, 1989) and others that don't roll food backwards but push, drag and carry it forwards.

To date, only morphological character sets have been used in phylogenetic studies to infer inter- and/or intra-generic relationships among members of the Scarabaeini (Mostert and Scholtz, 1986; Barbero *et al.*,1998; Harrison and Philips, 2003). These studies were based on relatively small amounts of data that may have generated inaccurate or biased phylogenetic reconstructions (see Hillis, 1998; Grandcolas *et al.*, 2001). Recent studies of the Scarabaeini (Forgie *et al.*, in press) and the Scarabaeinae (Pretorius *et al.*, 2001; Philips *et al.*, 2004.) were based on large morphological data sets comprising more than 200 characters in an attempt to improve phylogenetic signal and generate more robust hypotheses. Both studies support congruence in the polyphyletic evolution of ball-rolling and feeding behaviours deviating from coprophagy.

However, a high degree of character homoplasy was reported among the scarabaeines (Pretorius *et al.*, 2001; Philips *et al.*, 2004), likely the product of convergent evolution brought about by similar environmental influences (Hillis, 1987).

The advent of Polymerase Chain Reaction (PCR; Saiki *et al.*, 1988) marked a proliferation in the use of sequenced regions within mitochondrial DNA (see, Simon *et al.*, 1994), and more recently, nuclear ribosomal DNA in insect molecular systematics (see, Caterino, Cho and Sperling, 2000). Within the former of these classes, the COI and COII markers have historically proven useful in providing sufficient phylogenetic signal in estimating relationships corresponding to interspecific levels of recent divergence within Coleoptera (e.g. Emerson and Wallis, 1995; Langor and Sperling, 1997; Kobayashi *et al.*, 1998; Cognato and Sperling, 2000) including within the Scarabaeinae (Villalba *et al.*, 2002). In contrast, the highly conserved 3' region of the large ribosomal subunit (16S) of mitochondrial DNA has proven more effective at addressing deep levels of divergence evident among distantly related taxa (DeSalle, 1992; Derr *et al.*, 1992). Similarly, 18S nuclear ribosomal RNA has also been useful for resolving basal relationships in higher level phylogenetic studies (Chalwatzis *et al.*, 1996; Caterino *et al.*, 2002). Given that different genes evolve at different rates and the same gene may have different rates of evolution in different lineages (Lunt *et al.*, 1996), the quest to obtain suitable levels of variability has become increasingly important in attempting to resolve close, intermediate and deep levels of divergence where possible in any phylogenetic study.

Thus, the value of a total evidence approach to utilising multiple data sets and analysing them separately (Bull *et. al.* 1993; Miyamoto and Fitch, 1995), or combined (Kluge, 1998) and analysed simultaneously (Nixon and Carpenter, 1996; Baker and DeSalle, 1997) has become apparent. Indeed, multiple data sets are integral in many phylogenetic studies using molecular markers (Vogler and DeSalle, 1993; Funk *et al.*, 1995; Vogler and Welsh, 1997; Funk, 1999; Mardulyn and Whitfield, 1999; Durando, *et al.*, 2000) and morphology (Lafay *et al.*, 1995;

Whiting *et al.*, 1997; Silvain and Delobel, 1998; Joy and Conn, 2001; Wieblen, 2001; Wiegmann *et al.*, 2002).

An examination of the evolution of flightlessness in the Scarabaeini has been made possible by the inclusion of morphological characters from old and rare brachypterous specimens curated in museums (Harrison and Philips, 2003; Forgie *et al.*, in press). While this current study attempted to obtain amplifiable DNA from these museum specimens for comparative analyses, a major limitation of molecular phylogenetics is realised with the difficulties not only in obtaining uncontaminated DNA of sufficient molecular-weight to amplify but in achieving repeatability (for overview, see Wayne *et al.*, 1999). However, with the development of improved DNA extraction methods and materials, limited success is achievable in the amplification and study of short mitochondrial DNA sequences from museum pinned beetles (Cognato and Sperling, 2000).

For this study, we chose portions of the COI and 16S rRNA mitochondrial genes as likely candidates for simultaneous analyses with and without morphological data to resolve as many of the relationships as possible between the close and more distantly related exemplars of the Scarabaeini. In doing so, we aimed to compare the molecular evolution and phylogenetic utility of these two genes and assess the level of congruence these analyses held with the morphology-based hypotheses of Scarabaeini phylogenetics by Forgie *et al.* (in press). This would then provide for a better understanding of the evolution of ball-rolling, flightlessness, feeding specialisation and relatedness between the Scarabaeini and the morphologically similar eucraniines (see Philips *et al.*, 2002). We then used this combined set of analyses to assess the current classification of the tribe (Forgie *et al.*, in press) in which *Kheper* and *Sceliages* are assigned as subgenera of *Scarabaeus* and *Drepanopodus* lineages synonymised with *Scarabaeus s. str.*

Materials and methods

Taxa

The 27 ingroup taxa used in this study are a majority representation of the exemplars selected by Forgie *et al.* (in press) to represent the most morphologically and behaviourally discordant members in the Scarabaeini (Table 1). Some of these taxa were historically classified into genera (retained in square brackets) that have since become synonyms of *Scarabaeus* (e.g. *Scarabaeus [Neateuchus] proboscideus*). Selectionof the three outgroup species *Heliocopris hamadryas* (Coprini), *Circellium bacchus* (Canthonini), and *Eucranium arachnoides* (Eucraniini) were based on their topological positioning relative to the Scarabaeini inferred in the phylogenetic study of the Scarabaeinae by Philips *et al.* (2004). Moreover, outgroup selection criteria discussed by Nixon and Carpenter (1993) were taken into account. A second eucraniine *Anomiopsoides heteroclytus* was utilised for COI and 16S rRNA sequencing to improve phylogenetic signal of the Eucraniini in an effort to qualify recent tribal analyses based on molecular data supporting convergence as the most likely cause of similar morphology between the eucraniines and scarabaeines (Ocampo, unpubl.).

Morphology

We compared the phylogenetic utility and levels of congruence of the DNA data with a set of morphological characters described in detail by Forgie *et al.* (in press). Two hundred and sixteen adult morphological and three biological characters were utilised for this study. Twenty-eight characters from the original data set directly associated with flight and flightlessness were excluded from this study due to their obvious lack of character independence and potential convergent nature.

DNA preservation and extraction

Apart from the pinned museum specimens listed in Table 1, all other specimens were freshly collected in the field. Larger beetles were either split in half or injected with 96-100% ethanol immediately prior to preservation in order to accelerate the infusion of preservative throughout the body. All material was stored at -20°C and ethanol changed at least twice for each specimen shortly after the initial steps of preservation. One of two methods was used in the extraction of mitochondrial DNA from prothoracic or profemoral muscle of ethanol specimens and tarsi from pinned beetles. Tissue was rinsed in insect Ringer's solution (10xStock pH7.4: 58 mol/g NaCl at 1280mM; 147 mol/g $CaCl_2$ + $2H_2O$ at 15mM; 74.6 mol/g KCl at 50mM), dried, frozen with liquid nitrogen then ground separately in 1.5ml microfuge tubes prior to the extraction of DNA. The first method was based on Chelex DNA preparation protocols of Walsh *et al.* (1991) and Belshaw *et al.* (1999). Approximately 1mg of tissue was added to 100 µl stirred 5% solution of Chelex100® (1-800-4BIORAD, Cat.# 143-2832) and heated to 95-100°C for 15min. Without being removed from the microtubes, homogenised samples were stored at -20°C until required for short term use as templates in polymerase chain reaction (PCR). When required, samples were thawed, vortexed thoroughly and spun down in a centrifuge 14 000 rpm for 3 minutes to separate the Chelex beads from the supernatant. Alternatively, a DNeasy Tissue Kit (Qiagen Inc., Santa Clara, CA) was used for the extraction and long term storage of DNA from pinned and ethanol specimens. The manufacturer's protocol was followed, except for the DNA elution, which consisted of one elution in 150 µl of sterile ddH_2O. For DNA from pinned specimens, this step was increased to two 200 µl elutions of pure water incubated at R/T for 5 min. Both eluates of each specimen were then combined and spun in an evaporator to increase concentration of the overall DNA yield.

Table1. Representative ingroup and outgroup (in bold) taxa used in this study.

Taxa	Collection Locality	DNA Preserv[n] method	CO1 analysis (Genbank accession number)	16S analysis (Genbank accession number)	Morph[y] analysis	Combined analysis
***Circellium bacchus* Fabr.**	E Cape, SA	EtOH	AF499750	AF499690	✓	✓
***Heliocopris hamadryas* (Fabr.)**	NW Prov., SA	EtOH	AF499751	AF499691	✓	✓
***Eucranium arachnoides* Brullè**	Mendoza, Arg.	EtOH	AF499752	AF499692	✓	✓
***Anomiopsoides heteroclytus* (Blanchard)**	La Rioja, Arg.	EtOH	AF499753	AF499693		
S. [Drepanopodus] proximus (Péringuey)	Namib Des., Nam.	EtOH	AF499754	AF499694	✓	✓
S. (Kheper) lamarcki (M'Leay)	Gauteng, SA	EtOH			✓	
S. (Kheper) nigroaeneus (Boheman)	Gauteng, SA	EtOH	AF499755	AF499695	✓	✓
S. (Kheper) subaeneus (Harold)	N Prov., SA	EtOH	AF499756	AF499696	✓	✓
S. [Mnematidium] multidentatus (Klug)	Palestine	Pinned			✓	
S. [Mnematium] ritchiei M'Leay	Tripoli, Libya	Pinned			✓	
S. [Mnematium] silenus Gray	Sanai Pen., Egypt	Pinned			✓	
S. [Neateuchus] proboscideus (Guérin)	Namaqualand, SA	EtOH	AF499757	AF499697	✓	✓
S. [Neopachysoma] denticollis (Péringuey)	Namib Des., Nam.	EtOH			✓	
S. [Neopachysoma] rodriguesi Ferreira	Namib Des., Nam.	EtOH	Sole *et al.* Unpubl. Seq.		✓	
S. (Pachysoma) bennigseni Felsche	Namib Des., Nam.	EtOH	AF499758	AF499698	✓	✓
S. (Pachysoma) hippocrates M'Leay	Namaqualand, SA	EtOH	AF499759	AF499699	✓	✓
Pachylomerus femoralis Kirby	Gauteng, SA	EtOH	AF499760	AF499700	✓	✓
S. (Scarabaeolus) bohemani Harold	Gauteng, SA	EtOH	AF499761	AF499701	✓	✓
S. (Scarabaeolus) flavicornis (Boheman)	NW Prov., SA	EtOH	AF499762	AF499702	✓	✓
S. (Scarabaeolus) rubripennis (Boheman)	Namib Des., Nam.	EtOH	AF499763	AF499703	✓	✓
S. (Scarabaeolus) scholtzi Mostert & Holm	Chalbi Des. Som.	Pinned			✓	
S. galenus (Westwood)	Kruger NP, SA	EtOH	AF499764	AF499704	✓	✓
S. goryi Castelnau	Kruger NP, SA	EtOH	AF499765	AF499705	✓	✓
S. rugosus (Hausman)	W Cape, SA	EtOH	AF499766	AF499706	✓	✓
S. rusticus (Boheman)	NW Prov., SA	EtOH	AF499767	AF499707	✓	✓
S. satyrus (Boheman)	N Cape, SA	EtOH	AF499768	AF499708	✓	✓
S. westwoodi Harold	Sani Pass, Lesotho	EtOH	AF499769	AF499709	✓	✓
S. zambesianus Péringuey	N Prov., SA	EtOH	AF499770	AF499710	✓	✓
S. (Sceliages) adamastor (Serville)	W Cape, SA	EtOH	AF499771	AF499711	✓	✓
S. (Sceliages) brittoni zur Strassen	Namaqualand, SA	EtOH	AF499772	AF499712	✓	✓
S. (Sceliages) hippias Westwood	NW Prov., SA	EtOH	AF499773	AF499713	✓	✓

NOTE. Key to abbreviations: Desert (Des); Province (Prov.); National Park (NP); Argentina (Arg.); Namibia (Nam.); Somalia (Som.); South Africa (SA). In the taxa column, *S.* is an abbreviation for the genus *Scarabaeus.*

PCR amplification and DNA Sequencing

Primer sequences used for amplification of DNA fragments were obtained from Simon *et al.* (1994). Initial COI sequences comprising 1296 nucleotide bases were obtained by amplifying C1-J-1718 (5' GGAGGATTTGGAAATTGATTAGTTCC 3') in conjunction with the tRNA-Leucine (UUR) primer TL2-N-3014 (5' TCCAATGCACTAATCTGCCATATTA 3'). Over-lapping sequences were generated with C1-J-1718 in combination with C1-N-2329 (5' ACTGTAAATATATGATGAGCTCA 3' with the A in position 12 from the 5' end substituted

with a G for improved primer specificity) and C1-J-2183 (5' CAACATTTATT TTGATTTTTTGG 3') with TL2-N-3014 to yield the 1197 bp segments used in the analyses. The Lunt *et al.*, (1996) primer UEA7 (5' TACAGTTGGAATA GACGTTGATAC 3') was tested in conjunction with TL2-N-3014 for shorter length COI sequences of pinned museum specimens. For 16S rRNA, we obtained 450 bp fragments using the 16Sb2 primer (5' TTTAATCCA ACATCGAGG 3') in conjunction with LR-N-13398 (5' CGCCTGTTTAACAAAAACAT 3') after Vogler *et al.* (1993). PCR reactions contained 2.5 mM of each dNTP ,10x Reaction Buffer, 25 mM $MgCl_2$, 25 pmol of each primer and 1.5 units Super-Therm™ *Taq* (Hoffman-la-Roche, Cat# JMR-801) gave approximately 70-110ng of DNA template. PCR cycle conditions for COI were 2 min at 94°C initial denaturing, 30 sec at 94°C, 30 sec at 48°C and 1min at 72°C for 35 cycles and a final extension of 72°C for10 min. For museum specimens the COI cycle conditions were altered by increasing the annealing temperature to 52°C and the number of cycles to 40. Cycle conditions for 16S were 2 min at 94°C initial denaturing, 20 sec at 94°C, 20 sec at 52°C and 1min at 72°C for 35 cycles and a final extension of 72°C for10 min.

PCR products were visualised by electrophoresis through agarose gels (1xTAE), then cut from the gels and the DNA recovered and purified with a High Pure PCR Purification Kit (Roche Diagnostics Corp.; Cat#1732676) following the manufacturer's protocol. Approximate concentration of the purified product was determined by comparing the intensity of the products to known concentration pGEM vector in agarose gels (1xTAE). Purified PCR fragments were sequenced using the same primers at 3.2 pmol, 2 ul BigDye™ terminator reaction mix (PE Applied Biosystems, Foster City, CA.), 200ng template and the BigDye cycle sequencing conditions provided by the GeneAmp™ PCR system 9700 (PE Applied Biosystems), Cycle sequencing products were precipitated following their addition to 1.5 ml centrifuge tubes containing 3M NaOAc (2 µl), absolute EtOH (50 µl) and ddH_2O (10 µl), then pelleted, dried and sequenced using an ABI 377 automated sequencer. Automated DNA sequences for each species were inspected and corrected using Sequence Navigator® software (PE Applied Biosystems)

with largely overlapping reads from both heavy and light strands. All complete sequences were submitted to GenBank (Table 1). We failed to obtain sufficient repeatable fragments of both the COI and 16S genes for the pinned specimens and they were therefore not used in the phylogenetic analyses of molecular and combined data sets. These species, however, were included in the morphological tree (Fig. 1A) of Forgie, Philips and Scholtz (unpubl.) and served to provide us with an estimate of their relatedness to other members of the tribe and their topological placement in the trees generated in this study.

Alignment and Phylogenetic Analysis

Alignment of COI and 16S rRNA sequences was done using the default parameters of Clustal X (Gibson *et al.*, 1994) taking the theoretical considerations of Gatesy *et al.* (1993) into account. Any ambiguous base(s) appearing in the aligned sequences were checked against several specimens showing nucleotide congruency in the same codon position(s) as viewed in Sequence Navigator. Gaps were coded as missing characters.

In order to gain the best estimates of phylogeny, simultaneous analyses of the data sets individually and in combination (i.e. Morphology + CO1, Morph. + 16S, CO1 + 16S, Morph. + CO1 + 16S) were carried out (Nixon and Carpenter, 1996). All analyses were performed using PAUP* 4.0b10 (Swofford, 1999). Both the eucraniines, *Eucranium arachnoides* and *Anomiopsoides heteroclytus*, were included in theoutgroup to ascertain their relationship to *Scarabaeus* subgenus *Pachysoma sensu lato* (S. L. : incl. *Neopachysoma* [Syn.]) lineages. Unweighted parsimony analyses were based on heuristic and branch-and-bound search options with 100 random additions of sequences and tree-bisection-reconnection (TBR) swapping unless otherwise stated. Neighbour-Joining analyses (NJ; Saitou and Nei, 1987) of the molecular data sets used a randomised input order for the taxa (see Farris, 1995) and employed distances corrected with HKY85 (Hasegawa *et al.*, 1995) which adjusts for variance in transition:

transversion ratios and unequal nucleotide frequencies. COI and 16S rRNA gene fragments operate under discordant molecular constraints (Matthee and Robinson, 1997), evolve at different rates (Lunt *et al.*, 1996) and may accumulate some degree of saturation of substitutions among the nucleotide sites (Mardulyn and Whitfield, 1999). We therefore conducted the following analyses in an attempt to improve the overall phylogenetic signal of the sequence data.

Weighted parsimony analyses

Character reweighting of conserved first and second codon changes against variable third-position changes in the COI sequences was carried out on the basis of the frequency of change for each position within a codon for all taxa using PAUP* (characters were weighted according to the inverse of their variability). Substitution weighting of transversions (TV) relative to transitions (TI) for both COI and 16S rRNA was performed based on the frequencies of the two substitution types calculated using the 'State Changes and Stasis' option counting 100 "equiprobable" random parsimony trees with MacClade's version 3.08 chart menu (Maddison and Maddison, 1992). Since this is likely to underestimate the ratios between more closely related taxa, less conserved ratios were also examined in maximum-likelihood and parsimony analysis. Saturation/homoplasy of molecular data usually occurs as sequence divergence among taxa increases. Its prevalence in the sequence data is measurable by comparing consistency and retention indices (CI and RI respectively) resulting from unweighted parsimony analyses with and without the removal of transitions (see Matthee and Davis, 2001). Transitions were also used to assess levels of saturation/homoplasy among weighted analyses since their removal may not fundamentally increase general levels of resolution or congruence (Vidal and Lecointre, 1998; Broughton *et al.*, 2000). We opted for this method rather than using regression analysis to determine the slope of saturation plots to estimate rates of nucleotide variance in genes (Simon *et al.*, 1996).

Maximum likelihood analyses

Maximum likelihood analyses (ML, Felsenstein, 1981), were conducted on both molecular data sets using the heuristic search option with the as-is addition sequence. We employed the HKY85 + Invariant (I) + Gamma (G) distribution model for the ML analyses which takes into account unequal substitution rates, site rate heterogeneity allowing for a proportion of invariant sites, simultaneously estimating these three variables from the sequence data when the tree(s) are calculated (Gu *et al.*, 1995). A maximum of 100 bootstrap replications were performed due to computational time constraints.

Combined data parsimony analysis

A partition homogeneity test with 1000 iterations and no branch swapping was performed on the combined data to identify possible conflicting phylogenetic signals (Liu and Miyamoto, 1999). Weighting of characters is possible in combined data sets but is prone to subjectivity (Hillis, 1987). We have therefore chosen not to apply weighting in the combined data.

Statistical support

Statistical support for each clade node in all analyses was estimated by bootstrapping (Felsenstein, 1985) with heuristic search, with 1000 iterations, as-is addition sequence and TBR branch swapping. It is worth noting that bootstrapping values should not be strictly interpreted as confidence limits on monophyly (Kluge and Wolf, 1993; Lee, 2000). Rather, they should provide an indication only for the degree of support of a particular clade node, as the values are usually very conservative estimates of the probability that a particular clade is a true historical group (Hillis and Bull, 1993; See also argument on tree robustness and clade significance by Lee, 2000).

Results and discussion

Pinned specimens

All pinned specimens (Table 1) yielded small amounts of DNA which gave limited success being amplified with CO1 primer pairs UEA7 and TL2-N-3014, in addition to C1-J- 1718 and C1-N-2191. Only small portions of sequence were obtained from *Scarabaeus (Scarabaeolus) scholtzi*, and *S. [Mnenatium] ritchiei*. A GenBank BLAST search on the sequences we generated confirmed we had amplified beetle DNA and not contaminant DNA. These sequence fragments possessed varying degrees of background noise and with minimal to no overlapping of complementary sequences, problematic bases were unresolved. Moreover, we were unable to obtain repeatability and with a lack of material to sample decided not to continue. Two suggestions as to why we were unsuccessful came to light, (a) the fragments we were trying to amplify were too long, and (b) the DNA was too old and/or degraded. Similar-sized fragments (~ 400-550 bp) of COI were sequenced successfully from recently pinned specimens (<30 years old) of the genus *Ips* (Coleoptera: Scolytidae) by Cognato and Sperling (2000). By contrast, the rare pinned material used in our study lacked any reliable collection data or dates. Moreover, the method of preservation is unknown and therefore unreliable (Prost *et al.*, 1993). The degradation of the DNA of old pinned specimens, are invariably the product of bacterial activity and the oxidative processes associated with time (Pääbo, 1989). Nonetheless, dry tissues up to 14, 000 years old has been shown to yield short mitochondrial DNA fragments (~100-200 bp) of sufficient concentration for sequencing (e.g. Pääbo, 1989; Roy *et al.*, 1994).

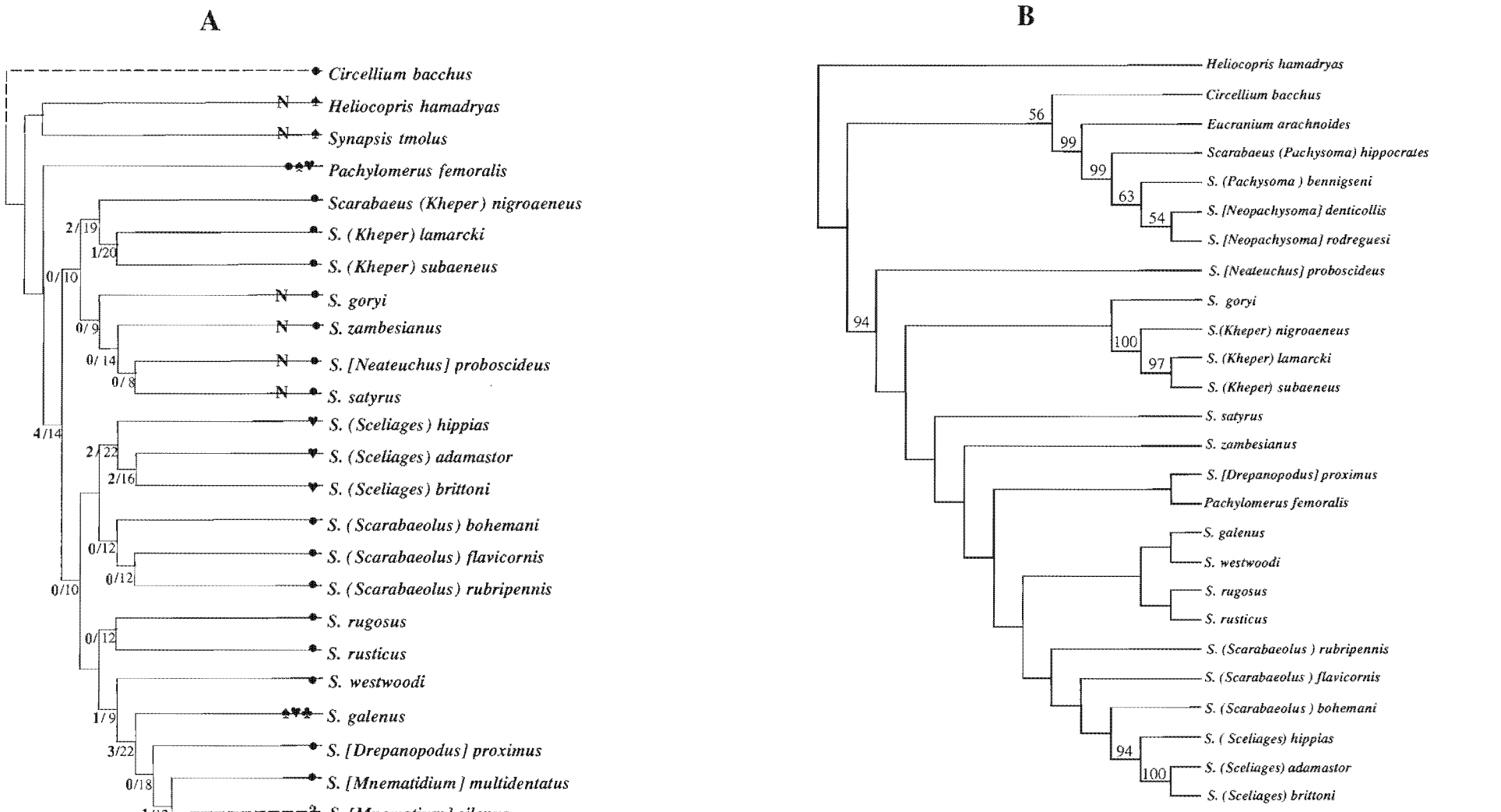

Fig. 1. Morphological Trees based on 216 morphological and 3 biological characters of the Scarabaeini (Coleoptera Scarabaeidae). (**A**) Single PIWE (Goloboff 1997) tree (CI=0.24, RI=0.49, Length =1673) from Forgie *et al.* (in press) following less strong weighting against homoplasious characters with concavity index = 4. Unique synapomorphies (Bold)/homoplasious synapomorphies provided below nodes. Flightless taxa (dashed branches). Nocturnal taxa (N). Modes of food relocation: 'rolling" ●; Tunnelling, ♠; Pushing, ♥; Dragging, ♦; Carrying, ♣; Unknown, ?. (**B**) Single MP tree recovered by heuristic search (CI=0.29, RI=0.51, Length =1336). All characters are unweighted. Bootstrap values above 50% support are given at nodes.

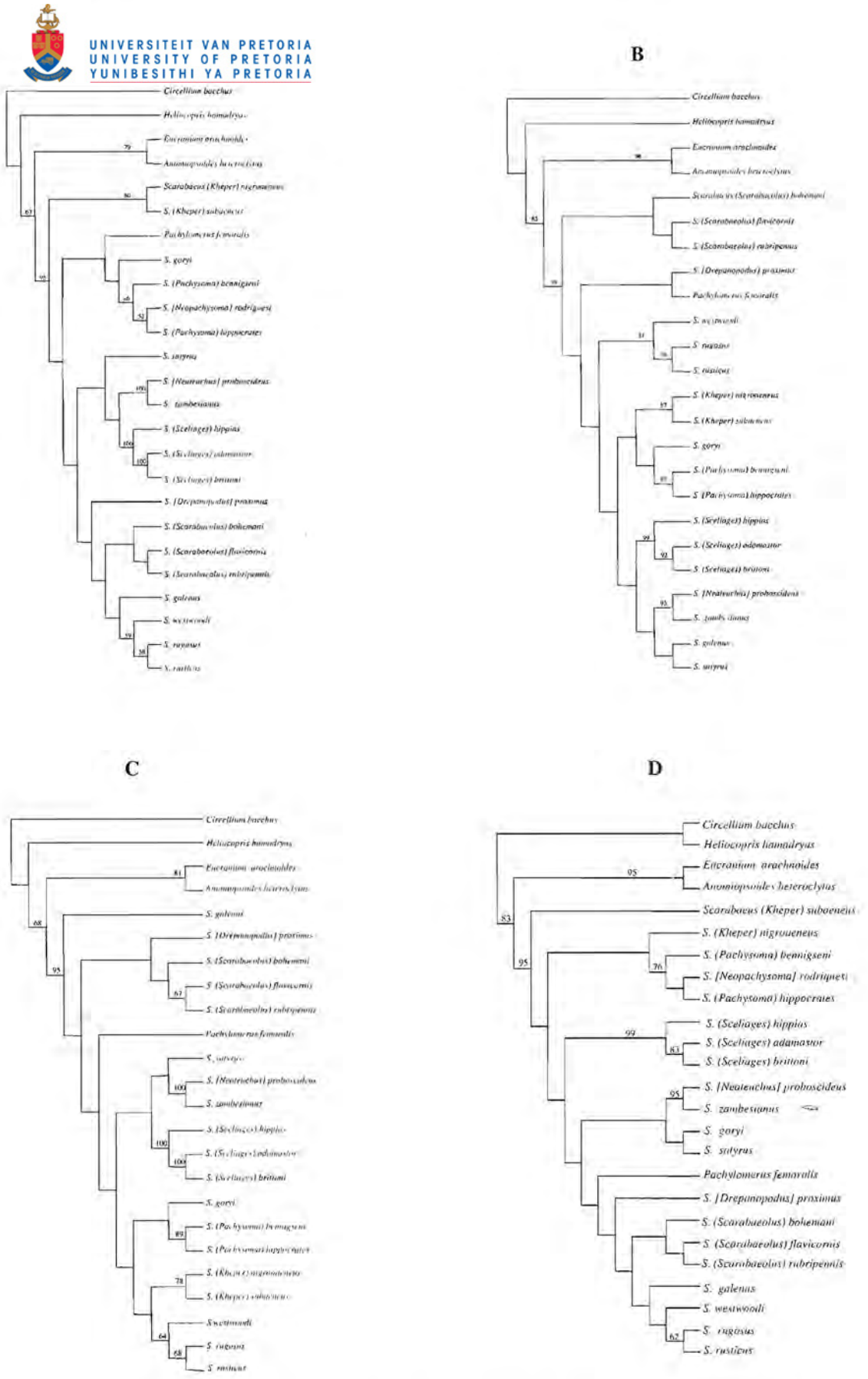

Fig. 2. Analysis of the Scarabaeini based on 1198 nucleotide sequences of the CO1gene of mitochondrial DNA. Single MP tree (A) recovered from heuristic search of unordered, unweighted parsimoniously informative characters. ML analysis with, branch swapping, TBR and 1000 iterations; recovery of single heuristic tree (B). Weighted MP analyses based on heuristic searches following reweighting codon positions 1-3 with a 3:10: 1 ratio respectively (C), and TI:TV = 1.3 weighting scheme (D). Bootstrap values above 50% support are given at nodes on all trees.

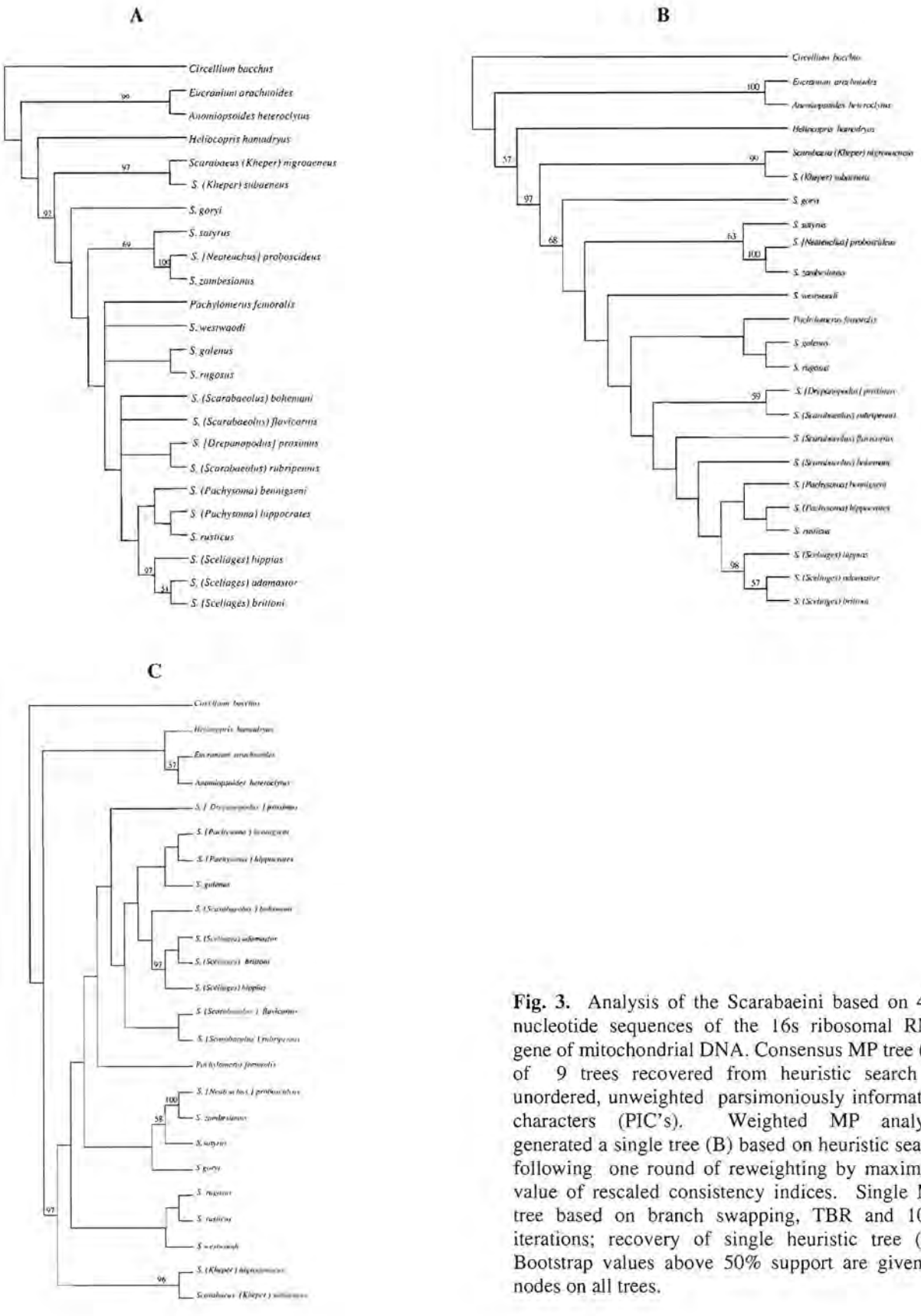

Fig. 3. Analysis of the Scarabaeini based on 461 nucleotide sequences of the 16s ribosomal RNA gene of mitochondrial DNA. Consensus MP tree (A) of 9 trees recovered from heuristic search of unordered, unweighted parsimoniously informative characters (PIC's). Weighted MP analysis generated a single tree (B) based on heuristic search following one round of reweighting by maximum value of rescaled consistency indices. Single ML tree based on branch swapping, TBR and 1000 iterations; recovery of single heuristic tree (C). Bootstrap values above 50% support are given at nodes on all trees.

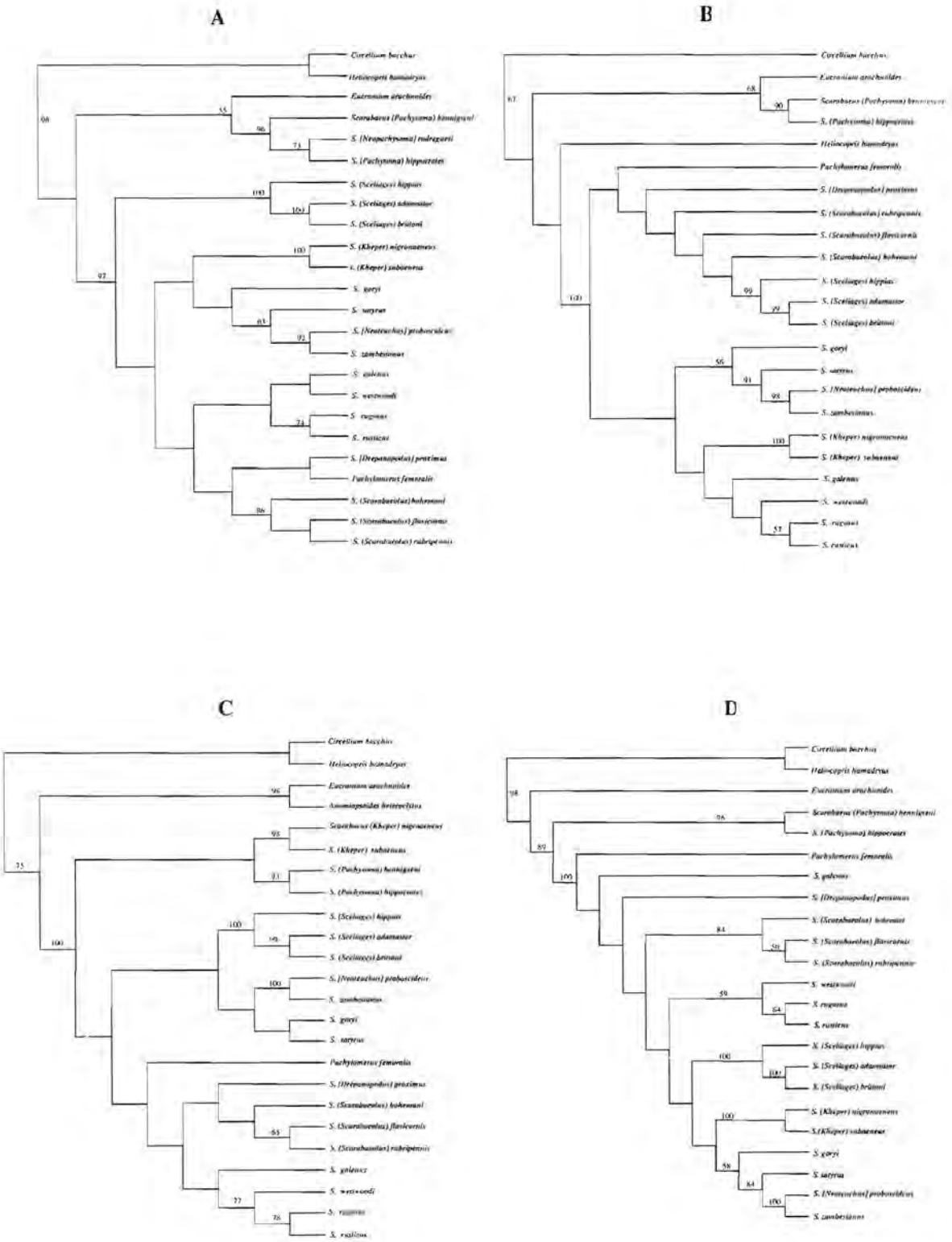

Fig. 4. Combined MP analyses of the Scarabaeini (Coleoptera: Scarabaeidae). Morphological + CO1 mtDNA (A), Morphological + 16S rRNA mtDNA (B), CO1+ 16S rRNA mtDNA (C), and all three data sets combined (D) recovered single tree topologies following Heuristic searches in each analysis. Bias in trees A, B and D towards a morphological topology (Fig. 1A) may be due to more weight being applied to morphological characters than molecular characters or more morphological than molecular characters supporting nodes that are topologically similar to those of the morphological trees. Bootstrap values above 50% support are given at nodes on all trees.

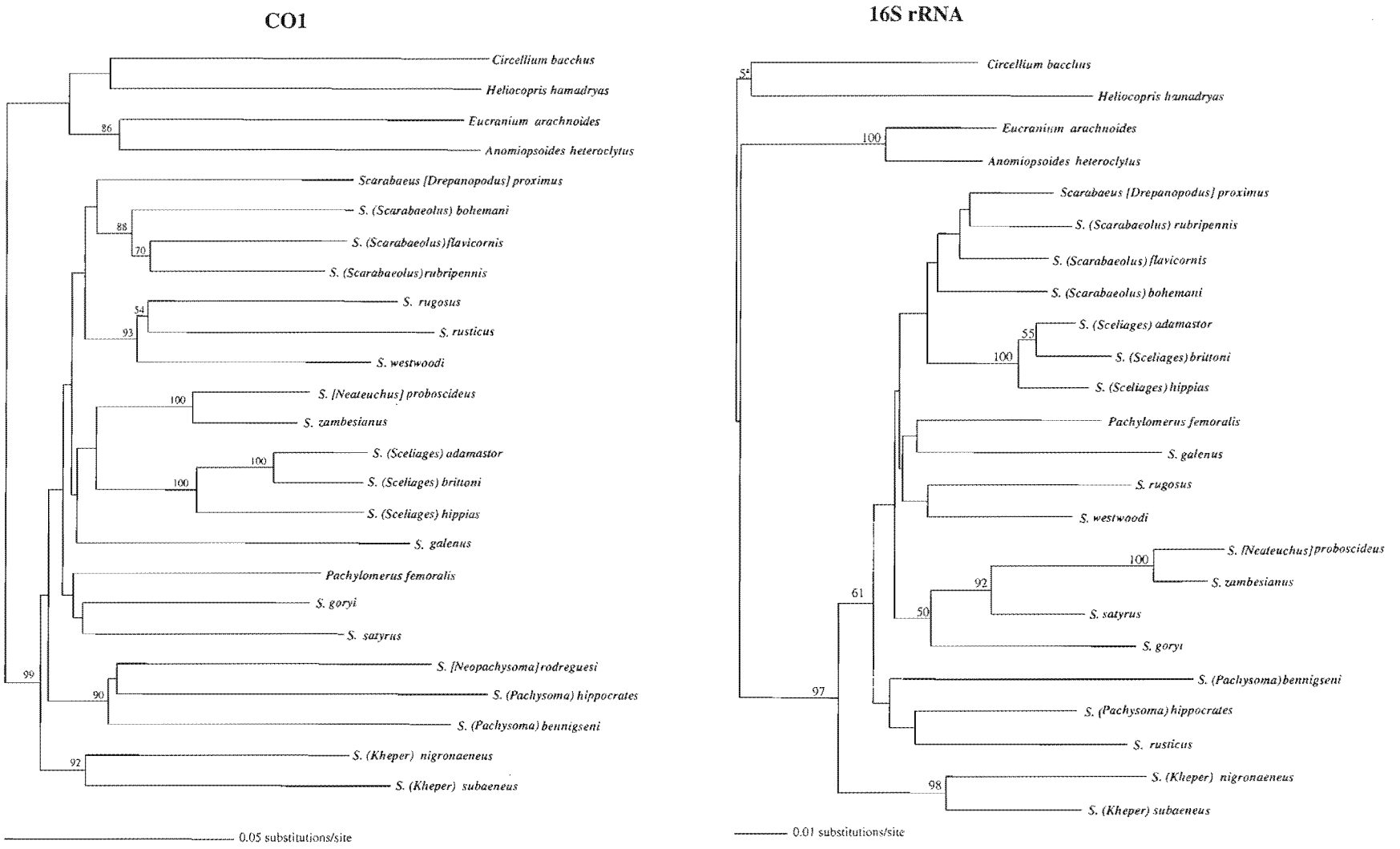

Fig. 5. Phylograms of Scarabaeini COI and 16S rRNA mitochomdrial genes based on Neighbour Joining analysis. Bootstrap values above 50% support are given at nodes. Arrows indicates node from which rapid divergence of the majority of lineages occurs. CO1 NJ trees shows predictably longer branches with greater accumulation of character changes over time than the 16S sequence data for the same taxa.

Morphological Analysis and Characteristics

Unweighted parsimony analysis of 219 characters in the morphological data set with 3 uninformative characters removed, generated a single tree 1336 steps in length (Fig. 1B; CI= 0.29, RI= 0.51). The majority of the relationships were unresolved. An over abundance of character homoplasy is evident in the morphological data (Fig. 1A). Even under the most stringent weighting (concavity indices 1 and 2) against homoplasious characters using PIWE (Parsimony Implied Weights; Goloboff, 1997), complete resolution of the resulting topologies (results not shown) was not achieved by Forgie, Philips and Scholtz (unpubl.).

COI mt DNA Analysis and Characteristics

A 1197 bp region of the mtDNA COI gene from 23 ingroup taxa and 2 outgroup taxa (*C. bacchus* and *H. hamadryas*) contained 480 variable characters (Appendix 1a) including 378 that were phylogenetically informative. The majority of phylogenetic information occurred at the 3rd codon position accounting for the vast majority (71%) of the variability followed by 1st and highly conserved 2nd codon positions (22% and 7% respectively), as seen for example in recent insect studies using COI (e.g. Emerson and Wallis, 1995; Langor and Sperling, 1997; Funk, 1999; Mardulyn and Whitfield, 1999; Cognato and Sperling, 2000, Villalba *et al.*, 2002). Mean base composition across all lineages showed an excess of A (31.7%) and T (39.6%) over C (14.8%) and G (13.9%) in our COI data corresponding to a similar A+T bias occurring at the 3rd codon position recorded in several of the studies cited above including Liu and Beckenbach, (1992) and Juan *et al.* (1995).

Unweighted parsimony analysis generated a single tree (CI = 0.30, RI = 0.34) with a length of 1917 steps (Fig. 2A). Reweighting 1st, 2nd and 3rd codon positions according to site variability also resulted in a single tree (CI = 0.34, RI = 0.37, length = 2904) sharing several topological

congruencies (Fig. 2C). A transversion weighting scheme (TI/TV= 1.3) generated a single tree (Fig 2D; CI = 0.34, RI= 0.35, length =2244.8). Maximum Likelihood analysis with an empirical estimated proportion of invariant sites of 0.49 is shown in Fig. 2 B (Log L score = -9567.485, gamma = 0.628). Although COI sequences yielded many parsimoniously informative characters (Appendix 2), they appear to be plagued by homoplasy due to saturation of the 3rd codon positions (but see Funk, 1999). As a result, approximately half of the tribal relationships depicted in figure 2 were unresolved at a variety of hierarchical levels. Interestingly, COI provided strong support for the deepest nodes that clearly differentiate the Scarabaeini from the outgroup lineages including the two eucraniines (*A. heteroclytus* and *E. arachnoides*).

16S rRNA mt DNA Analysis and Characteristics

Maximum Parsimony and Likelihood analyses were carried out using 461 bp sequences (including alignment gaps) of the 16S ribosomal RNA gene of mtDNA obtained from 22 ingroup taxa and 2 outgroup taxa. One hundred and forty-six characters were variable (Appendix 1b) with 107 of these being parsimony informative. While our analysis did not ascertain whether the majority of character variability occurred in stems or loops, studies by Funk *et al.*, (1995), Matthee and Robinson (1997) and Funk (1999), among others, report general consensus in site variability, A+T bias and indeed saturation being more prevalent in loops than in stems. Our data showed predictably high A+T richness (76%) compared to C (15%) and G (9%) nucleotide frequencies in the variable informative characters. The strict consensus tree (CI = 0.43, RI = 0.50) of nine most parsimonious trees recovered following an unweighted heuristic search is shown in Fig. 3A. This consensus tree contains two sets of tetratomies resulting from the collapse of several poorly supported nodes located medially in the topology. One round of consistency index-based reweighting of the parsimony informative sites (based on the individual character CI's in the nine trees) assigned 15 characters with a weight of 1 and 92 characters with weights less than 1. A single tree (Fig. 3B; CI = 0.52, RI = 0.57, Length = 157) was recovered

some 211 steps shorter than the consensus tree. A Maximum Likelihood analysis based on an empirically estimated proportion of invariant sites at 0.6 is shown in Fig. 3C (Log L score = -2574.429, gamma = 1.628).

Distance Analysis

The presence of many short branches stemming from the more basal nodes in the trees of both COI and 16S NJ trees (Fig. 6) suggests a short time period of rapid radiation of a majority of the ingroup taxa. Virtually all lineages in both COI and 16S NJ trees exhibit long branches following on from the short burst of rapid speciation. In the COI tree however, these branches are predictably longer with an accumulation of more character changes over time than in the 16S tree given the faster evolution rate in COI. Despite the difficulties in achieving good resolution of [ancient] rapid radiations, bootstrap support was given to more relationships recovered by the neighbour joining method in both molecular data sets thereby providing perhaps the most robust hypotheses of tribal evolution.

Sequence Divergence

Pairwise sequence divergences exhibited up to 19.1% divergence recorded between flight capable *S. rugosus* and flightless *S. (Pachysoma) hippocrates* within the Scarabaeini and up to 22.7% between ingroup and outgroup lineages in the COI data (Table 2). These values fall within the range of divergences reported in several insect COI studies (reviewed by Funk, 1999; Mardulyn and Whitfield, 1999, Cognato and Sperling, 2000). Maximum sequence divergences for 16S within ingroup taxa were up to 14.4% recorded between *S. rusticus* and *S. (Kheper) nigroaeneus*, and similarly up to 14.8% scarabaeine divergence from outgroup taxa (Table 3).

TABLE 2. CO1 mtDNA uncorrected sequence divergence values (%) between members of the Scarabaeini (Coleoptera: Scarabaeidae) used in this study. Square brackets denote genera synonymised with *Scarabaeus* (*S.*) and parentheses denote subgenera of *Scarabaeus* recognised by Forgie *et al.* (in press). Outgroup taxa are highlighted in bold.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	***Circellium bacchus***	-																								
2	***Heliocopris hamadryas***	16.28	-																							
3	***Eucranium arachnoides***	17.93	18.22	-																						
4	***Anomiopsoides heteroclytus***	17.25	17.79	15.30	-																					
5	*S. [Drepanopodus] proximus*	17.46	17.47	16.41	17.29	-																				
6	*S. (Kheper) nigronaeneus*	18.75	18.33	16.28	18.04	13.73	-																			
7	*S. (Kheper) subaeneus*	20.19	17.89	16.85	19.22	14.42	12.41	-																		
8	*S. [Neateuchus] proboscideus*	15.57	16.61	17.37	16.33	11.12	12.90	13.30	-																	
9	*S. [Neopachysoma] rodriguesi*	20.33	19.88	20.31	19.64	15.58	15.57	16.94	14.12	-																
10	*S. (Pachysoma) bennigseni*	19.79	18.65	17.61	19.80	14.69	16.55	16.88	15.34	14.71	-															
11	*S. (Pachysoma) hippocrates*	21.60	22.60	19.39	21.24	17.34	15.58	16.65	15.38	14.92	15.45	-														
12	*Pachylomerus femoralis*	17.46	16.73	16.51	17.25	11.81	12.69	14.62	11.01	14.76	14.47	14.99	-													
13	*S. (Scarabaeolus) bohemani*	17.78	15.26	18.33	17.90	10.84	14.33	14.73	11.60	13.60	14.88	16.00	10.93	-												
14	*S. (Scarabaeolus) flavicornis*	17.03	17.13	17.29	17.16	11.32	13.92	14.63	11.03	14.74	15.10	15.90	11.83	9.58	-											
15	*S. (Scaerabaeolus) rubripennis*	17.58	17.24	17.46	16.84	10.94	13.91	14.32	10.44	13.84	14.76	16.10	11.31	8.89	8.08	-										
16	*S. galenus*	18.77	19.01	17.64	18.60	13.88	14.45	16.10	11.81	15.89	16.58	16.54	13.54	13.51	13.94	12.81	-									
17	*S. goryi*	18.74	16.82	16.73	18.11	12.30	11.70	13.91	10.63	15.03	12.72	14.64	10.72	11.53	11.33	10.91	13.73	-								
18	*S. rugosus*	18.86	18.66	19.87	19.66	12.72	13.81	14.44	13.31	15.27	17.56	16.43	13.21	12.91	12.83	11.71	14.67	12.60	-							
19	*S. rusticus*	19.41	18.84	18.91	19.10	14.37	14.96	15.56	13.82	16.29	17.23	19.06	14.74	13.72	12.42	12.01	14.22	13.01	11.64	-						
20	*S. satyrus*	18.67	18.47	18.03	17.83	12.12	13.10	14.75	10.74	14.36	15.38	16.75	11.11	13.12	12.34	11.91	13.53	10.62	12.91	14.63	-					
21	*S. westwoodi*	19.76	19.11	19.70	19.59	12.23	13.32	13.72	12.11	15.60	16.42	17.64	11.62	11.44	11.04	11.63	13.64	11.22	10.60	11.73	13.94	-				
22	*S. zambesianus*	16.31	17.68	16.09	15.80	10.82	12.70	13.61	4.86	13.82	15.51	14.94	10.82	11.89	9.85	9.66	11.81	10.23	12.61	13.71	10.05	11.80	-			
23	*S. (Sceliages) adamastor*	17.06	18.54	17.15	17.80	12.42	14.53	14.43	10.34	15.14	16.38	16.54	12.41	13.81	13.13	11.91	14.11	11.61	12.12	14.44	12.11	13.05	10.24	-		
24	*S. (Sceliages) brittoni*	16.94	18.35	17.60	18.04	12.45	14.35	14.76	10.15	15.48	15.98	16.15	11.72	13.23	12.64	11.73	14.02	11.83	12.24	14.44	12.01	12.45	9.86	4.00	-	
25	*S. (Sceliages) hippias*	17.16	19.00	17.61	17.82	11.83	14.31	14.13	10.44	15.27	16.21	16.89	12.03	12.91	12.94	11.12	13.81	11.93	12.14	13.74	12.24	13.27	10.73	7.01	7.59	-

TABLE 3. 16S rRNA mtDNA uncorrected sequence divergence values (%) between members of the Scarabaeini (Coleoptera: Scarabaeidae) used in this study. Square brackets denote genera synonymised with *Scarabaeus* and parentheses denote subgenera of *Scarabaeus* recognised by Forgie *et al.* (in press). Outgroup taxa higlighted in bold

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	***Circellium bacchus***	-																							
2	***Heliocopris hamadryas***	11.86	-																						
3	***Eucranium arachnoides***	11.48	12.68	-																					
4	***Anomiopsoides heteroclytus***	10.42	12.13	3.47	-																				
5	*Scarabaeus [Drepanopodus] proximus*	9.83	12.39	12.42	10.78	-																			
6	*S. (Kheper) nigronaeneus*	13.26	13.73	10.79	10.78	12.40	-																		
7	*S. (Kheper) subaeneus*	10.17	10.80	9.95	9.93	7.94	7.45	-																	
8	*S. [Neateuchus] proboscideus*	11.43	14.28	14.30	13.20	7.10	13.41	10.41	-																
9	*S. (Pachysoma) bennigseni*	14.72	15.47	12.32	12.36	8.18	12.69	9.41	13.68	-															
10	*S. (Pachysoma) hippocrates*	10.95	10.75	9.69	10.28	7.92	10.26	9.67	12.05	8.38	-														
11	*Pachylomerus femoralis*	12.77	13.38	10.42	9.89	6.51	11.79	8.54	11.37	9.39	7.12	-													
12	*S. (Scarabaeolus) bohemani*	12.48	13.80	12.75	11.65	4.97	11.88	9.21	10.07	7.94	5.90	7.03	-												
13	*S. (Scarabaeolus) flavicornis*	10.84	11.88	10.57	9.49	3.52	11.11	7.16	8.95	6.42	5.65	5.28	3.25	-											
14	*S. (Scaerabaeolus) rubripennis*	10.72	12.68	13.30	12.20	3.99	13.86	9.76	8.42	8.46	8.19	7.86	4.73	3.73	-										
15	*S. galenus*	12.85	14.33	13.60	13.34	7.71	13.27	9.00	12.39	10.08	10.56	7.79	8.47	7.47	8.49	-									
16	*S. goryi*	12.26	13.21	13.48	13.51	7.19	11.61	8.43	8.14	11.34	9.46	7.74	8.00	6.94	8.52	8.69	-								
17	*S. rugosus*	14.81	13.52	13.75	14.37	8.42	14.30	8.98	12.88	9.72	8.67	7.28	6.94	7.72	9.80	7.93	8.47	-							
18	*S. rusticus*	12.52	11.88	14.58	15.20	7.20	14.36	12.67	11.78	11.86	7.39	10.16	6.40	8.20	8.67	11.64	10.00	7.47	-						
19	*S. satyrus*	11.09	11.30	11.20	10.15	7.39	9.72	8.37	7.58	11.23	8.90	6.75	7.68	5.67	8.18	9.76	6.40	9.98	10.51	-					
20	*S. westwoodi*	11.96	12.09	10.77	10.77	6.40	10.49	9.75	9.17	10.01	7.95	6.04	5.24	5.49	6.53	9.77	7.21	8.42	7.98	4.21	-				
21	*S. zambesianus*	10.99	12.44	11.57	10.54	7.46	12.43	9.23	2.29	11.86	9.46	8.82	9.28	7.73	9.81	10.81	7.39	11.36	11.57	5.41	8.48	-			
22	*S. (Sceliages) adamastor*	12.63	13.20	11.61	10.51	5.17	11.83	7.43	9.18	7.68	6.39	6.79	4.46	3.02	4.95	6.91	7.69	8.41	9.00	5.16	6.15	6.94	-		
23	*S. (Sceliages) brittoni*	12.39	12.37	10.75	10.20	4.95	11.58	7.62	8.89	7.37	6.61	6.54	4.20	2.76	4.68	6.95	7.93	8.71	8.68	5.64	5.95	7.19	1.15	-	
24	*S. (Sceliages) hippias*	12.38	12.09	11.82	11.03	5.43	12.12	8.38	8.90	7.87	5.63	7.05	4.69	3.24	5.17	7.20	6.90	8.46	8.68	4.92	5.94	6.69	1.60	2.77	-

Mean sequence divergences between ingroup taxa with relationships supported by bootstrap analysis particularly in NJ and ML trees ranged from 4.0-15.3% in COI and 1.2-7.5% in 16S data. This, in conjunction with high A-T richness and multiple substitutions in the most variable sites of the molecular data, suggests saturation of substitutions occurs in taxa with sequence divergence values above the mean range of each data set (i.e. above 15.3% in COI and 7.5% in 16S). Mean divergences within the ingroup subgenera varied between 6.2 % (COI):1.9% (16S) (*Sceliages*), 8.9%:3.9% (*Scarabaeolus*), 12.4%:7.5% (*Kheper*) and 15%:8.4% (*Pachysoma*) S. L. Among the closely related taxa within each subgenus, the overall divergence was relatively low, homoplasy was therefore low, hence, the noise:signal ratio was in favour of truly homologous base substitutions. Highest divergence values occurred between ingroup and outgroup taxa yet the basal node differentiating the scarabaeines from the outgroup taxa in both trees received strong bootstrap support. Resolution, at least in the COI data, was likely to be gained from the highly conserved sites as reported by Mardulyn and Whitfield (1999).

Several studies have provided molecular clock calibrations for insect mt DNA ranging from approximately 0.98-2.3% divergence per million years to estimate phylogenetic time frames (DeSalle *et al.*, 1987 (2.0%); Brower, 1994 (2.3%); Prüser and Mossakowski, 1998 (0.98-2.3%)). Based on this range, our data suggests the Scarabaeini appeared around 8-19 Mya. The current school of thought suggests the Scarabaeini may have evolved around the same time as other Scarabaeines during the Eocene epoch (37-54 Mya) of the Cenozoic (Crowson 1981; Cambefort 1991a; Scholtz and Chown 1995). Our molecular data suggest this tribe is more recently derived than previously speculated. Clay covered brood balls and nests recovered from the Chadian Pliocene Australopithecine levels (Duringer *et al.*, 2000) suggest brood ball construction and nesting behaviour practiced by many of the Scarabaeini were well established at least 3-3.5 million years ago. According to our estimates, this advanced level of reproductive behaviour had at least 4 million years to evolve in the tribe.

Combined Data Analysis and Characteristics

A partition homogeneity test of the three data sets supported their combinability (P = 0.001). Analysis of each combination recovered a single most parsimonious tree (Fig. 4). Analysis of the molecular data sets either simultaneously (Figs 4A; CI = 0.30, RI = 0.39, length = 3.173 and 4B; CI = 0.33, RI = 0.45, length = 1.680) or together (Fig 4D; CI = 0.30, RI = 0.39, length = 3.413) with the morphological data yielded topologies reflecting an over proportional impact by the latter (but see Hillis, 1987). This may stem from a greater weight given to individual homoplasious morphological characters than saturated molecular characters supporting nodes at a number of hierarchical levels. The majority of topologies recovered from separate COI and 16S analyses were largely conflicting and poorly supported apart from relationships between closely related taxa (see discussions below). The tree recovered from the combined molecular data (Fig. 4C; CI= 0.33, RI= 0.37, length = 2,200) bears little to no improvement in topological robustness and prompts us to question their effectiveness in further resolving morphologically based Scarabaeini phylogeny. Given that our molecular data sets markedly differ in rates of evolutionary change (see Brown *et al.*, 1982; DeSalle *et al.*, 1987; Knight and Mindell, 1993), the pooling of these heterogeneous data may yield incorrect topologies (Bull *et al.*, 1993) that are poorly supported (Brower and DeSalle, 1994), thus, providing us with less confident estimations of relationships. Nonetheless, inclusion of the molecular with morphological data recovered a single "total evidence" tree (Fig.4D; 683 PIC's) that contained some groupings compatible with certain elements of the preferred weighted morphological tree by Forgie, Philips and Scholtz (Fig. 1A, unpubl.).

Saturation

The mean empirical Scarabaeini TI:TV scores calculated in PAUP for all COI and 16s rRNA characters were 1.29 and 1.00 respectively. In contrast, the overall transversions exceeded

transitions in ratios of 0.64 and 0.72 for all COI and 16S rRNA characters respectively when analysed in MacClade. Similarly, by comparing the CI and RI values, with and without the removal of transitions and/or uninformative characters, it is evident a high degree of homoplasy is present in most of the unweighted data, particularly COI. The same holds true even following weighting against presumably highly homoplasious nucleotides most prevalent at third codon positions in the molecular data. This site tends to become saturated quickly due to a higher frequency of silent substitutions than replacement substitutions in genes, particularly in genes subject to strong selection (Swofford *et al.*, 1996). Equally, saturation is problematic when inferring relationships among taxa that have been diverged for a long time (Brower and DeSalle, 1994; Källersjö *et al.*, 1999). Brown *et al.* (1982) estimated 10-30 million years for saturation of silent sites in parts of two protein genes (URF 4 and 5 after Anderson *et al.*, 1981) and three transfer RNAs (i.e. His, Ser, Leu) of hominoid primate mtDNA. Whilst third codon positions are therefore thought to be less informative indicators in phylogenetic studies than more slowly evolving first and second codon positions, Källersjö *et al.* (1999) have shown the contrary: third positions, although highly homoplasious, contain most of the phylogenetic signal in their data. Weighting schemes aimed to reduce or eliminate highly variable [homoplasious] positions from nucleotide sequences tend to decrease phylogenetic signal rather than noise (Philippe *et al.*, 1996) and therefore does not fundamentally increase general congruence (Vidal and Lecointre, 1998). Indeed, weighting of our data had no significant effect in improving congruency or resolution of many of the mid to deep-level nodes in tribal phylogeny as 16S and the third codon positions of COI gene appear to be too saturated. For similar problems encountered with COI and 16S data Mardulyn and Whitfield (1999: 290-1) suggest the level of divergence of the generic relationships examined are located in a window in which rapidly evolving sites are too saturated and highly conserved sites are not variable enough to provide sufficient phylogenetic signal.

Despite the lack of topological congruence, all trees support the following behavioural inferences: Feeding specialisation in terms of shifts from coprophagy (*Kheper, Pachylomerus, Scarabaeus, Scarabaeolus*) to necrophagy (*Sceliages, Scarabaeolus, Scarabaeus* (In part)) and saprophagy (*Pachysoma* S. L. (In part)) is polyphyletic in Scarabaeini evolution. The same inference can be made for food relocation behaviour when "ball-rolling" is defined in terms of mode and direction i.e., rolling backwards (*Scarabaeus, Kheper, Pachylomerus*) pushing forwards (*Pachylomerus, Sceliages, S. galenus*), dragging forwards (*Pachysoma*) S. L. or carrying forwards (*S. galenus*) (refer to Fig. 1A). Both *Pachylomerus femoralis* and *S. galenus* practice tunnelling behaviour in addition to "ball-rolling". Their disparate placement within all topologies suggests a polyphyletic reversal back to an ancestral tunnelling behaviour (Philips, Pretorius and Scholtz, unpubl.).

The general pattern emerging from all analyses indicates significant disagreement in the hypotheses generated. For instance, *S. (Sceliages)* taxa have a medially derived placement within the ingroup in virtually all trees recovered. The exception lies with the 16S MP trees (Figs 3A, B) where its representatives become the most highly derived clade within the tribe. Moreover, conflicting inferences are made as to the clade's relatedness with other lineages between the different data sets. *S. (Sceliages)* is closely related to the noturnal taxa *S. satyrus, S. proboscideus, and S. zambesianus* in all COI trees recovered, *S. (Pachysoma)* S. L. in 16S MP trees and members of *S. (Scarabaeolus)* in the morphology and 16S MP trees. The majority of trees reconstructed place the *Kheper* lineages most basal in the ingroup as sister to the remaining members of *Scarabaeus* S. L. This inference has strong bootstrap support only from distance analyses of both genes (Fig. 6). In contrast, the morphology and reweighted COI MP and COI ML trees place *Kheper* taxa among the more derived lineages of the Scarabaeini. Flightlessness in the ingroup is limited to *S. (Pachysoma)* S. L. in the molecular data and is subject to the

highest degree of conflict in terms of topological placement and relationship with fully winged taxa. In the morphological data where flightless lineages are well represented, there is virtually complete reversal in ingroup polarity from a paraphyletic flightless origin of scarabaeine evolution (results not shown) to the loss of flight becoming a derived condition that has evolved monophyletically (Fig. 1A). Flightlessness is generally accepted as a derived condition from a macropterous ancestry (Darlington, 1936; Goldschmidt, 1940; Southwood, 1962; den Boer *et al.*, 1980; Harrison, 1980; Kavanaugh, 1985; Roff 1986, 1990. Cited by Emerson and Wallis, 1995. See also Scholtz, 2000) and is derived in the morphological trees recovered from a majority of weighted schemes conducted by Forgie *et al.*, (in press).

All analyses were characterised by low consistency and retention indices and a low degree of resolution. Bootstrap analysis collapsed the majority of the nodes in all trees recovered apart from the basal node differentiating the Scarabaeine clade from the outgroup taxa and those on the apical branches exhibiting strongly supported relationships among closely related taxa. These include the two nocturnal species *S. satyrus* + (*S. proboscideus* + *S. zambesianus*), *S. rugosus* + *S. rusticus* and *Scarabaeus* subgenera *Sceliages* and *Kheper*. *Scarabaeus (Pachysoma)* S. L. taxa were only moderately supported as a monophyletic clade following bootstrapping in the COI data but gained strong support in the COI NJ tree (Fig. 6). In contrast, members of *S. (Scarabaeolus)* (with the exception of *S. (S.) scholtzi*) were well supported by bootstrap analysis as a monophyletic clade only in the COI NJ tree (Fig. 6) and Combined MP trees (Figs 4A, D).

The conflicting topologies and poor resolution in the separate analyses are problematic in establishing to what degree the data are able to support both phylogenetic signal and hypotheses of intra-tribal relationships that emerge when the data are combined (Vogler and Welsh, 1997; Durando *et al.*, 2000). Moreover, it becomes difficult to consider whether or not the combined analysis is the best phylogenetic estimate for the tribe (Vogler and Welsh, 1997). We suspect not. Although the "total evidence" tree shares several elements in common with the preferred

weighted morphological tree by Forgie *et al.* (Fig. 1A, in press), heterogeneity among data sets as mentioned is apparent even though none of the strongly divergent topologies are well supported. There are many arguments for and against combined analysis of multiple data sets in phylogenetic inference (see review by Nixon and Carpenter, 1996). In essence, combination of different data sets providing phylogenetic signals at different but complementary hierarchical levels results in improved overall resolution. Where our combined data falls short of this notion, it does provide us with useful phylogenetic information when analysed separately for two reasons: (1) heterogeneity can be circumvented when looking separately at areas of congruence and conflict of trees (Nixon and Carpenter, 1996); (2) the independence of separate analyses increases the significance of corroboration (Miyamoto and Fitch, 1995).

Comparison with Forgie, Philips and Scholtz (in press) classification of the Scarabaeini

The proposed classification of the Scarabaeini by Forgie *et al.*, (in press) is based on comprehensive morphological phylogenetic evidence. Genera forming monophyletic clades within *Scarabaeus* S. L. lineages that remained well supported after subjection to even the most stringent weighting schemes against character homoplasy (PIWE concavity indices 1 and 2) were given subgeneric status (i.e. *Kheper* and *Sceliages*). The genus *Pachylomerus* was maintained as it appeared basal to *Scarabaeus* S. L. in the majority of trees recovered, while the only remaining genus, *Drepanopodus*, was synonymised with *Scarabaeus* due to its derived placement within *Scarabaeus* S. L. and lack of both statistical support and distinct apomorphic characters. The two existing subgenera, *Pachysoma* S. L. and *Scarabaeolus* (excluding *S. (S.) scholtzi*) were adequately supported as derived lineages within *Scarabaeus* S. L. to warrant their taxonomic maintenance. With the exception of *Pachylomerus*, phylogenetic signal in our molecular data is apparent in several congruently supported clades representing the principal genera and subgenera of the Scarabaeini according to Forgie *et al.*, (in press). The placement of

Pachylomerus within the molecular framework of *Scarabaeus* S. L. lineages however is discordant with the proposed classification. It is worth noting that NJ analysis of both molecular data sets suggest *Kheper* lineages are sister to those of *Scarabaeus* S. L. . Pairwise sequence divergences between *Kheper* and *Scarabaeus* lineages are moderate to high in COI (11.7-16.9%) and 16S (7.5-14.4%), but not more so than divergences between *Scarabaeus* lineages. Unfortunately, there is insufficient weight from either the COI or 16S data alone to argue against the proposed classification of either *Kheper* or *Pachylomerus*. However, further investigation with more conserved genes is likely to clarify any ambiguity.

Eucraniinii vs Scarabaeini

The putative close association between the Neotropical Eucraniini and the Scarabaeini was believed to be based on morphological convergence of characters (Zunino *et al.*, 1989; Philips *et al.*, 2002) likely to be associated with existence in arid environments where all flightless lineages of these tribes occur. The molecular component of this study was able to test the hypothesis that the close relationship between the Eucraniini and the Scarabaeini is the result of morphological convergence and is not due to common ancestry. Both molecular data sets infer an obvious genetic dissimilarity between the eucraniines and the morphologically congruent scarabaeines particularly members of the subgenus *Scarabaeus* (*Pachysoma*) MacLeay with 19.6% and 11.2% mean sequence divergences between the *E. arachnoides* + *A. heteroclytus* and *S. (Pachysoma)* S. L. lineages for the COI and 16S data respectively. These values are above the saturation thresholds indicated for each but are no more divergent than the comparisons scored between each eucraniine and the remaining scarabaeine lineages. While the eucraniines and *S. (Pachysoma)* S. L. lineages are morphologically very similar, their genes are not. All molecular topologies recovered from MP, ML and NJ analyses do not support a close relationship between them with the MP and ML COI trees (apart from the TI:TV=1.3 reweighting tree; Fig. 2D) inferring the least measure of relatedness. Our molecular data

therefore concur with Philips *et al.*(2002) morphological and Ocampo's (unpubl.) molecular findings supporting a convergence hypothesis in the evolution of eucraniines and *S. (Pachysoma)* S. L. lineages rather than one of parallel evolution, stemming from a common "Gondwanan" ancestor. For us to accept the latter hypothesis, there would have to be a high degree genetic similarity between them and a tribal evolution far preceding the earliest estimates of African-South American separation of West Gondwanaland around 120-150 Mya (Thayer, 1985).

Conclusions

It is obvious the limitations of this study result from conflicting signals and poor resolution between the morphological and molecular data sets making it difficult to estimate the true phylogeny of the tribe when these data are combined. Resolution is achieved in all analyses between the closely related taxa providing good support for Scarabaeini systematics and several phylogenetic inferences. Of these, the NJ trees provided the hypotheses most strongly supported by bootstrap analysis. Nonetheless, COI and 16S genes have not contributed to fully resolving tribal relationships other than those between closely related lineages and this is likely due to the ancient rapid radiation of the group. Sampling of rare North and North East African flightless taxa is necessary to gain molecular perspective on the levels of divergence between these lineages and those of South Western Africa and whether or not there is molecular support for the monophyly of flightlessness.

Acknowledgments

The authors would like to thank Alfried Vogler and Daegan Inward of the Britsh Museum of Natural History for their collaboration. Thanks also to Miguel Archangelsky, CRILLAR, Argentina, and Daegan Inward for ethanol preserved material. This study was funded by the National Research Foundation (NRF), South Africa, and the University of Pretoria.

References

Anderson, S., Bankier, A.T., Barrell B.G., de Bruijn, M.H.L., Coulson, A.R., Drouini, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger A.F., Schreier, P.H., Smith A.J.H., Staden, R. and Young, I.G. (1981) Sequence and organisation of the human mitochondrial genome. *Nature*, **290**, 457-465.

Baker, R.H. and DeSalle, R. (1997) Multiple sources of character information and the phylogeny of Hawaiian drosophilids. *Systematic Biology*, **46**, 654-673.

Barbero, E., Palestrini, C. and Roggero, A. (1998) Phylogenetic relationships in the genus *Scarabaeus* Linnaeus (Coleoptera, Scarabaeidae, Scarabaeinae). *Atti I Colloquio Nazionale di Sistematica Cladistica*, Verona, Italy, 6-8 Nov., 87-96.

Belshaw, R., Quicke, D.L.J., Völkl, W. and Godfray, H.C.J. (1999) Molecular markers indicate rare sex in a predominantly asexual parasitoid wasp, *Evolution,* **53**, 1189-1199.

Broughton, R. E., Stanley, S.E. and Durrett, R.T. (2000) Quantification of homoplasy for nucleotide transitions and transversions and a reexamination of assumptions in weighted phylogenetic analyses. *Systematic Biology*, **49**, 617-627.

Brower, A.V.Z. (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Science, USA*, **91**, 6491-6495.

Brower, A.V.Z. and DeSalle, R. (1994) Practical and Theoretical considerations for choice of a DNA sequence region in insect molecular systematics, with a short review of published studies using nuclear gene regions. *Entomological Society of America*, **87**, 702-716.

Brown, W.M., Prager, E.M., Wang, A. and Wilson, A.C. (1982) Mitochondrial DNA sequences of primates: tempo and mode of evolution. *Journal of Molecular Evolution*, **18**, 225-239.

Bull, J.J., Huelsenbeck, J.P., Cunningham, C.W., Swofford, D.L. and Waddell, P.J. (1993) Partitioning and combining data in phylogenetic analysis. *Systematic Biology*, **42**, 384-397.

Cambefort, Y. (1991a) Biogeography and Evolution, pp. 51-74. *In*: I. Hanski and Y. Cambefort (Eds.), *Dung Beetle Ecology*. Princeton University Press, Princeton.

Cambefort, Y. (1991b) From saprophagy to coprophagy, pp. 22-35. *In*: I. Hanski and Y. Cambefort (Eds.), *Dung Beetle Ecology*. Princeton University Press, Princeton.

Caterino, M.S., Cho, S. and Sperling, F.A.H. (2000) The current state of insect molecular systematics: A thriving tower of babel. *Annual Review of Entomology*, **45**, 1-54.

Caterino, M.S., Shull, V.L., Hammond, P.M. and Vogler, A.P. (2002) Basal relationships of Coleoptera inferred from 18S rDNA sequences. *Zoologica Scripta*, **31**, 41-49.

Chalwatzis, N., Hauf, J., Van De Peer, Y., Kinzelbach, R. and Zimmermann, F.K. (1996) 18S ribosomal RNA genes of insects: Primary structure of the genes and molecular phylogeny of the Holometabola. *Annals of the Entomological Society of America*, **89**, 788-803.

Chin, K. and Gill, B. D. (1996) Dinosaurs, Dung Beetles, and Conifers: Participants in a Cretaceous Food Web. *Palaios*, **11**, 280-285.

Clary, D.O. and Wolstenholme, D.R. (1985) The mitochondrial DNA molecule of *Drosophila yakuba*: Nucleotide sequence, gene organisation and genetic code. *Journal of Molecular Evolution*, **22**, 252-271.

Cognato, A.I. and Sperling, F.A.H. (2000) Phylogeny of *Ips* DeGeer species (Coleoptera: Scolytidae) inferred from mitochondrial cytochrome oxidase 1 DNA sequences. *Molecular Phylogenetics and Evolution*, **14**, 445-460

Crowson, R.A. (1981) *The Biology of the Coleoptera*, Academic Press, London, 802 pp.

Darlington, P.J. (1936) Variation and atropy of flying wings of some carabid beetles (Coleoptera). *Annals of the Entomological Society of America*, **29**, 136-179.

den Boer, P.J., Van Huizen, T.H.P., Den Boer-Daanje, W., Aukema, B. and Den Beiman, C.F.M. (1980) Wing polymorphism and dimorphism in ground beetles as stages in an evolutionary process (Coleoptera: Carabidae). *Entomologia Generalis*, **6**, 107-134.

Derr, J.N., Davis, S.K., Woolley, J.B. and Wharton, R.A. (1992) Variation in the phylogenetic utility of the large ribosomal subunit of mitochondrial DNA from the insect order Hymenoptera. *Molecular Phyolgenetics and Evolution*, **1**, 136-147.

DeSalle, R. (1992) The phylogenetic relationship of flies in the family Drosophilidae deduced from mtDNA sequences. *Molecular Phylogenetics and Evolution*, **1**, 31-40.

DeSalle, R., Freedman, T., Prager, E.M. and Wilson, C.A. (1987) Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *Journal of Molecular Evolution*, **26**, 157-164.

Durando, C.M., Baker, R.H., Etges, W.J., Heed, W.B., Wasserman, M. and DeSalle, R. (2000) Phylogenetic analysis of the *replete* group of the genus *Drosophila* using multiple sources of characters. *Molecular Phylogenetics and Evolution*, **16**, 296-307.

Duringer, P., Brunet, M., Camberfort, Y., Likius, A., Mackaye, H.T., Schuster, M., Vignaud, P. (2000) First discovery of fossil dung beetle brood balls and nests in the Chadian Pliocene Australopithecine levels. *Lethaia*, **33**, 277-284.

Emerson, B.C. and Wallis, G.P. (1995) Phylogenetic relationships of the *Prodontria* (Coleoptera: Scarabaeidae; Subfamily Melolonthinae), derived from sequence variation in the mitochondrial cytochrome oxidase II gene. *Molecular Phylogenetics and Evolution*, **4**, 433-447. unpubl.

Farris, J.S., Källersjo, M., Kluge, A.G. and Bult, C. (1995) Constructing a significance test for incongruence. *Systematic Biology*, **44**, 570-572.

Felsenstein, J. (1981) A likelihood approach to character weighting and what it tells us about parsimony and compatibility. *Biological Journal of the Linnaen Society*, **16**, 183-196.

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

Forgie, S.A., Philips, T.K. and Scholtz, C.H. (in press) Evolution of the Scarabaeini. *Systematic Entomology*, **X**, XX-XX.

Funk D.J. (1999) Molecular systematics of cytochrome oxidase 1 and 16S from *Neochlamisus* leaf beetles and the importance of sampling. *Molecular Biology and Evolution* **16**, 67-82.

Funk, D.J., Futuyma, D.J., Ortí, G. and Meyer, A. (1995) Mitochondrial DNA sequences and multiple data sets: A phylogenetic study of phytophagous beetles (Chrysomelidae: *Ophraella*). *Molecular Biology and Evolution* **12**, 627-640.

Gatesy, J., DeSalle, R. and Wheeler, W. (1993) Alignment –ambiguous nucleotide sites and the exclusion of systematic data. *Molecular Phylogenetics and Evolution*, **2**, 152-157.

Gibson, T., Higgins, D.G. and Thompson, J. (1994) CLUSTAL X: Program and documentation available at ftp://ftp.ebi.ac.uk/pub/software/mac/clustalw/clustalx.

Goldschmidt, R. (1940) *The Material Basis of Evolution*. Yale University Press, New Haven, CT.

Goloboff, P.A. (1997) *PIWE*, Version 2.6, (1993-1997). Published by the author, Instituto Miguel Lillo, Miguel Lillo 205, 400 Sierra Madre de Tucuman, Argentina.

Grandcolas, P., Deleporte, P., Desutter-Grandcolas, L. and Daugeron, C. (2001) Phylogenetics and ecology: as many characters as possible should be included in the cladistic analysis. *Cladistics*, **17**, 104-110.

Gu, X., Fu, Y.X. and Li, W.H. (1995) Maximum-likelihood estimation of the heterogeneity of sunstitution rate among nucleotide sites. *Molecular Biology and Evolution*, **12**, 546-557.

Halffter, G. and Halffter, V. (1989) Behavioral evolution of the non-rolling roller beetles (Coleoptera: Scarabaeidae: Scarabaeinae). *Acta Zoologica Mexicana (Nueva Serie)*, **32**, 1-53.

Harrison, R.G. (1980) Dispersal polymorphisms in insects. *Annual Review of Ecology and Systematics*, **11**, 95-118.

Harrison, J. du G. and Philips, T.K. (2003) Phylogeny of *Scarabaeus (Pachysoma* MacLeay) stat.n., and related flightless Scarabaeini (Scarabaeidae: Scarabaeinae). Submitted to: *Annals of the Transvaal Museum*, **40**, 47-71.

Hasegawa, M., Kishino, H. and Yano, T. (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **21**,160-174.

Hillis, D.M. (1987) Molecular versus morphological approaches to systematics. *Annual Review of Ecological Systematics*, **18**, 23-42.

Hillis, D.M. (1998) Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Systematic Biology*, **47**, 3-8.

Hillis, D.M. and Bull, J.J. (1993) An empirical test of boostrapping as a method of assessing confidence in phylogenetic analysis. *Systematic Biology*, **42**, 182-192.

Joy, D.A. and Conn, J.E. (2001) Molecular and Morphological phylogenetic analysis of an insular radiation in Pacific Blackflies (*Simulium). Systematic Biology*, **50**, 18-38.

Juan, C., Oromi, P. and Hewitt, G.M. (1995) Mitochondrial DNA phylogeny and sequential colonisation of Canary Islands by darkling beetles of the genus *Pimelia* (Tenebrionidae). *Proceedings of the Royal Society of London, B*, **261**, 173-180.

Källersjö, M., Albert, V.A. and Farris, J.S. (1999) Homoplasy *increases* phylogenetic structure. *Cladistics*, **15**, 91-93.

Kavanaugh, D.H. (1985) On wing atrophy in carabid beetles (Coleoptera: Carabidae), with special reference to Nearctic *Nebria,* pp 408-431 *In:* G.E. Ball (Ed.). *Taxonomy, Phylogeny and Zoogeography of Beetles and Ants*. Dr.W. Junk, Dordrecht, The Netherlands.

Kluge, A.G. and Wolf, A.J. (1993) Cladistics: What's in a word? Cladistics, **9**, 183-199.

Knight, A. and Mindell, D.P. (1993) Substitution bias, weighting of DNA sequence evolution and the phylogenetic evolution of fea's viper. *Systematic Biology*, **42**, 18-31.

Kobayashi, N., Tamura, K., Aotsuka, T. and Katakura, H. (1998) Molecular phylogeny of twelve Asian species of epilachnine ladybird beetles (Coleoptera: Coccinellidae) with notes on the direction of host shifts. *Zoological Science*, **15**, 147-151.

Lafay, B., Smith, A. B. and Christen R. A (1995) Combined morphological and molecular approach to the phylogeny of Asteroids (Asteroidea: Echinodermata). *Systematic Biology*, **44**,190-208.

Langor, D.W. and Sperling, F.A.H. (1997) Mitochondrial sequence divergence in weevils of the *Pissodes strobi* species complex (Coleoptera: Curculionidae). *Insect Molecular Biology*, **6**, 255-265.

Langor, D.W. and Sperling, F.A.H. (1995) Mitochondrial DNA variation and identification of bark weevils in the *Pissodes strobi* species group in western Canada. *Canadian Entomologist*, **127**, 895-909.

Lee, M.S.Y. (2000) Tree Robustness and Clade Significance. *Systematic Biology*, **49**, 829-836.

Liu, F.R. and Miyamoto, M.M. (1999) Phylogenetic assessment of molecular and morphological data for eutherian mammals. *Systematic Biology*, **48**, 54-64.

Liu, H. and Beckenbach, A.T. (1992) Evolution of the cytochrome oxidase II gene among 10 orders of insects. *Molecular Phylogenetics and Evolution*, **1**, 41-52.

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

Maddison, W.P. and Maddison, D.R. (1992) "MacClade: Analysis of Phylogeny and Character Evolution, Version 3.08," Sinauer Associates Inc., Sunderland, Massachusetts.

Mardulyn, P. and Whitfield, J.B. (1999) Phylogenetic signal in the COI, 16S, and 28S genes for inferring relationships among genera of Microgastrinae (Hymenoptera: Braconidae): Evidence of a high diversification rate in this group of parasitoids. *Molecular Phylogenetics and Evolution*, **12**, 282-294.

Mathee, C.A. and Robinson, T.J. (1997) Molecular phylogeny of the Springhare, *Pedetes capensis*, based on mitochondrial DNA sequences. *Molecular Biology and Evolution*, **14**, 20-29.

Miyamoto, M.M. and Fitch, W.M. (1995) Testing species phylogenies and phylogenetic methods with congruence. *Systematic Biology*, **44**, 64-76.

Mostert, L. E. and Scholtz, C. H. (1986) Systematics of the subtribe Scarabaeina (Coleoptera: Scarabaeidae). Entomology Memoir, Department of Agriculture and Water Supply, Republic of South Africa, **65**, 1-25.

Nixon, K.C. and Carpenter, J.M. (1993) On outgroups. *Cladistics*, **9**, 413-426.

Nixon, K.C. and Carpenter, J.M. (1996) On simultaneous analysis. *Cladistics*, **12**, 221-241.

Pääbo, S. (1989) Ancient DNA: Extraction, characterisation, molecular cloning, and enzymic amplification. *Proceedings of the National Academy of Science USA*, **86**, 1939-1943.

Philips, T.K., Pretorius, R. & Scholtz, C.H. (2004) A phylogenetic analysis of dung beetles (Scarabaeinae: Scarabaeidae): unrolling an evolutionary history. *Invertebrate Systematics*, **18**, 1-36.

Philips, T.K., Scholtz, C.H. and Ocampo, F.C. (2002) A phylogenetic analysis of the Eucraniini (Coleoptera: Scarabaeidae: Scarabaeinae). *Insect Systematics and Evolution*, **33**, 241-252.

Pretorius, R., Philips, T.K. and Scholtz, C.H. (2001) Geometric morphometrics, the metendosternite and its use in phylogenetics of the Scarabaeinae (Coleoptera). *Elytron*, **14**, 125-148.

Prost, R.J., Flook, P.K. and Millest, A.L. (1993) Methods for the preservation of insects for DNA studies. *Biochemical Systematics and Ecology*, **21**, 85-92.

Prüser, F. and Mossakowski, D. (1998) Low substitution rates in mitochondrial DNA in Mediterranean carabid beetles. *Insect Molecular Biology*, **7**, 121-128.

Roff, D.A. (1986) The evolution of wing dimorphism in insects. *Evolution*, **40**, 1009-1020.

Roff, D.A. (1990) The Evolution of Flightlessness in Insects. *Ecological Monographs*, **60**, 389-422.

Roy, M.S., Girman, D.J., Taylor, A.C. and Wayne, R.K. (1994) The use of museum specimens to reconstruct the genetic variability and relationships of extinct populations. *Experientia*, **50**, 551-557.

Sanderson, M. J. and Wojciechowski, M. F. (2000) Improved Bootstrap Confidence Limits in Large-Scale Phylogenies, with an Example from Neo-Astragalus (Leguminosae). *Systematic Biology*, **49**, 671-685.

Saiki, R.K., Chang, C.A., Levenson, C.H., Warren, T.C., Boehm, C.D., Kazazian, H.H. and Erlich, H.A. (1988) Diagnosis of sickle cell anemia and beta-thalassemia with enzymatically amplified DNA and nonradioactive allele-specific oligonucleotide probes. *The New England Journal of Medicine*, **319**, 537-541.

Saitou, N. and Nei, M. (1987) The neighbour-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406-425.

Scholtz, C.H. (2000) Evolution of flightlessness in Scarabaeoidea (Insecta, Coleoptera). *Mitteillung des Museums fu"r Naturkunde Berlin, Deutsche entomologische Zeitschrift*, **47**, 5-28.

Scholtz, C.H. and Chown S. L. (1995) The evolution of habitat use and diet in the Scarabaeoidea: A phylogenetic approach, pp 354-374. *In*: Pakaluk, J. and Slipinski, S.A. (Eds.) *Biology, Phylogeny and Classification of Coleoptera: Papers celebrating the 80th birthday of Roy A. Crowson*. Museum I Instytut Zoologii PAN, Warszawa.

Silvain, J.-F. and Delobel, A. (1998) Phylogeny of West African *Caryedon* (Coleoptera: Bruchidae): Congruence between Molecular and Morphological Data. *Molecular Phylogenetics and Evolution*, **9**, 533-541.

Simon, C., Frati, F., Bechenbach, A., Crespi, B., Liu, H., and Flook, P. (1994) Evolution, Weighting, and Phylogenetic Unity of Mitochondrial Gene Sequences and a Compilation of Polymerase Chain Reaction Primers. *Annals of the Entomological Society of America*, **87**, 651-701.

Simon, C., Nigro, L., Sullivan, J., Holsinger, K., Martin, A., Grapputo, A., Franke, A., and McIntosh, C. (1996) Large differences in substitutional pattern and evolutionary rate of 12S ribosomal RNA genes. *Molecular Biology and Evolution*, **13**, 923-932.

Swofford, D. L. (1999) PAUP*: Phylogenetic analysis using parsimony (*and other methods), Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Swofford, D.L., Olsen, G.L., Waddell, P.J. and Hillis, D.M. (1996) Phylogenetic inference, pp 407-514. *In*: D.M. Hillis, Morowitz, C., and Mable, B.K. (Eds.), *Molecular Systematics*, 2nd Ed., Sinauer, Sunderland, MA, USA.

Thayer, M.K. (1985) Revision, phylogeography and biogeography of the austral genus *Metacorneolabium* Steel (Coleoptera: Staphylinidae: Omaliinae), pp113-179. *In*: G.E.Ball (Ed.) *Taxonomy, Phylogeny and Zoogeography of Beetles and Ants*. (Junk: Dordrecht).

Vidal, N. and Lecointre, G., (1998) Weighting and congruence: A case study based on three mitochondrial genes in pitvipers. *Molecular Phylogenetics and Evolution*, **9**, 366-374.

Villalba, S., Lobo, J.M., Martín-Piera, F. and Zardoya, R. (2002) Phylogenetic Relationships of Iberian Dung Beetles (Coleoptera: Scarabaeinae): Insights on the Evolution of Nesting Behaviour. *Journal of Molecular Evolution*, **55**, 116-126.

Vogler, A.P. and DeSalle, R. (1993) Phylogenetic patterns in coastal North American tiger beetles, *Cicindela dorsalis* Say, inferred from mitochondrial DNA sequences. *Evolution*, **47**, 1192-1202.

Vogler, A.P., DeSalle, R., Assmann, T., Knisley, C.B. and Schultz, T.D. (1993) Molecular population genetics of the endangered tiger beetle, *Cicindela dorsalis* (Coleoptera: Cicindelidae). *Annals of the Entomological Society of America*, **86**, 142-152.

Vogler, A.P. and Welsh, A. (1997) Phylogeny of North American *Cicindela* tiger beetles inferred from multiple mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **8**, 225-235.

Walsh, S.P., Metzger, D.A. and Higuchi, R. (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*, **10**, 506-513.

Wayne, R.K., Leonard, J.A. and Cooper, A. (1999) Full of sound and fury: History of ancient DNA. *Annual Reviews of Ecology and Systematics*, **30**, 457-477.

Whiting, M., Carpenter, J.C., Wheeler, Q.D. and Wheeler, W.C. (1997) The Strepsiptera problem: Phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology*, **46**, 1-68.

Wieblen, G.D. (2001) Phylogenetic relationships of fig wasps pollinating functionally dioecious Ficus based on mitochondrial DNA sequences and morphology. *Systematic Biology*, **50**, 243-267.

Wiegmann, B.M., Regier, J.C. and Mitter, C. (2002) Combined morphological and molecular evidence on the phylogeny of the earliest lepidopteran lineages. *Zoologica Scripta*, **31**, 67-81.

Zunino, M., Barbero, E. and Luzzatto, M. (1989) Food relocation behaviour in Eucraniina beetles (Scarabaeidae) and the constraints of xeric environment. *Tropical Zoology*, **2**, 235-240.

Appendix 1a. Variable character matrix for Cytochrome Oxidase subunit I (COI). Question marks (?) denote unknown nucleotides. Names of taxa starting with '*S*' belong o the genus *Scarabaeus* L. Moreover, subgenera and synonyms of *Scarabaeus* are denoted by parentheses and square brackets respectively. Outgroup taxa are highlighted in bold.

Taxon/Node	1111111111111111111111111111111111 1122223333344444456666677777888999900001111122222233334444555666667778 9582457036790235691123461257814734692568124780367892568145806902568145 70
Circellium bacchus	TATCTCCCTTTATATTTTAATAATATGCTATAAGTTCATATGTTCTCTAGAAATATTCGATATATCAAAATA
Heliocopris hamadryas	TATCTTAAAATAAACTTTAATAGAGAACCTAATGTCCACATGTTGACTATAACTATTTGTTCTATTATAATT
Eucranium arachnoides	TATCTCAATTCAATCTTTAATAGAGTACTTATAGTCCATATGTTATCTGGTTTTAATTATCACTTTATAGTC
Anomiopsoides heteroclytus	CATCTCAAATTAATTATTAATAGAGCACAATGTGTCATTAAGTTATTTGGTTAGATTCGATCCCTCATAACA
S. [Drepanopodus] proximus	TATTATATATATAATTTTAATAGTGATCTTTTTGTTAACTCATTATCTACAATTATCTGATTTATCTTAATC
S. (Kheper) nigroaeneus	TATTATACTAATAATTTTATTAGAGAGCATAAAGATTTTATATTATTTACTATTAATTGTTCCTTTACAACC
S. (Kheper) subaeneus	TATCTTATTAATATTTTTATTAGAGATCAAAATGCTCACTCATTATTTACATATAACTGTTTCTTTATTATT
S. [Neateuchus] proboscideus	TATCATAATAATAATTTTAATAGTGAACAATTAGTTAATATATTATATACTAATATTTGTTATATTAAAATA
S. [Neopachysoma] rodriguesi	TATTATATTCATATTTTTAATAGAGAACTATGTGTCTTTAAATTATCTACTTATTTTTAATCTATTAGAATC
S. (Pachysoma) bennigseni	TATCATACTTATATTTTTAATAGAGTCCCATTTGGCCTCAAATTACCTACAGATTTTTGTTGCTTCATAATA
S. (Pachysoma) hippocrates	TACTATACCCCTATTTTTAGCTGAACGTTTTTTACCTTTATATTATTTACCGATTTTTGTTCTATCACAATA
Pachylomerus femoralis	TATTATAAATATAATTTTAATAGAGAACTTATTGTTAATAAAATGTTTACAATTATTTGTTACTTCAAAACT
S. (Scarabaeolus) bohemani	TATTATATAAATAATTTTAATAGAGATCAAAATGTCTTCTTAACATCCACATCTATTTGTTACACCAGAATT
S. (Scarabaeolus) flavicornis	TATTACATAAATAATTTCAATAGAGAGCCATTAGACATCTAAACATCTACTGATACTTGTCTCTTCAAAATT
S. (Scarabaeolus) rubripennis	TATTATATAAATAATTCTAATAGAGAGCCATAAGTCCACTAAATATCTACTAATATTTGTTTCTTTAAAATT
S. galenus	CATCTTACCAATATTTTTAATAGAGGGCAATATGTTATTAAATTATATACAGTTAATTGTTTTATTAGAATT
S. goryi	TATTATATTAATAATTTTAATAGAGAGCAAATAGTTAATAAAATGAATACAATTATTTGATACTTTATAATC
S. rugosus	TATTACATAAATAATTTTATTAGTGATCTTTTTGATACCTAATTGTTTACAATTAATTGATTCTTTAATATA
S. rusticus	TTTTATACATATAATTTCATTAATGGGCAATTAGATATCTAAATGTTTACGATTAATTGTTGCTTTATAATT
S. satyrus	TACTATAATTATAATTTTAATAGAGAGCTTTATGATAATAAAATGTCTACATATATTTGATTCTTCTTTACC
S. westwoodi	TATTACACATATAATTTCATTAGTGATCAAATTGATTCCTAACTATCTACAATTAATTGTTTCTTTAATATT
S. zambesianus	TATTATAATAATAATCTTAATAGTGAGCGATTAGACAATATATTATCTACTAATATTTGTTACATTAAAATA
S. (Sceliages) adamastor	TATTATATATATAATTTTGATAGTGATCTATTTGTTAATAAATTATTTACTAATAACTGATACTTTATTATT
S. (Sceliages) brittoni	TATTATATATATAA?TTTAATAGTGATCTATTTGTTAATAAATTATCTACTAATAATTGATACTTTATTATT
S. (Sceliages) hippias	CATTATAATTATAATTTTAATAGTGATCTACTTGTTAATAAATTATTTACAAATATTTGATACTTTGTTATT

COI variable characters (continued):

Taxon/Node	1111111122222222222222222222222222222222222222222233333333333333333333 8888999900011111233333444455556666677777888889999990000001111222333344 3679256814803469913478013925891456703678902589124578013679256814703679 02
Circellium bacchus	ATCTCTATTTTTTGTTGTCAAAATATAACCTTGCTAATTCTATATTACTCCCTTATCTTTCTTGATAACTCT
Heliocopris hamadryas	ATTAGTGATTAATGCTATCAAAAAATAATCTTGCTATAGCTTAAATATACTCTATTCACATATGATTACTTA
Eucranium arachnoides	TTCTTTGATTAAAGATTACTAGAATTTAATATGTTAAAGTTTAATGATACACATAATAATTAAAACAACATA
Anomiopsoides heteroclytus	TTCTTAGATTATAGATTTCAAATAATAACTATGTTAAAGTATAACGATATACTTAATAATCTATATAATATA
S. [Drepanopodus] proximus	TTTAGTATTTATAATTTACTAAAAATTATTATGCTATAGGTATATTATACTCATATTAATCAAATTTACATA
S. (Kheper) nigroaeneus	ATCTCTATTTATTATTTTCTAAAATTAAATATGCTATAGGAATATTATATATACTTTATTTAAGTTAACATA
S. (Kheper) subaeneus	TTCTGTATTTATTATCTACTAAAATTAATTATGCTATAGGAATATTATATATACTATATATATATTAACATA
S. [Neateuchus] proboscideus	ATTAAAATTTATTATTTTCTTAAATTAATTATGCTAATGGTATTTTATATATACTTTAACTATTATTATATA
S. [Neopachysoma] rodriguesi	ATTAACATTTATCATTTACAAAAATTTATTATGCTGAAGGGATTCCTCTCTTACCCTACATATCATAACATA
S. (Pachysoma) bennigseni	A?TATTATTTAAAATTTACCAAAATCTGTTATGCGATAGGTATAATATATACTTACC?ATTA?CATTATATA
S. (Pachysoma) hippocrates	ATTAGTATCTAAAATCTATTAAAATTTATTATGCTGAAGGAATTACATATATATAATAAATATGTTAATATA
Pachylomerus femoralis	ATTAATATTTATAACTTTCTTAAATTAATTACACTATAGGAATATTATATACATAATATATATTATTACGTA
S. (Scarabaeolus) bohemani	ATTAGCATTTATTATTTACTAAAAATAACTATGCTGAAGGAATATCACACTCTTAATAATTATTTTTACTCA
S. (Scarabaeolus) flavicornis	ATCAGTATTCAATATTTACCAAAAATTATTATGCTATAGGTATATTATACTTATAATAAACAAATTAACTTA
S. (Scarabaeolus) rubripennis	ATTAGTATTTATTATTTACTAAAAATTATTATGCCGACGGAATACCATACCCATAATATATAATTTAACTTA
S. galenus	ATTAATATTTATAATTTACAAAAATTTATTATGCTAATGGAATTACATACTCTCTTTGATTAAATTAACACA
S. goryi	ACTAGTATCTAAAATTTACTAAAATTAATTATGCTATAGGTACATTATATATACTATAATTATGTTTATATA
S. rugosus	ATTAAAATCTATTATTTACTAAAAATTAATACCCTATAGGAATTTCTTACTCAGTTTATATATATTAACTTA
S. rusticus	ATTATTATTTATTATTTACTAAAAATTATTATGCTATTGGAACACCTTACACACTATAATTAAGTTAACTTA
S. satyrus	ATCAATATTTATTATTTTCAAAAATCGAATATGCTATTGGAATTTTATATATACTTTAAATAAAATAATGTA
S. westwoodi	ACCAGTATTTATTATTTACTAAAAATTATTATGCTATAGGAATATTATACATACTTTAATTATTTTAACTTA
S. zambesianus	ATTAATATTTATTATTTTCTTAAATTAATTATGCTAAAGGAATTTTATACTTACTTTAAATAAAATAATATA
S. (Sceliages) adamastor	ATCAATATTTAATATTTACTAAAAATTATTATACAATTGGAACTTCACTCATACTTCTACTATTCTTATACT
S. (Sceliages) brittoni	ATTAATATTTAATATTTACTAAAAATTATTATGCAACTGGAATTTCACTCATACTTCTACTATTTTTATGCT
S. (Sceliages) hippias	ATTAATATTTATTATTTACTAAAAATTATTATGCAATTGGAATTCTACTTATACTACTATCATATTTGCACT

COI variable characters (continued):

Taxon/Node	333333333333333333334444444444444444444444444445555555555555555555555555 445555666777788889999000011112223344455566677888000001111222334444555556 581457369258901470369025812473692514846925847039014790369238470379256894
Circellium bacchus	CTTTTAATCTTTATAATGTTTTATATTATTTATTATTTAAATATTTTTGAAAAAATTTTGTCAAGATATATA
Heliocopris hamadryas	TTTTATTTTTTTATTAAGTTTCTTATTATTTCTTTTTTATATTCTTTTGAAGAATTAATATTTAGTTTTACC
Eucranium arachnoides	TTATTATATTCTGTATTTTTTTACATTACTTCTTTTTCAATTTTTTTGACAAAAATTTTTTTTAGTTTTATT
Anomiopsoides heteroclytus	ATACTACCCTCAACTAGATATTACATTATCTATATTTTATGTATTTTAGAAGGAATTATATTAAGACTCATT
S. [Drepanopodus] proximus	ATACTAAACTTAATAAAATATTATATTATTTTTATCTTATATTTATTAGAAAAATTAACATTTAAATATATA
S. (Kheper) nigroaeneus	ATATCTCTTTTTATAAATCATTATATTATTTACTTTCTAATTTTATTAGAAAAATTTTTTTTTAAATTTATA
S. (Kheper) subaeneus	ATATCTTCTTTAATAAATTATTATATTATTTTTTTTTTATTCGTATTGGAAAAATTTATATTTAATTTTATA
S. [Neateuchus] proboscideus	TTATTATTTTTAATAAAATATCACACTATTTATTTTTTATGTATCTTAGAAAAATTTACATTTAAATCTATA
S. [Neopachysoma] rodriguesi	ATATTAATTTCAATAAGATTTCTTATCTTTTTCTTTCCTAATGTCTTCGAAAAACTATTCCTAAATCGCTCA
S. (Pachysoma) bennigseni	ATATTAACTTTTATACGATATTATATTATTTCCATTCCCTATGTTTCGGTAAAATTATTGTTTAATTTTATA
S. (Pachysoma) hippocrates	ACACTAATTTCTATAAGACTTTATGTTATTTTTTTTTCTCTTATTCTCATAAAATTATTTTTTAATTGTATA
Pachylomerus femoralis	ATATTAACTTCTATAAATTATCTTATTATTTTTTTTTCTTGTATTTTAGAAAAATTATTATTCAAATGTATA
S. (Scarabaeolus) bohemani	ATATTAATTTTAATAATTTTTCATATTATCTTTATTTCCTATTCATTAGAAAAATTAACATTTAAATATATT
S. (Scarabaeolus) flavicornis	TTACTATATTTAATAATCTATCATATTATTTTCATTTTATATACTTTAGAAAAATTCATATTTAAATCCATA
S. (Scarabaeolus) rubripennis	TTATTAAATTCAATAAAACATTATATTATCTTCAATTTATATGTATTAGAAAAATTAATATTTAAATTTATA
S. galenus	TTATTAAATCCTATAAGTTATCTTACTATTTTTTTTCCTAATTTCTTAGAAAAATCTTTTTTTAAATATATA
S. goryi	ATATTATCTTCAATAAGTCATTATATTATTTACATTCCTAATTTTTTTGAAAAATTAATTTTTAATTTTATA
S. rugosus	ATATTAATCTTTATAATATTTTATATTATCTTCAATTTAAATTCTTTGGAAAAATTAATATTAAAATATACA
S. rusticus	ATTTTAAATTTTATACTTTTTCATATTATTTTTATTCTATATATATTTGGAAAGTTAATATTTAAATATACA
S. satyrus	ATATTATCTTCAATAAGATTTTATATTATTCACTTTTCTTATTTTTTAGAAAAATCAATTCTAAAATTTATA
S. westwoodi	ATATTATCTTCTATAAATTATCATATTATTTTTTATTCTATGTTTTTTAGAAAAACTAACATTTGAATACATA
S. zambesianus	ATACTATATTCAATAAAATATTATATTATTTATTTTTTATATATTTTAGAAAAACTTATATTAAATTATATA
S. (Sceliages) adamastor	TTATTATATTCTATAAAATACTATATTATTTACATTTTATTTTTTTTAGAAAAATTTATATTAAAATTTATA
S. (Sceliages) brittoni	TTATTATATTCTATAAAATACCATATTATTTATATTTTATCTTTTTTAGAGAAATTTATATTAAAATTTATA
S. (Sceliages) hippias	TTACTATATTTAATATAATTTTATATCATTTATATTTTATATTTCTTAGAAAAATTTATATTTAAATTTATA

COI variable characters (continued):

Taxon/Node	555555555556666666666666666666666666666666677777777777777777777777777777 677778899990000112223333445555667788889999900000111222222333344455555666 723692514570369581470369251247035812470136901268147013469356814701369025
Circellium bacchus	TTGGTTTTTGTAATTCATTTTTTTATAGTATATAATATTCTTTATTTATTCCGCTATGATATTATCATTGTG
Heliocopris hamadryas	TTTACTTCAGATATACACTTTAATATAGTTAATAATAACCATCACATATTCCTATATGTAAATATTACAATG
Eucranium arachnoides	TTAATCTTTATAATTTATTTATATATAGAAATCAGTATACTTCTCTCTCTTTTATATGTTAATATTATAGCT
Anomiopsoides heteroclytus	TTTAACTTAATAACTTATTTAATAATTGATAACTATAAATATTACTTATCTATATATGATAATACTATAGTT
S. [Drepanopodus] proximus	TTAACTTTAGAAATTAATTCTATAATTGAAATTAATATATTTTGCTCTTTAATATATGTTATTATCATAGAG
S. (Kheper) nigroaeneus	TGTAATTTAGAAATTTATTCTTAAATTGGCATTAACACATTTAACTATTTATTATATGTTAGTATTATTGTC
S. (Kheper) subaeneus	TTTAATTTAGTAATTAATCTAAAAATTGTAATTTATATTTTTAACTATTTATTATATAATAATTCTATTAAT
S. [Neateuchus] proboscideus	TTAAATTTAGTAATTAAATTAATAATTGTTAATAATAATTTTTTCTCTTTAATATATGAAAATTCCATAGTA
S. [Neopachysoma] rodriguesi	TTAAATTTAGAAATTTATTTCATTATCATTAATTATATTTCTATCTCTTTTATTAATTTTAGTATTATAGTG
S. (Pachysoma) bennigseni	TTAGTCTTAGAAATATGATTTATAATTGTAAATGATATCTCTCTCCTACTCATATATGTTAATATTACTGAG
S. (Pachysoma) hippocrates	TGTATTTTAGAAATATTCTTTAAAATTATCAACTGTACTTTCAACTCTTTAATATATTCCTTCTCTATTGAA
Pachylomerus femoralis	TTAATTTTAGATATTATTTCTAATATCGTTTATAATATCTTTAACACTTTATTATATGTTATTATTATAGAA
S. (Scarabaeolus) bohemani	TTAATTTTAGATATGAATTTTTATATTGTTAATAATAAATCTATCTCTTTATTTTATGTAAATATTATTGTG
S. (Scarabaeolus) flavicornis	TTAAATTTAGTTATAAGTTTTTTAATTGTTTATAATACTTTTATCTCTTTATTTTATGTAAATATTATAGTC
S. (Scarabaeolus) rubripennis	TTAAATTTAGTAATTAAATTTACAATTGATTACTATGATTTTATCTCTTCATTTTATGTAAATATTATAGTG
S. galenus	TTTATTTTTGGTGTTATTTTTAGAATAGACTTTAATATATTTATCTTATCATTATATGACAATACTATAATG
S. goryi	TTAACTTTAGATATATAATTTTAAATTGATAATAATAATTTTAACTTACTACTATATGTAAATATTGTTGAT
S. rugosus	TTAAATTTAGAAATTAAATTTAAAATTATTATTGATAATTTTAACTTATTATTTTTGTGATAGTTTTATAGAA
S. rusticus	CTAAATTTAGTAATTATACTCAATATTATTAATAATATTCATAACTCTTCAATATACTATAATATTATAGTT
S. satyrus	TTAAATCTAGATATTAAATCCTAAGTTGTAAATAATATATTTATCTCTTTCATATATGTAAATACTACAGTC
S. westwoodi	TTAAATTTTGATATTAATCTTAAAATTATTAATCACATATCTCTCTCTTCACTTTATTATAATATTATAGAA
S. zambesianus	TTAAATTTAGTAATTAAATTAATAATTGATAATAATAATTTTTTCTCTCTATTATATGAAAATTCTATAGTA
S. (Sceliages) adamastor	TTAAACTCTGAAATTAAATTTAATATTGATATTAATAATTCTAACTCTTTTTTACATGACATTACTATAGAT
S. (Sceliages) brittoni	CTTAATTTAGAAATTAAACTTAATATTGATATTAATAATTTTGACTCCTTCTTATATGTAATTACTATAGAA
S. (Sceliages) hippias	TTAAATTTAGGAACTATATTAAATACTGTTATTAATAATTTCAACTCTTCTTTACATGTTATTATTATAGAC

COI variable characters (continued):

Taxon/Node	77777777777777788888888888888888888888888888888899999999999999999999999 66777788888999900001111222333444455556677788999990000111122233333455556 89147801349258914580136258147036923584703658123472367258923401236904567O
Circellium bacchus	TTATAGAGATCTAAGTTAGTCTTTTTAATCTTTATAATTATTTATATTTTCCTTTATACATTCTATTTATTT
Heliocopris hamadryas	ATATCGAATTATTTATTAGTTATACTAATTTTTTTAAGTATTTACACATTCTTTATTGTATAGAATTTTTAT
Eucranium arachnoides	ATGATGAGTTTTAAATTGGATATTTTTTCTTTTATAAGAATATATGTCTTTTTTTTATATATTCTATTTTTAT
Anomiopsoides heteroclytus	ATATGGAATTATAAACTAGTTATTCCAATTTTTATATATATACACATTTTCATCAATACTTACTATTTCTAT
S. [Drepanopodus] proximus	CTAAGGAGTCTTAAATTAGTTATTTTATCCCTTTTAAGTATTAATGTTCTTTTTAGTATATGCTATTTTTAT
S. (Kheper) nigroaeneus	ATATGGAATTTTAAATTAGATATTTTATTTTTTTCTTATATAAATGTCTTTTTTTTACACTTTTAATCTTTAA
S. (Kheper) subaeneus	TTAAAGTGATATATATTAGATATTCTTTTTTTTTATAAAATTAATTACACTTTTTAATACATTCTATTTTTAT
S. [Neateuchus] proboscideus	TTAAAGTGATTCTAATTAGTCTTTTTTTACCTCTTCTAATATTTATGTTTTTTTTTAATACATTCAATTTTTAT
S. [Neopachysoma] rodriguesi	ATAATGTGATTTCTATTAATCATTCTATATTTTTCTAATATTAGTATTCATATTAATACTTTCTATTTTTAT
S. (Pachysoma) bennigseni	ATACAGCGATTTTAATTAG?CCTTTTATATTTTACCA?AATAATCGTATTCTTTTTATATATTCTATTTTGAT
S. (Pachysoma) hippocrates	ATATTGTGATTTCAGTCAGACTTTTCAACTTTTATAAGTATAAATGTACTTTTTTAATACATTCTACTTTTAT
Pachylomerus femoralis	ATATTGAGATTTATATTAGACTTTTTTAATTCTTTCTTGAATCTATGTATTCATTTATACTTACTATTTTTAT
S. (Scarabaeolus) bohemani	ATAATGAGACTTTTATCAGTCTTACTAACTTTTACCACTATTAATGTATTTTTTTAATACATACTATTTTTAT
S. (Scarabaeolus) flavicornis	ATAATGAGATTTAAGTTAGACTTACTAAATTTTACTAATATTAATGTTCTCATTAACACATGCTATTTTTAT
S. (Scarabaeolus) rubripennis	ATAAAGAGATTTAAATTAGTCTTCTTAAATTTTACTAGAACTTATGTTTTTTTTTAATACTTTCTATTTTTAT
S. galenus	GCATAGTAACATAAATCAATCACTTTTTATTTTTCTTAAATATATGTTTTTTTTTAACACATACTACTTTTAT
S. goryi	ATAAAGTGATTTAAACTAGTCTTATCAAATTTTTCTAGAATAATTGTATTTACTTTTACATCCAATTTTTAT
S. rugosus	ATATAATGTTATAAATTAGTTATTTTTATTTTTTACCAAAATTAATGTATTTATTAATATATACTACTCTTAT
S. rusticus	ACAAAGTACTCTTAATTAGTTATACTTAATTCTACTAAAATAATTATTTTCTCTAACACTTACTATTTCTAT
S. satyrus	TTATAGAGATTTAAATTAGATATTTTTAATTCTTCTAGTATTAATGTATTTTCTAATACATTCTGTTTTTAT
S. westwoodi	ATATAATACCACAAATCAATCTTATTTATTCTCTCTAAAATTAACGTTCTTATTTTTACATCATACTTTTAT
S. zambesianus	TCAAGGTGACTTAAATCAGTCTTTTTTACTCTTTCTAATATTCATGTTTTCTTTAATACATTCAATTTTTAT
S. (Sceliages) adamastor	CTAATGTATTTTAAATTAGTCCTTTTAACTTCTATAATAATTTATGTATTTATTAATACATTCTATTTTTAT
S. (Sceliages) brittoni	ATAATGTATTTTAAATTAGCCTTTTTAACTTCTATAATAATCAATGTGTTCATTAATACATTCTATTTTTAT
S. (Sceliages) hippias	ATAAAGTATTTTAAATTAGTCTTCTTAACTTTTATAATAACTTATGTATTTATTAATACACTCTATTTTTAT

COI variable characters (continued):

Taxon/Node	1111111111111111111111111111111111111111111111111111111111 9999999999990000000000000000000000000000000000000000111111111111111 6677778889999000001111122223334445555566666777888899999000001111122222 6925891470369235891345703692581470345925689147036792568901478013671245
Circellium bacchus	TTAATAACCCCACTTAATAGTGCATAATTTCACCTACAATATATAATAATATCTAGTAATTTTATTTTTTTA
Heliocopris hamadryas	TTTATAACTCCATTTATTAGAGTAACTTATTATCTACTGTATACTATAATATATAGTAATATTATGTTTTTA
Eucranium arachnoides	TTAATGAATTCATCTAATAGTGAGAAATATCTTATATAATGTATAACAAAATCTAATTTTTTTATCTTTTTA
Anomiopsoides heteroclytus	TTAATAAACCTATTTAATAGAGTAAATTATCATTTATAATATATAATATAGCATAGTTATTCTACTTTTTTA
S. [Drepanopodus] proximus	TTAATAACTCTATCCATTAGAGTGATATATTTTTTCTAATAAGAATTACTATCTAATTATTTGACCTTTTTG
S. (Kheper) nigroaeneus	TTTACAATTCGATTCAATAGTGAATTATACCTTACTTTATAAGATTTAATATATAACTATTTAATTTTTTTA
S. (Kheper) subaeneus	TTATTAACTCCATCCTTTATAGAATAATATCTTATTTAATAAGATACAATATATAATTATTTAATTTTTTCA
S. [Neateuchus] proboscideus	TTAATAATTCTATCTATTAGAGTATAATTTTATATTTAATAAGATATAATACTTAATTATTTTATTTTTTTA
S. [Neopachysoma] rodriguesi	TTATTAACTCCATTCTTCTGAGAATCTTTCTTTATTTAATAAAACTTCATATATAATTATTTTACTGATTTA
S. (Pachysoma) bennigseni	CTATTAACTCTGTTCTCCTGAGAGATTTATCACATTTAATAAGACTTTACATTTAATTTTTTTATCGATTTG
S. (Pachysoma) hippocrates	TAATTAATTTAATTCAATAGAGAATCATTTCCTACTTAATAATATACAATATTCCATTACTATATTGATTCA
Pachylomerus femoralis	TTATTAACTCTATCCAATAGAGCATATTTTCTTTTTTAATAAGAAATAATACATAATTTTTTCTTTTTTTCA
S. (Scarabaeolus) bohemani	TCTTTAACTCTATTTATTAGAGTAAATTATTTTTTTTAATAAGAAATGCTATATAATTATTTGATTTTTTCA
S. (Scarabaeolus) flavicornis	TTATTAATTCCATTTAATAGAATATAATATTATATTTAATAAGAAATACTATACAGTTATTTTATTTTTTTA
S. (Scarabaeolus) rubripennis	TTATTAACTTTATTTATTAGAGTAATATATTATATTTAATAAGAAATGCTACATAATTATTTACTTTTTTTA
S. galenus	TTAACAAATTTATCTATTAGAGTAAAACTTTATCTTTAATAAGATATACCATTCAATTATTTGATTTTTTTA
S. goryi	TTATTAAATCAATTCAATAGAGAAAAACATTTTATTTAATAAGAAATAATATATAATTATTCAATTTTTTTA
S. rugosus	TTATCTAATTTACTTATTAGAGAATAACACTTTACTTAATAATATATAATATATGATTATTCATCTTTCATA
S. rusticus	TTAATTAATTAATTTATTAGTGAAAAACATTTTATTTAATAACCATTGATATCCAATTATTTATCTTTTTTA
S. satyrus	TTATTAAACTTATTCTTTAGAGTATAATTCTTTTTTTTAATAAGAAATGTCATATAATTATTCCACTTTTTTA
S. westwoodi	CTATTAGCTCTATTCATTAGAGAATAATATTTTACTTAATAAGAAATTATATATAATTACTTATTTTTTTTA
S. zambesianus	TTAATAAATTTATCTAATAGAGTATAATTTTATATTTAATAAGACATAATATATAATTTTTTCATTTTTTTA
S. (Sceliages) adamastor	TTATTAATCTAATCCAATAGCGTATAACTTTATCTCTAATAAGATATAATATATAATTTTTTGACTTTTTTA
S. (Sceliages) brittoni	CTATCAATCTAACCCAACAGTGTATAATTTTATTTCTAATAAGATATAACATATAATTTTTTAACTTTTTTA
S. (Sceliages) hippias	TTACCAACTTTATCTATTAGAGAATAACTCCATTTTTAACAAGAAATAACATATAATTTTTTGATTTTTTTA

COI variable characters (continued):

Taxon/Node	1111111111111111111111111111111111111111111111111 1111111111111111111111111111111111111111111111111 3344444444455566666667777777778888888889999999 4701234678926812345780123456789012456789013456 7
Circellium bacchus	TATATTTTTTAAGTTAGAACTAAGAATATCAACACCCTTGAAGAATA
Heliocopris hamadryas	AATATTTTATGATTTGGAATTAAGACTGCCAACTCTTTAGAAGATAT
Eucranium arachnoides	TACATTTAATAATATGGAATTAAATATAAATTCTCTCTTGAAGATTT
Anomiopsoides heteroclytus	TATGTATTATAATATGGAATTAAGAATAAACTCACTATTGAAGATTA
S. [Drepanopodus] proximus	TATATTGTATAATATGCCCTCAAGAACATCATCACAATTGAAGGTTC
S. (Kheper) nigroaeneus	AATACTTTATAATATATCCTTAAGAACAAACTCACTATTGAGGATTT
S. (Kheper) subaeneus	AATATTTAATAATATAGACTTAGAACAATTCATCATTGATGATCAAT
S. [Neateuchus] proboscideus	TATATTGTATAATATACACTTAAGAACTTCATCACAATTGAAGATTT
S. [Neopachysoma] rodriguesi	AATTCCTTATATTTTATCCATAAGAACATCTTCACTATTGAAGATTT
S. (Pachysoma) bennigseni	ATTACCTAACAATACATATTTAAGAACGTCATCACAATTGAAGATTT
S. (Pachysoma) hippocrates	ATTACCTAATAATATACAAACAAGAACATCGTCACTATTGAAGATTT
Pachylomerus femoralis	TTTATTGTATAATATGCACTTAAGAACATCTTCACAATTGAAGATTT
S. (Scarabaeolus) bohemani	CATATTTTATAATATACTCTTAAGAACATCATCACAATTGAAGATTT
S. (Scarabaeolus) flavicornis	TATATTTAATAATTTGCACTTAAGAACATCATCACAATTGAAGATTC
S. (Scarabaeolus) rubripennis	TTTATTTTATAATTTGCTCATAAGAACATCATCTCAATTGAAGATTC
S. galenus	TATACAGAATAATATGCACCTAAGAACATCATCTCAATTGAAGATTT
S. goryi	TATATTGTATAATATACACATAAGAACTTCATCACAATTGAAGATTC
S. rugosus	AATATTTTATAATTTACACACTAGAACATCATCACAATTGAAGATTT
S. rusticus	AACATTTTATAATTTGTACTCTAGAACTTCATCACAATTGAAGAATA
S. satyrus	TATATTGTATAATTTGCACTCAAGAACTTCAACTCAATTGAAGATTT
S. westwoodi	TATATTTTATAATATACACATAAGAACATCATCACAATTGAAGATTA
S. zambesianus	TATATTGTATAATATACACTTAAGAACTTCTTCACAATTGAAGATTC
S. (Sceliages) adamastor	TATATCTTATAATTTATACATAAGAACCTCTACACAATTGAAGATTT
S. (Sceliages) brittoni	TATATCTTATAATATACACATAAGAACATCTACACAATTGAAGATTT
S. (Sceliages) hippias	TATATTTTATAATTTATACACAAGAACATCCTCACAATTGAAGATTT

Appendix 1b. Variable character matrices for 16S ribosomal RNA (16S). Dashes (-) represent gaps/deletions. Names of taxa starting with '*S*' belong o the genus *Scarabaeus* L. Moreover, subgenera and synonyms of *Scarabaeus* are denoted by parentheses and square brackets respectively. Outgroup taxa are highlighted in bold.

	1111111111111111111111111111111111111111111111111112222
	1112223337788888899990000111111122333334444455566666677777777888889999990000
Taxon/Node	6901413905909023456045824792345679010345735679078456789234567890348901235901 23
Circellium bacchus	ATTTTAGAAATATA-ATTAACTCTCCATTATAA-TTGCATCTTAATGCCCTAATTTATATAACTATTTAAGATTCAAC
Heliocopris hamadryas	ATTTTAACAAAATAATATAACTCTACATTATTA-AAGCATCTTAACGTCTTGAATTTTAAAAATATTAATAATATAAC
Eucranium arachnoides	ATTTAAACAATATA-ACTAACTATTCATTATTA-TTATATTTTGATGCCTAATTTAATATATTTATTTAAAATTTAAC
Anomoipsodes heteroclytus	ATTTTAA-AATATA-ATTAACTATTCATTATTA-TTACATTTTAATGCCTAATTTAATATATTCATCTAAAATTTAAC
S. [Drepanopodus] proximus	CTTTTTAAAAAATA-ATAAATTCTTCATTTTAA-TAGAATTTTAGTGCTTTTATATAAATACCTATCTTAAAATAAAC
S. (Kheper) nigronaeneus	ATCTTAAAAAAATAAATTAACTATATATTTTAA-TAGCATACCAGTGCTTTATAATAAATATTTATTATATAATAAAT
S. (Kheper) subaeneus	ATTTTTAAAAAATA-ATTAACTCTTCAATTGAAATAGTATATTAGTGCTTTAATTTAAATATTTATTTTATAATAAAT
S. [Neateuchus] proboscideus	GCTTTTAAAGTAAAAATATACTCTTCATATTATATAGATTTTTAGTGCCTAAATATATATATCTATCTATAACTAACC
S. (Pachysoma) bennigseni	ATTTTTAAAAAATA-ATTAATTTATCTTTATAA-TAGTATATTAATATCTATACCTATATGCTTTTTTAATGAAAAAT
S. (Pachysoma) hippocrates	ATTTTTAAAAAAAAAAATAATTTTATATTTTAA-TAGCATATCGATGTCTTTCTATAAATATTTTTTTAAAATTAAAT
Pachylomerus femoralis	ATTTTTAAAAAACA-AATGACTCTTCAGTTTAATTAGTATCTTAGTGCCTGTCTATAAATACTTACCTAAAATTACAC
S. (Scarabaeolus) bohemani	ATTTTTAAAATGTA-AATAATTTTCTATTTTAA-TAGACTTTCAGTGCTTTTATATATAAACTTATCTAAAAATAAAC
S. (Scarabaeolus) flavicornis	ATTTTTAAAAAATA-ATTAATTTTTCATTTTAA-TAGAATTTTAGTGCTTTTATATAAATACTTATCTAAAAATAAAC
S. (Scarabaeolus) rubripennis	ATTTTTAAAAAATA-AATAATCTTTCATTTTAA-TAGATATTTAGTGCTTTTATATAAATACCTATCTAGAAAAAAAC
S. galenus	ATTTTTAAAAAATA-AAAAAAAATTCAGTTTAA-TAGTATATTAGTACTCTTATATAAATAATAAATAATAAATAAAC
S. goryi	GTTTTTAAGAAATT-AATAACTCTTCAGTTTAA-TAGATTATTAGTGCCTGTATATTTATATTCAATAAAAAATAAAC
S. rugosus	ATTTTAAAAAAAATAAAAAACTCTATAGTTTAA-TAGTATTTCAGTGCTTGTATATATAAGTTTAATTAAAAATAAGC
S. rusticus	ATTTTAAAAAATTA-TATAATTCTATATTTTAA-TAGAATATCAATGTCTGTATACTTTAACCTAGTAAAAAGTAAAT
S. satyrus	ATTTTTAAAATATA-AATAACTCTTCATTTTTA-TAGAATTTTAGTGCCTATATATATATATTCATCTTTAAATAAAC
S. westwoodi	ATTTTTAAAAAATATAATACCTCTACATTTTTA-TAGACTTTTAGTGCTTGTATACATATATTTATCTAAAAAAAAAC
S. zambesianus	GCTTTTAAAGTAAA-ATATACTCTTCATATTAT-TAGAATTTTAATGCCTAAATATATATATTCATCTTTAATTAAAC
S. (Sceliages) adamastor	ATTTTTAAAATATAAAATAATTTTTCATTTTAA-TAGAATTTTAGTGCTTTTATACTTATATTCATCTATAAATAAAC
S. (Sceliages) brittoni	ATTCTTAAAATATAAAATAATTTTTCATTTTAA-TAGAATATTAGTGCTTTTATACTTATATTTATCTATAAATAAAC
S. (Sceliages) hippias	ATTTTAAAAAAATAAAATAATTTTTCATTTTAA-TAGAATATTAGTGCTTTTATATTTATATTCATCTATAAATAAAC

16S variable characters (continued):

Taxon/Node	22222222222222222222222222222233333333333333333333333344444444444444 00000111111111222223566677889990001112222233344556899911111233444555 45678012356789023483901246040892897890234901519476526901237425289089
Circellium bacchus	GATCAATTAATAAACGAAACGATTTTTAGTTATTTCCTCTCCAAATCTGTGGCAAAT-AACGAAATAC
Heliocopris hamadryas	TCTCTATTCTAAAATAAAACGAATCTTAATTATTTTTTTTCCCGTACTATGGCATAT-AACAATAACC
Eucranium arachnoides	TTTCAAAAAATCAAATTAACGTTTTTTAATTATTTCTTTTCGAGTACTATGGCAAACATACGAAATAC
Anomoipsodes heteroclytus	TTTCGAATAATTAAAGTAATGTTTTTTGATTATTTCTTTTCTCGTACTATGACAAGCATATGAAATAC
S. [Drepanopodus] proximus	TTTCAAATATTATATAAAATTATTTTTAATTATTTCTTCCCCAATATTATGGCGAATGTATGTAATAC
S. (Kheper) nigronaeneus	ATCCAAATATTAATATAAATAATATTTAATTATTTATTTCTCAATTTTAGGACAAAA-TACGATATCC
S. (Kheper) subaeneus	TTCTAATTATTAAATTAAATATATTTTAATAATTTATTTCTCAATTTTAGGACAATT-AATGATATCC
S. [Neateuchus] proboscideus	TTTCAAATAATATAGGGAATAATCTTTAAATATCTATTCCCCAATATTATGGCGAGTATTCGTAATCC
S. (Pachysoma) bennigseni	TTTTAAACAATACTTATTATAGTTTTTAATATTTTTTTTCCCAATATTATGACGAATATGTGTAATAC
S. (Pachysoma) hippocrates	TTTTAAATAATATATATAATAATTTTTAAAAATTTCTTCCCCAATATTATGACATATATACGAAATAC
Pachylomerus femoralis	TTTTAAATACTATATAAAATAGTTTTTAATTATTTCTTCCCCAATATTATAGTGAA---ATGTGATAC
S. (Scarabaeolus) bohemani	TTTTAAATAATATATAAAATAATTTTTAATTATTTTTCCCCCAATATTATGACGAATATATGTACTAC
S. (Scarabaeolus) flavicornis	TTTCAAATAAAATATAAAATAATTTTTAATTATTCTTTCTCCAACATTATGACGAATATATGTTATAC
S. (Scarabaeolus) rubripennis	TTTCAAATATTATATAAAATAATTTTTAATTATTTCCCCTCCAATATTATGACGAATATATGTAATAC
S. galenus	ATATAATTATTATATAAATTATTTTATTATTATTTCTCCCCCAAAATTATGACGAATATATGTAATAC
S. goryi	TCCCAAATAAAAAATAAAATAATTTTTAGTTATTTCTTCCCCAATGTTATGGCGAAT-TACGTAAGCG
S. rugosus	TTTTATACTATATATAAAATATATTTTAATAATTTCTCCCCCAATATTATGACGATTATATGAAATAC
S. rusticus	TTTTAGATTAAATATAAAATAATTTTTAATTATTTCTCCCCCAATATTATGGCGTAT-TACGAACTAC
S. satyrus	TTTCAAATAATATACGAAATAATTTTTAGTGATTTCTTCCCCAATATTATGGCGAGAATACGTTATCC
S. westwoodi	TTTTAAATAATATATAAAATAATCTTTAATTATTTCTCCCCCAATATTATGACGATA-TACGAAATAC
S. zambesianus	TTTCAAATAATATAGGGAATAATTTTTAAATATCTATTCCCCAATATTATGGCGAGTATCCGTAATCC
S. (Sceliages) adamastor	TTTTAAATAAAATATTAAATAATTTATAATAACTTTTTTCCTCAATATTATGACGAAT-TATGTAATAC
S. (Sceliages) brittoni	TTTTAAATAAAATATTAAATAATTTACAATTACTTTTTTTCCCAATATCATGACGATT-AATGTAATAC
S. (Sceliages) hippias	TTTTAAATAAATTATAAAATAACTTATAATAACTTTTTTCCTCAATATTATGACGAAT-AACGTAATAC

General Conclusions

Due to the size of the morphological character set and the robustness of the analyses, we do not expect any significant changes in overall tree topologies with the inclusion of additional morphological data or taxa.

High levels of homoplasy in the morphological data prevented the complete resolution of all intermediate and deep level nodes supporting relationships between less closely related lineages. It was predicted the inclusion of molecular data of the same taxa would further resolve these relationships. However, it became obvious the limitations of this study resulted from poor resolution achieved by the COI and 16S genes and conflicting signals and between the molecular and the morphological data sets. Whilst Resolution was achieved in all analyses between the closely related taxa providing good support for Scarabaeini systematics and several phylogenetic inferences, estimations of the true phylogeny of the tribe became difficult when all morphological and molecular characters were combined and analysed simultaneously.

In the first chapter we were able to describe some behavioural characteristics of the adult *Sceliages* beetles provisioning nests with millipedes for nidification. Many questions, however, remain unanswered: We know quinonous secretions of millipedes are responsible for attracting *Sceliages*, however, this was tested by stimulating a defensive reaction by millipedes. In a natural situation, are *Sceliages* beetles attracted to these secretions produced as allomones in response to the millipede being threatened or injured, and/or to these secretions being used as pheromones during millipede mate attraction and copulation? Do *Sceliages* beetles kill uninjured millipedes they may have been attracted to, or, must they rely solely on the demise of injured millipedes? Is *Sceliages* truly an obligate necrophage or are other food types also utilized? Are millipedes utilized for maturation feeding or nuptial courtship? Exactly how is the millipede

disarticulated? A leverage action using the clypeal teeth and protibial external denticles is inferred (Villalobos *et al.*, 1998) but has not been witnessed. We hope that these questions will stimulate further study on the biology of *Sceliages*.

Further attention needs to be directed towards the biology of the flightless Scarabaeini, *Mnematium ritchiei*, *M. silenus* and *Neomnematium sevoistra*. Sampling of these flightless taxa is necessary to gain molecular perspective on the levels of divergence between these lineages and those of South Western Africa, and whether or not there is molecular support for the monophyly of flightlessness.

Additionally, it would be interesting to further test relationships between the "ball-rolling" Scarabaeini that never or rarely horizontally relocate food resources (Halffter and Halffter, 1989) and those that do so exclusively. Both *Pachylomerus femoralis* and *Scarabaeus galenus* represent links between rolling and tunnelling by exploiting both behavioural strategies whilst equipped with true telecoprid morphologies