Syst. Biol. 53(5):711–734, 2004

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150490468710

The Origin and Radiation of Macaronesian Beetles Breeding in *Euphorbia*: The Relative Importance of Multiple Data Partitions and Population Sampling

BJARTE H. JORDAL[1,2] AND GODFREY M. HEWITT[1]

[1]*School of Biological Sciences, University of East Anglia, NR47TJ Norwich, United Kingdom*
[2]*(Current Address) Museum of Natural History and Archaeology, Section of Natural history, NTNU, NO-7491, Norway; E-mail: bjarte.jordal@vm.ntnu.no*

Abstract.—Species-level phylogenies derived from many independent character sources and wide geographical sampling provide a powerful tool in assessing the importance of various factors associated with cladogenesis. In this study, we explore the relative importance of insular isolation and host plant switching in the diversification of a group of bark beetles (Curculionidae: Scolytinae) feeding and breeding in woody *Euphorbia* spurges. All species in the genus *Aphanarthrum* are each associated with only one species group of *Euphorbia* (succulents or one of three different arborescent groups), and the majority of species are endemic to one or several of the Macaronesian Islands. Hence, putative mechanisms of speciation could be assessed by identifying pairs of sister species in a phylogenetic analysis. We used DNA sequences from two nuclear and two mitochondrial genes, and morphological characters, to reconstruct the genealogical relationships among 92 individuals of 25 species and subspecies of *Aphanarthrum* and related genera. A stable tree topology was highly dependent on multiple character sources, but much less so on wide population sampling. However, multiple samples per species demonstrated one case of species paraphyly, as well as deep coalescence among three putative subspecies pairs. The phylogenetic analyses consistently placed the arborescent breeding and West African—Lanzarote–distributed species *A. armatum* in the most basal position in *Aphanarthrum*, rendering this genus paraphyletic with respect to *Coleobothrus*. Two major radiations followed, one predominantly African lineage of succulent feeding species, and one island radiation associated with arborescent host plants. Sister comparisons showed that most recent divergences occurred in allopatry on closely related hosts, with subsequent expansions obscuring more ancient events. Only 6 out of 24 cladogenetic events were associated with host switching, rendering geographical factors more important in recent diversification. [Biogeography; data combinability; *Euphorbia*; host plant use; molecular systematics; Scolytinae; speciation; taxon sampling.]

Reconstruction of species-level phylogenies is a powerful tool in the study of macroevolutionary patterns of diversification, where the assessment of putative sister species provides a quantitative measure of the factors associated with cladogenetic events. These measures can only be as accurate as the phylogenies themselves, however, and various methodological problems may reduce the precision needed for such studies. The best practical approach to overcome these problems involves reconstructions under various optimality criteria, sampling every species and multiple populations of the focus group, and sampling characters from as many independent character sources as necessary to stabilize the phylogenetic hypothesis. Most studies on higher level phylogenies so far have sacrificed details for scale by including only one (or zero) individual per species, using only a single or a few sources of phylogenetic characters (Barraclough and Nee, 2001). Our aim in this study is to establish a robust phylogenetic hypothesis to explore important factors in the diversification of a group of bark beetles associated with *Euphorbia* spurges, by sampling five different character partitions and nearly all species and populations in the reconstruction of the phylogeny. During the process of exploring these data, we investigate the effect of partial sampling of characters and taxa, providing empirical results to complement recent theoretical discussions on these issues (Cummings et al., 1995; Poe and Swofford, 1999; Rosenberg and Kumar, 2001; Pollock et al., 2002).

Putative diversifying factors associated with cladogenetic events may be of two main categories, geographical isolation or ecological shifts (Barraclough and Nee, 2001). The relative strength of the comparative approach to assess speciation modes, however, depends logically on how accurately geographical and ecological differences can be defined, and how stable these are over evolutionary time scales (Losos and Glor, 2003). Perhaps the most precise model systems to explore patterns of colonization and geographic speciation are groups of species endemic to different oceanic islands, where each island in an archipelago has very similar barriers of open water between them (Juan et al., 2000; Emerson, 2002; Gillespie and Roderick, 2002). Similar or even stronger barriers to gene flow, in terms of ecological isolation, occur in organisms that specialize on narrow resources for feeding or reproduction. Some of the best examples on ecological isolation involves plants that serve as the sole host for herbivores, especially when the herbivores complete their entire life cycle within a single resource (Kelley and Farrell, 1998; Machado et al., 2000; Pellmyr and Leebens-Mack, 2000; Scheffer and Wiegmann, 2000; Despres et al., 2002).

Whereas many phylogenetic studies to date have focused on either geographical or ecological factors in speciation, rather few have tried to include both in a simultaneous study (Joy and Conn, 2001; Jordan et al., 2003; Percy, 2003). This paper explores the potential influence from both of these factors in the diversification of crypturgine bark beetles breeding in arborescent *Euphorbia* spurges on the Macaronesian islands. Macaronesia is made up of four archipelagos west of the African coast: the Azores (which do not host any crypturgine beetles), Cape Verde, and the Canary and Madeiran island groups. The oldest islands in each archipelago

(Fuerteventura, Porto Santo, and the eastern Cape Verde islands) are known to be at least 18 My, with younger islands gradually decreasing in age westward, with the most recent island only 1.1 My (El Hierro; see Hess et al., 2000, and Brown et al., 2001). Hence, these islands provide opportunities for diversification over an extended period of time, similar to the Hawaiian and Galapagos archipelagos, which have played prominent roles as evolutionary model systems (e.g., Gillespie and Roderick, 2002).

Arborescent growth forms have evolved several times independently in *Euphorbia* (Molero et al., 2002), and are particularly frequent on islands and coastal areas, as in many other oceanic plant groups (Bohle et al., 1996; Panero et al., 1999). Their prevalence on the Macaronesian islands and in drier parts of the African mainland has provided an unique opportunity for niche shifts and radiations in bark beetles of the weevil subfamily Scolytinae (Jordal et al., 2003). The largest clade of beetles associated with *Euphorbia* are found in the tribe Crypturgini, where all species of *Aphanarthrum* (19) and *Coleobothrus* (4) breed exclusively in dead or moribund parts of these plants. The majority of species are found on one or several of the Macaronesian islands (17) or western Africa (5), with the remaining species (5) reaching eastern parts of Africa and India. Five species have also been divided into two or three subspecies based on genitalic differences and allopatric island distributions (Israelson, 1972), suggesting that speciation is currently an ongoing process. The closest relatives to *Aphanarthrum* and *Coleobothrus* are most likely *Cisurgus* and *Crypturgus* whose species are morphologically very similar to each other (Wood, 1986). However, all species of the mainly Holarctic *Crypturgus* (10) breed in conifers only, whereas species of the Mediterranean *Cisurgus* (6) breed in various woody herbs, including two species breeding in succulent euphorbs in Morocco and the Canary Islands. The last two genera in Crypturgini are the monotypic *Dolurgus* (breeding in North American pines) and *Deropria* (associated with *Rubus* in the Canary Islands), believed to be distantly related to the other four genera (Wood, 1986).

Dead or moribund plant material serves simultaneously as both the food resource and as breeding space for adults and larvae of these beetles, where they complete the entire life cycle as in all other Scolytinae bark beetles. The level of host plant fidelity is high, and each species show distinct preference for only one of the four main groups of *Euphorbia* which have proven suitable as host plants (Table 1; Jordal, submitted). All host plants produce various forms of latex (Carter and Smith, 1988; Seigler, 1994) that are still present during the early stage of colonization, and are further distinguished by growth form and structural differences in tissue (Webster, 1994) as illustrated by the cactus-like succulents (e.g., *E. canariense*, *E. handiense*), and the shrub-like groups consisting of *E. lamarckii* complex (includes the *artropurpurea*-complex), *E. balsamifera*, and *E. longifolia*. The four groups are also deeply separated by molecular data (Molero et al., 2002), further confirming the different nature of potential host plants. The wide and largely overlapping distributions among each of the three first and most important host groups, together with the beetle's distinct preference for one of the groups, make cladogenetic events associated with host switching less ambiguous in this model system.

Current classification of Crypturgini, with reciprocal monophyly of all genera, implies at least three origins of the Macaronesian species and two origins of *Euphorbia* associations. The aim of this study is to establish a new and well-supported phylogeny to allow a more exact estimation of the rate and direction of island-mainland transitions and host plant switching, and hence, to estimate the relative frequency of allopatric divergence versus host shifts in diversification. The largely overlapping distribution within several guilds of ecologically similar species, and the lack of externally diagnostic characters between most of these (Israelson, 1972), may suggest that not all of the species are reciprocally monophyletic. Recent splits often reveal problems with paraphyletic species, or worse, lineage sorting of ancestral polymorphism or even reticulate evolution under low frequency hybridization. To be able to assess these problems, if they do occur, we included multiple specimens per species, from several different islands, and characters from five different partitions. We used sequences from two genes each from the nuclear and the mitochondrial genomes, and morphological characters, to avoid biased gene tree representation of the species tree, and explored the relative contribution of each partition to the various depths of the combined data phylogeny.

Material and Methods

Taxon Sampling

We collected and sequenced 92 specimens, including all Macaronesian and African species of *Aphanarthrum* and *Coleobothrus* (Table 1), two species each of the outgroups *Crypturgus* and *Cisurgus*, and *Dolurgus pumilus* (Mannerheim). Preliminary analyses of 839 EF-1α coding characters, from 12 crypturgine and 45 additional species of Scolytinae and Platypodine, confirmed the previously suggested monophyly of Crypturgini (Wood, 1986), and that *Dolurgus*, and *Crypturgus* plus *Cisurgus*, are nested sets of outgroups suitable for all further analyses of a monophyletic *Aphanarthrum* plus *Coleobothrus*. An identical outgroup topology also resulted from the analysis of 316 aligned characters from the large ribosomal subunit (28S) for 23 crypturgine taxa and 3 scolytine outgroups.

Collected specimens were compared to the original descriptions or type material in the Natural History Museum, London, and Zoologische Staatssammlung, Munich. Based on these comparisons, three species were treated as synonyms as follows: *A. saturatum* Peyerimhoff and *A. duongi* Villiers = *A. mairei* Peyerimhoff, and *A. mododi* Paulian & Villiers = *A. armatum* Wollaston. Also *Cisurgus resiniferae* Peyerimhoff was treated as a synonym of *Ci. occidentalis* Payerimhoff. Three *Aphanarthrum* species described from India (Wood, 1988) were

TABLE 1. Species and populations included in the phylogenetic analyses. Samples not collected by the senior author are indicated by collectors name (in parentheses).

Species	Locality	Host
Outgroups:		
Dolurgus pumilus	USA: AK, Juneau 24.6.2002 (M. Schultz)	*Pinus sitchensis*
Crypturgus borealis	USA: NH, Mt. Monadnock, 26.5.1998	*Picea* sp.
Cr. hispidulus	Norway: Nordland, Svenningdal, 26.7.2001	*Picea* sp.
Cisurgus occidentalis	Morocco: S Agadir, Massa, 16.4.2002	*E. echinus*
Ci. occidentalis	Morocco: N Agadir, Cap Rhir, 17.4.2002	*E. echinus*
Ci. occidentalis	Morocco: S Agadir, Bouzacarne, 20.4.2002	*E. echinus*
Ci. wollastoni	CI: Tenerife, Punta Teno, 6.8.1998 (L. Kirkendall)	*E. canariense*
Ci. wollastoni	CI: El Hierro, Timijiraque, 7.3.2002	*E. canariense*
Ci. wollastoni	CI: La Palma, Punta del Eden, 4.3.2002	*E. canariense*
Ingroup:		
Aphanarthrum aeoni	CI: La Gomera, Alajero, 10.3.2002	*E. lamarckii*
A. aeoni	CI: La Gomera, El Cercado, 10.3.2002	*E. lamarckii*
A. affine	CI: Lanzarote, Mirador de Haria, 27.2.2002	*E. regis-jubae*
A. affine	Morocco: N Agadir, Cap Rhir, 17.4.2002	*E. regis-jubae*
A. affine	Morocco: S Agadir, Sidi Ifni, 20.4.2002	*E. regis-jubae*
A. affine	CI: Gran Canaria, Galdar, 22.2.2002	*E. regis-jubae*
A. affine	CI: Fuerteventura, Lajares, 26.2.2002	*E. regis-jubae*
A. armatum	CI: Lanzarote, Playa Blanca, 27.2.2002	*E. balsamifera*
A. bicinctum (*vestitum*)	CI: Tenerife, Tamaimo, 11.8.1999 (K. Harkestad)	*E. lamarckii*
A. bicinctum (*vestitum*)	CI: La Gomera, Alajero, 14.8.1999 (K. Harkestad)	*E. lamarckii*
A. bicinctum (*vestitum*)	CI: La Gomera, San Sebastian, 8.3.2002	*E. lamarckii*
A. bicinctum (*obsitum*)	CI: Gran Canaria, Galdar, 22.2.2002	*E. regis-jubae*
A. bicinctum (*bicinctum*)	Morocco: N Agadir, Cap Rhir, 17.4.2002	*E. regis-jubae*
A. bicinctum (*bicinctum*)	CI: Lanzarote, Mirador de Haria, 27.2.2002	*E. regis-jubae*
A. bicinctum (*bicinctum*)	CI: Fuerteventura, Lajares, 26.2.2002	*E. regis-jubae*
A. bicolor	CI: El Hierro, La Dehesa, 6.3.2002	*E. lamarckii*
A. bicolor	CI: La Palma, El Time, 1.3.2002	*E. lamarckii*
A. bicolor	CI: La Gomera, Vallehermoso, 9.3.2002	*E. lamarckii*
A. bicolor	CI: Tenerife, San Marcos, 15.2.2002	*E. lamarckii*
A. bicolor	CI: Tenerife, Bco. Balbyo, 12.3.2002	*E. lamarckii*
A. bicolor	Ma: Porto Santo, 7.9.2002	*E. piscatoria*
A. bicolor	Ma: Madeira, Pico das Furnas, 8.9.2002	*E. piscatoria*
A. canariense (*canariense*)	CI: Tenerife, Los Cristianos, 16.2.2002	*E. canariense*
A. canariense (*canariense*)	CI: Fuerteventura, Cofete, 25.2.2002	*E. canariense*
A. canariense (*canariense*)	CI: La Gomera, Bco. Rincon	*E. canariense*
A. canariense (*canariense*)	CI: El Hierro, Timijiraque, 7.3.2002	*E. canariense*
A. canariense (*neglectum*)	CI: La Palma, Playa de Nogales, 28.2.2002	*E. canariense*
A. canariense (*neglectum*)	CI: La Palma, Punta del Eden, 4.3.2002	*E. canariense*
A. canescens (*canescens*)	CI: La Gomera, Airport, Playa Santiago, 10.3.2002	*E. balsamifera*
A. canescens (*canescens*)	CI: La Gomera, Bco. Tapachuga, 11.3.2002	*E. balsamifera*
A. canescens (*canescens*)	CI: La Gomera, Punta Llana, 11.3.2002	*E. balsamifera*
A. canescens (*polyspiniger*)	CI: Tenerife, El Escobonal, 12.8.1999 (K. Harkestad)	*E. balsamifera*
A. canescens (*polyspiniger*)	CI: Tenerife, Playa de la Gaviota, 12.3.2002	*E. balsamifera*
A. canescens (*polyspiniger*)	CI: Fuerteventura, La Matilla, 26.2.2002	*E. balsamifera*
A. euphorbiae	Ma: Madeira, Chao de Laoros, 5.9.2002	*E. longifolia*
A. euphorbiae	Ma: Madeira, Encumeada, 5.9.2002	*E. longifolia*
A. euphorbiae	Ma: Madeira, 1998	FIT trap
A. glabrum (*glabrum*)	CI: El Hierro, Ermita Virgen, 6.3.2002	*E. lamarckii*
A. glabrum (*glabrum*)	CI: La Gomera, Alojera, 9.3.2002	*E. lamarckii*
A. glabrum (*glabrum*)	CI: Tenerife, Bco. Balbyo, 12.3.2002	*E. lamarckii*
A. glabrum (*glabrum*)	CI: Tenerife, Los Corrizales, 13.3.2002	*E. artropurpurea*
A. glabrum (*nudum*)	CI: La Palma, Brena Baja, 2.3.2002	*E. lamarckii*
A. glabrum (*nudum*)	CI: La Palma, El Faro, 2.4.2002	*E. lamarckii*
A. hesperidum	CV: Sao Vicente, Monte Verde, 7.10.2002	*E. tuckeyana*
A. hesperidum	CV: Santo Antao, Cova de Paul, 10.10.2002	*E. tuckeyana*
A. hesperidum	CV: Santo Antao, Rabo Curto, 9.10.2002	*E. tuckeyana*
A. jubae (*jubae*)	CI: Fuerteventura, Pajara, 24.2.2002	*E. balsamifera*
A. jubae (*jubae*)	CI: Lanzarote, Playa Blanca, 27.2.2002	*E. balsamifera*
A. jubae (*jubae*)	CI: Tenerife, Roque de las Bodegas, 14.2.2002	*E. balsamifera*
A. jubae (*tuberculatum*)	CI: El Hierro, Faro de Orchilla, 6.3.2002	*E. balsamifera*
A. jubae (*tuberculatum*)	CI: El Hierro, Charco Manso, 6.3.2002	*E. balsamifera*
A. mairei	Morocco: Massa, S Agadir, 16.4.2002	*E. echinus*
A. mairei	Morocco: Cap Rhir, N Agadir, 17.4.2002	*E. echinus*
A. mairei	CI: Fuerteventura, Jandia, Gran Valle, 25.2.2002	*E. handiense*
A. orientalis	Uganda: Fort Portal, 1.7.1998	*E. teke*
A. orientalis	Uganda: Queen Elisabeth NP, 25.6.1998	*E. teke*

(Continued on next page)

TABLE 1. Continued

Species	Locality	Host
A. piscatorium	Ma: Madeira, Pico das Furnas, 8.9.2002	*E. piscatoria*
A. piscatorium	Ma: Porto Santo, 7.9.2002	*E. piscatoria*
A. piscatorium	CI: Gran Canaria, 4 km S Galdar, 22.2.2002s	*E. regis-jubae*
A. piscatorium	CI: Tenerife, Bco. Balbyo, 12.3.2002	*E. lamarckii*
A. piscatorium	CI: La Gomera, Vallehermoso, 9.3.2002	*E. lamarckii*
A. piscatorium	CI: El Hierro, La Guancha, 6.3.2002	*E. lamarckii*
A. piscatorium	CI: La Palma, Mazo, 4.3.2002	*E. lamarckii*
A. pygmeum	CI: Gran Canaria, 3 km S Tasarte, 20.2.2002	*E. canariense*
A. pygmeum	CI: Tenerife, Los Cristianos, 16.2.2002	*E. canariense*
A. pygmeum	CI: La Gomera, Bco. Rincon, 8.3.2002	*E. canariense*
A. pygmeum	CI: La Palma, Playa de Nogales, 28.2.2002	*E. canariense*
A. pygmeum	CI: El Hierro, Timijiraque, 7.3.2002	*E. canariense*
A. subglabrum	CI: La Palma, El Time, 1.3.2002	*E. lamarckii*
A. subglabrum	CI: La Palma, Punta del Eden, 4.3.2002	*E. lamarckii*
A. subglabrum	CI: La Palma, Brena Baja, 2.3.2002	*E. lamarckii*
A. wollastoni	CI: La Gomera, Bco. Rincon, 8.3.2002	*E. canariense*
A. wollastoni	CI: La Gomera, El Cercado, 10.3.2002	*E. lamarckii*
A. wollastoni	CI: La Gomera, Playa de la Caleta, 9.3.2002	*E. lamarckii*
A. wollastoni	CI: La Gomera, Valle Gran Rey, 10.3.2002	*E. lamarckii*
Coleobothrus allaudi	Morocco: Cap Rhir, N Agadir, 17.4.2002	*E. echinus*
Co. allaudi	Morocco: Bouzacarne, 20.4.2002	*E. echinus*
Co. germeauxi	Uganda: Fort Portal, 1.7.1998	*E. teke*
Co. germeauxi	Uganda: Queen Elisabeth NP, 25.6.1998	*E. teke*
Co. luridus	CI: Gran Canaria, 3 km S Tasarte, 20.2.2002	*E. canariense*
Co. luridus	CI: Tenerife, Punta Teno, 6.8.1998 (L. Kirkendall)	*E. canariense*
Co. luridus	CI: La Gomera, Bco. Rincon, 8.3.2002	*E. canariense*
Co. luridus	CI: La Palma, El Time, 3.3.2002	*E. canariense*

not available for DNA extraction, but we note that these species are very similar to the two Ugandan species included. Multiple specimens per species were included for all species possible, with coverage of nearly all islands within their known range. Current studies on the population genetics of the Canary Island species confirmed that our samples were unbiased with respect to the populations included here.

Character Sampling

Morphological and behavioral characters.—Twenty-eight morphological characters were coded and included in the parsimony analyses (Appendices 1, 2). We also included *Deropria elongata* (Eggers) in this matrix to test the monophyly of all other crypturgine genera. Host plant use and geographical distribution were based on literature records (Israelson, 1972, 1976, 1979) and corrected and augmented after intensive sampling throughout the Macaronesian islands (Jordal, in preparation). Previous records from multiple *Euphorbia* groups were probably based on rare stray specimens on the less preferred host, as documented during recent sampling. Geographical distributions of the beetles were largely congruent with literature records, with only minor corrections needed.

DNA amplification and sequencing.—We sequenced portions of two mitochondrial and two nuclear genes, using mostly the same primer for sequencing as for PCR amplification. Mitochondrial cytochrome oxidase I was amplified with the primers s1718 5′-GGA GGA TTT GGA AAT TGA TTA GTT CC-3′ (forward), a2237 5′-CCG AAT GCT TCT TTT TTA CCT CTT TCT TG-3′ (reverse), or a2411 5′-GCT AAT CAT CTA AAA ACT TTA ATT CCW GTW G-3′ (reverse) (Normark et al., 1999). PCR reactions were performed in a 25-μL volume containing 0.2 μM of each primer, 0.25 mM of each dNTP, 0.5 unit of Bioline BioTaq DNA polymerase, 1× buffer with $MgCl_2$ to a final concentration of 1.5 mM. Typical PCR cycles consisted of 90 s initial denaturing at 94°C, followed by 38 cycles of 94°C for 30 s, 46°C annealing for 60 s, and 72°C extension for 60 s, followed by a final extension for 7 min.

The mitochondrial large ribosomal subunit, 16S, was amplified under the same PCR conditions, using the primers 16S 5′-TTT AAT CCA ACA TCG AGG-3′ (forward) and 16S 5′-CGC CTG TTT AAC AAA AAC AT-3′ (reverse) (Cognato and Vogler, 2001).

The nuclear gene encoding Elongation Factor 1α was amplified with the primers efs149 5′-ATC GAG AAG TTC GAG AAG GAG GCY CAR GAA ATG GG-3′ (forward) and efa1043 5′-GTA TAT CCA TTG GAA ATT TGA CCN GGR TGR TT-3′ (reverse) (Normark et al., 1999). We used a touchdown profile consisting of 43 cycles where all cycles had 72°C extension for 40 s and 94°C denaturing for 30 s, with the first 16 cycles having decreasing annealing temperature from 60°C to 47°C by 2°C (last 1°C) every second cycle, with the final 27 cycles at 46°C for 60 s.

The nuclear encoded *Enolase* gene (2-phospho-D-glycerate hydrolase) (Tracy and Hedges, 2000; Farrell et al., 2001) was amplified with nested PCRs using initially the primers ens57c 5′-GAC TCC CGT GGN AAC CCC ACN GTG GAG GT-3′ (forward) and ena886 5′-CCA GTC RTC YTG RTC RAA XGG-3′ (reverse), with the latter replaced by ena780 5′-TCT TGA AGT CCA AAT CGT A-3′ (reverse, sequencing primer) in the nested

PCR. Initial reactions were performed under similar conditions as for the mtDNA genes, but with 2.5 mM $MgCl_2$ concentration and 44°C annealing temperature. The nested PCR reactions used 1.5 mM $MgCl_2$ concentration, with 1.0 μg bovine serum albumin added. Cycles consisted of 90 s initial denaturing at 94°C, followed by 40 cycles where all cycles had 72°C extension for 60 s and 94°C denaturing for 30 s, with the first 16 cycles having increasing annealing temperature from 40°C to 55°C by 2°C (last 1°C) every second cycle, with the final 12 cycles at 55°C for 60 s. Purified PCR products were sequenced using a standard protocol for big-dye version 3 (Perkin Elmer).

Sequences were assembled and edited with the Lasergene software (DNASTAR, Inc.). There are multiple copies known for EF-1α and *Enolase* and we used the single-intron copy (C1) of EF-1α and intron free copy of *Enolase* (Eno1) in this study (Farrell et al., 2001; Jordal, 2002). Sequences can be downloaded from GenBank under the accession numbers AY500900 to AY500991 (EF-1α), AY514121 to AY514211 (*Enolase*), AY514212 to AY514303 (16S), AY514904 to AY514995 (COI).

Phylogenetic Analyses

Alignments.—For *Enolase* and the EF-1α coding region, alignments were unambiguous due to the lack of any insertions or deletions in the coding region. Three COI sequences had indels (in multiples of 3 bp), but none of these putative pseudogenes had stop codons and were cautiously included; these sequences grouped together with conspecific indel free sequences (*A. aeoni* El Cercado; *Co. allaudi*, Morocco). Sequences of 16S and the EF-1α intron were aligned in ClustalX (Thompson et al., 1997) under a wide range of parameters to assess alignment stability (gap cost 2, 4, 6, 8, 12 16; gap extension equal or half the cost of gap opening; transitions weighted equally or half the cost of transversions). Preliminary alignments were chosen based on their topological congruence with the topology resulting from the elision matrices which combines all alignments from each partition into one matrix, providing a gradual down weighting of ambiguous sites (Wheeler et al., 1995; Lee, 2001). Final alignments were produced under gap opening to extension cost of 6:3, with transitions half the cost of transversions, resulting in 68 ambiguous sites for 16S, and 54 ambiguous sites in the much more variable EF-1α intron. These alignments were also the most congruent with the unambiguously aligned portions of the EF-1α exon and 16S, respectively. A secondary structure model for insects (Buckley et al., 2000) was used to assist in editing the final 16S alignment. All indel ambiguous regions were located in the terminal loops of helices 68 (9 bp), 75 (48 bp, including all but 5 bp of the helix), 81′ (8 bp), and 84′ (4 bp). Various alignments (or inclusion/exclusion of ambiguous sites) had no influence in the analyses of combined data. Neither did gaps as a fifth character change the combined data topology. Hence we scored gaps as missing characters in the parsimony analyses, to make comparisons to likelihood analyses more realistic.

Sequence pruning.—To assess the relative importance of sampling multiple sequences per species, and to increase the speed in likelihood analyses, we pruned one sequence from each of the species which had an excess of sequences before each heuristic parsimony search. The phylogenetic depth and number of node switches was noted for each partition, genome, and the combined data, on their respective strict consensus trees.

Tree searching.—Paup* 4.0 (Swofford, 1999) was used in most phylogenetic analyses (the complete data matrix and Figure 1 can be downloaded from TreeBase, accession number SN1709, http://www.treebase.org/treebase/). Under the maximum-parsimony (MP) criterion, we performed 500 random addition replicates of heuristic searches for each of the data partitions and the combined matrix. Bootstrap support for individual nodes was assessed by 200 bootstrap replicates (Felsenstein, 1985) of 10 random additions per replicate, with uninformative characters excluded. Bremer support was calculated as the number of additional steps necessary to collapse a node (Bremer, 1994), using TreeRot v.2 (Sorenson, 1999). As DeBry (2001) recently pointed out, the significance of this metric is dependent on the number of informative characters assigned to each node. We have used minimum branch lengths as a proxy estimate of informative characters, and applied DeBry's Table 3 to assess approximate significance of each Bremer support value.

Maximum-likelihood (ML) analyses were performed on the pruned data matrix only (54 terminals), with parameters estimated through the application of Modeltest 3.06 (Posada and Crandall, 1998). This application tests a nested set of substitution models using either the log-likelihood ratio test (LRT) or the Akaike information criterion (AIC). When the two statistics disagreed, we applied the AIC criterion to select the best model for each partition, or for all data combined. Ten random additions were used in each of the final likelihood searches. ML branch lengths were also optimised with the molecular clock enforced, and likelihood scores were compared to those obtained with no constraints using the LRT.

MrBayes 3.1 (Ronquist and Huelsenbeck, 2003) was used in the Bayesian analyses of the combined data, with a GTR +Γ+ model of substitution (selected by Modeltest) fitted to each molecular partition, and a rate variable model fitted to the morphological data. Parameters were estimated on each of the genes and morphology (5 partitions), or to each coding position, EF-1α intron, 16S, and morphology (12 partitions). Initial searches of 500,000 generations were used to assess convergence in likelihood values obtained from four running Markov chains, sampling trees every 500 generations. Trees from the first 200,000 generations were discarded in the final analysis of 1,000,000 generations, as a 'burn-in' before stability in likelihood values was achieved.

Incongruence.—The incongruence length difference test (ILD) (Farris et al., 1995) has been frequently applied to assess combinability of data partitions. However, both empirical (Yoder et al., 2001) and simulated data (Barker and Lutzoni, 2002, and references therein) have

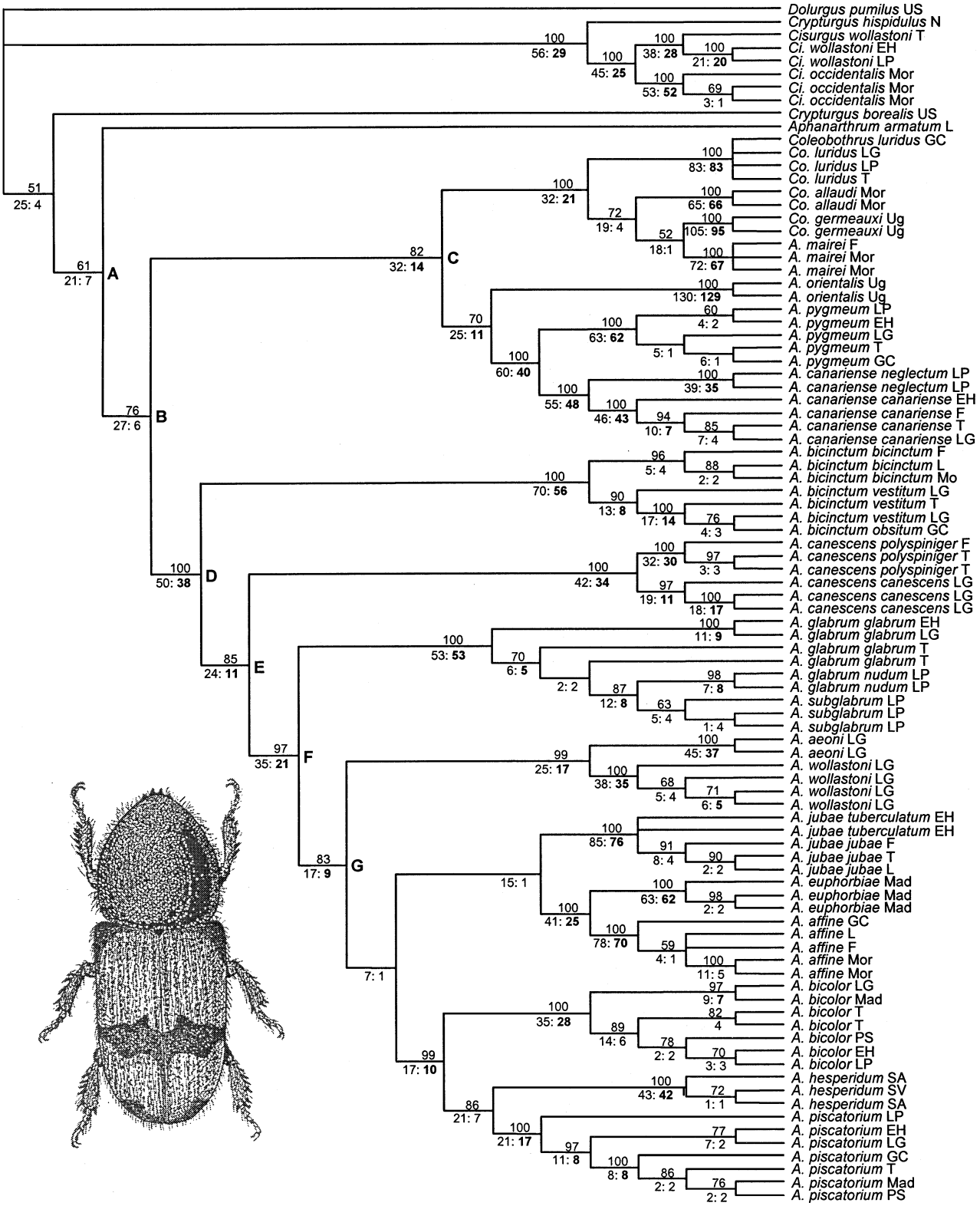

FIGURE 1. Strict consensus of 48 most parsimonious trees of length 5014 steps (CI = 0.35, RI = 0.82) resulting from the unweighted analysis of all data. Bootstrap support values are shown above the nodes, with minimum branch length and Bremer support values below (ml:BS, 'significant' support shown in bold). Capital letters in bold indicate nodes discussed in the text. Geographic origin of each terminal is indicated by capitals: SV, Sao Vicente; SA, Santo Antao (Cape Verde); Mad, Madeira; PS, Porto Santo (Madeira); EH, El Hierro; LP, La Palma; LG, La Gomera; T, Tenerife; GC, Gran Canaria; F, Fuerteventura; L, Lanzarote (Canary Islands); Mor, Morocco; Ug, Uganda, US, United States of America; N, Norway. The line drawing illustrates *Aphanarthrum armatum*.

demonstrated the potential failure of this test to allow combination of data when they should be. Potential negative effects from combining statistically incongruent gene partitions can now be tested in a Bayesian framework, by analyzing all data under a uniform model of substitution versus a mixed model where each partition fits their own model. In a parsimony framework, incongruence can further be explored by accumulating data from an increasing number of partitions, aiming towards a predictive hierarchy; that is, it remains stable with the addition of more data (Nixon and Carpenter, 1996). Incongruence may then be quantified by the level of node support in conflicting nodes across different partitions. A node-specific approach to detect conflicting taxa and partitions simultaneously was implemented through the calculation of partitioned Bremer support (Baker and DeSalle, 1997; Baker et al., 1998), using TreeRot 2.0 (Sorenson, 1999). We estimated the relative contribution from each data partition to the combined analysis, using default settings with 20 random addition sequences for each node in each of the partitions.

The overall support from different data partitions may also be asymmetric, such that one partition perform significantly worse on the topologies resulting from the other partitions, but not vice versa. We applied the Wilcoxon sign rank test (parsimony) in Paup* to test this asymmetry, with the critical value Bonferroni corrected for 49 consecutive comparisons among partitions. For comparison, we also applied the Shimodaira-Hasegawa test (Shimodaira and Hasegawa, 1999) on the molecular partitions under the likelihood criterion, using 1000 bootstrap pseudoreplicates.

Lineages that were topologically unstable across different analyses were also examined for potential deviation in their base frequencies. We used the base heterogeneity test as implemented in PAUP*, and performed tests on all sites, and on informative sites only, to detect confounding opposite bias in invariant sites (Buckley et al, 2001). Log-determinant distance analysis (Lockhart et al., 1994) is the only tree reconstruction method that is specifically designed to avoid attraction of unrelated lineages with similar base composition. We used this method to explore the possibility that biased base frequencies had any effect on conflicting or unstable nodes.

Character evolution.—MacClade 3.04 (Maddison and Maddison, 1992) was used for editing matrices, calculating substitution frequencies, and tracing character transformations over the most parsimonious or likely tree topologies from the combined data analyses. We applied the "chart state changes" option to count the minimum and maximum transformations in host plant use and island-mainland distribution under ACCTRAN and DELTRAN optimisations. Although accelerated character transformations may be preferred due to conforming more closely to primary homology statements (by implying reversals rather than parallel evolution of characters), there is no particular reason to expect one or the other for labile island-mainland transitions or host switching.

Ancestral character states were also reconstructed using maximum likelihood, incorporated in the computer program Discrete (Pagel, 1994, 1999). The likelihood approach has the advantage (over parsimony methods) of estimating the probability of each ancestral character state, for each ancestor at a time, taking character frequencies and branch lengths (but not prior node estimates) into account.

RESULTS

Combined Analysis of All Data

The parsimony analysis of all 2688 characters (2660 aligned base pairs, 28 morphological characters; Table 2), resulted in two trees differing only in the placement of the two El Hierro specimens of *A. jubae* (Fig. 1). *Aphanarthrum* plus *Coleobothrus* were monophyletic in this analysis, with the latter nested within *Aphanarthrum*. The predefined hypothesis on reciprocal monophyly of *Aphanarthrum* and *Coleobothrus* added 55 steps to the tree length, and was firmly rejected by the Kishino-Hasegawa and Wilcoxon tests ($t = 4.01$, $z = -3.99$, $P < 0.001$).

All but one of the species with multiple individuals sampled were monophyletic, and most sister species were clearly defined. *Aphanarthrum glabrum* was paraphyletic with respect to *A. subglabrum* which formed the sistergroup to the La Palma subspecies of *A. glabrum* (*nudum*). Among the higher level relationships, only the positions of *A. jubae*, and the internal arrangement of two *Coleobothrus* species with respect to *A. mairei*, were uncertain as shown by minimum Bremer and bootstrap supports. *Aphanarthrum armatum* was the basal lineage, with all species associated with succulent *Euphorbia* in one

TABLE 2. Tree statistics from the parsimony analyses of the various data partitions (54 taxa). Informative characters for the ingroup only (*Aphanarthrum* and *Coleobothrus*) are given in parenthesis. Incongruence refers to the number of node shifts compared to the combined data tree topology.

Partition	Characters	Informative	Treelength	CI	RI	No. of trees	Incongruence
All data	2588	917 (860)	4116	0.37	0.76	2	—
Molecular	2560	893 (840)	4050	0.37	0.75	1	0
MtDNA	946	367 (353)	1969	0.33	0.72	3	3
NucDNA	1614	526 (487)	2057	0.41	0.79	1	2
EF-1α	927	272 (252)	858	0.48	0.82	16	4
Enolase	687	254 (235)	1182	0.36	0.77	9	3
16S	454	154 (142)	618	0.43	0.81	8	1
COI	492	213 (211)	1320	0.28	0.67	1	14
Morphology	28	24 (20)	55	0.76	0.95	294	7

well supported clade (C), and all Macaronesian species breeding in arborescent euphorbs making up a highly supported sistergroup (clade D). All clades that received maximum bootstrap support had also significantly high Bremer support, with only two significantly supported nodes having bootstrap values below 80%. There was a generally high correlation between the two node support indices for the combined data.

Outgroup manipulation.—Because the *Dolurgus pumilus* sequences were relatively more diverged from the ingroup than the *Crypturgus* and *Cisurgus* sequences, we excluded this taxon in the remaining analyses. We also excluded *Cr. borealis* because this taxon had the entire *Enolase* partition missing. The exclusion of these taxa did not change the topology resulting from the combined parsimony analysis (or any of the molecular, nucDNA, or EF-1α topologies). Support for the most basal nodes (A to G) increased after exclusion of *D. pumilus* and *Cr. borealis*, further suggesting that the analyses should be improved by using more narrowly defined outgroups.

Sequence Pruning

Rather few contradicting changes occurred with increasingly pruned data matrices in each of the gene partitions (Table 3); all data combined, nucDNA, and the 16S data did not produce any contradictory nodes with increased pruning of sequences. Most changes occurred in the two genes with highest substitution rates, *Enolase* and COI (cf. Fig. 4), but these changes were unstable and sometimes reversed to the original topology in the next step of pruning. Overall, there were only marginal effects of pruned matrices in each gene partition, down to three sequences per species, and all bootstrap supported nodes (>50%) were recovered when only one or two sequences per species or subspecies were retained. Hence, we used two sequences per species or subspecies in all further comparative analyses of the various gene partitions.

Separate Analyses—Parsimony

Morphology.—The analysis of morphological characters resulted in 294 equally parsimonious trees of length 55 steps (Table 2). The inclusion of all outgroups and *Deropria elongata* resulted in an identical topology for the ingroup (*Aphanarthrum* plus *Coleobothrus*), and confirmed the distant relationship of *D. elongata* in Crypturgini (Fig. 2) as previously suggested (Wood, 1986).

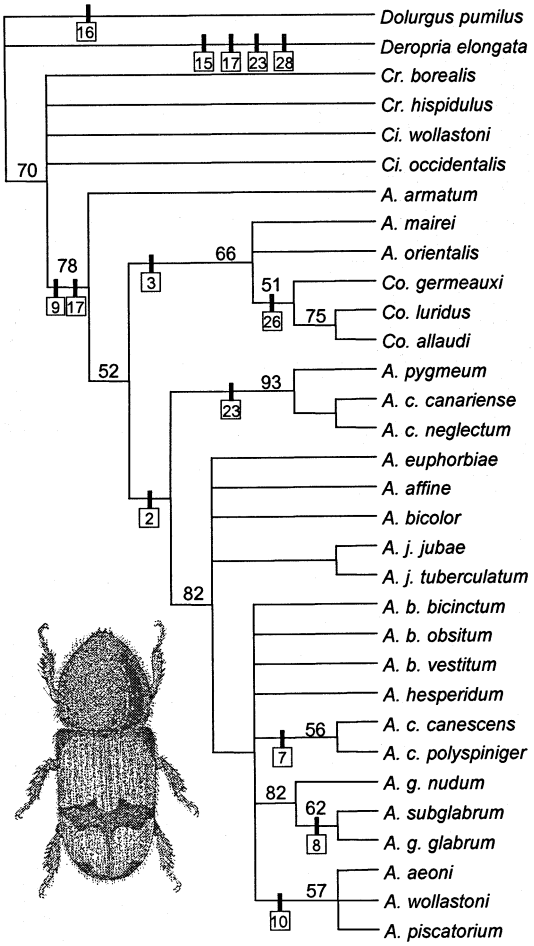

FIGURE 2. Strict consensus of 432 MP trees of length 73 steps each (CI = 0.71, RI = 0.87), resulting from the parsimony analysis of morphological characters. *Deropria elongata* was included in this analysis to test the monophyly of the ingroup. Bootstrap support values are shown above the nodes. Black bars marked by numbers in squares indicate characters with synapomorphic character states (except autapomorphies also shown in *Deropria* and *Dolurgus*, see Appendices 1 and 2). The line drawing illustrates *Aphanarthrum armatum*.

Our morphological data did not enable us to distinguish *Crypturgus* and*Cisurgus*, but grouped *Aphanarthrum* and *Coleobothrus* with high support (including two uniquely derived characters), with the latter genus nested within *Aphanarthrum*. Topological congruence with the combined data tree was further exemplified by the basal placement of *A. armatum*, the close relationship between *A. mairei*, *A. orientalis*, and *Coleobothrus*, and the

TABLE 3. Number of shifted interspecific nodes (total/deep only) after deletion of sequences. Only those taxa with excess sequences had sequences deleted, one per species (or subspecies) before each search. The first column shows the number of sequences retained per species. Deep nodes were defined as clades A to D (see Fig. 1).

Sequences	All data	nDNA	mtDNA	EF-1α	Enolase	16S	COI	Total
5	0/0	0/0	0/0	0/0	0/0	0/0	1/1	1/1
4	0/0	0/0	1/0	0/0	0/0	0/0	3/2	4/2
3	0/0	0/0	0/0	0/0	2/0	0/0	2/1	4/1
2	0/0	0/0	1/0	0/0	4/2	0/0	0/0	5/2
1	0/0	0/0	3/1	2/0	1/0	0/0	3/1	10/2

well-supported clade of Macaronesian shrub feeders (clade D). The lack of resolution in the latter clade was apparently due to the low number of informative characters at this topological level.

Molecules.—The four molecular partitions contributed from 154 (16S) to 272 (EF-1α) informative characters (Table 2), and produced partly resolved topologies for each partition, except for the fully resolved COI data (Fig. 3). Although several discrepancies appeared among the four consensus trees (Table 2), only three topological differences were supported by a bootstrap support above 70%, and only barely so (EF-1α: *A. orientalis*, *A. bicinctum*, and *A. canescens*. COI: *A. bicolor*). Major differences of low or no support were mostly observed in the COI topology where the basal structure was strikingly absent when compared to the combined data tree. Other noteworthy differences were seen in *Enolase*, which differed by the changed position of the otherwise basal *A. armatum* and the more basal position of *A. orientalis*, and in EF-1α, which differed in their basal position of *A. euphorbiae* plus *A. affine* in clade D.

The *Enolase* and COI genes had also significant biases in base composition across taxa, caused by nonstationarity in third positions (Table 4). EF-1α had a similarly significant bias in third position, but all parsimony informative sites combined were not significantly biased for this gene. For EF-1α and *Enolase*, the outgroups and *A. armatum* had generally intermediate base frequencies, and all succulent feeding taxa, and all remaining shrub feeding taxa, had divergent base frequencies that were more similar within each of these clades. The pattern was less clear in COI, and varied much more between close relatives. We further noticed that *A. orientalis* had frequencies more similar to the *Coleobothrus* clade than to the *A. canariense* clade for the two nuclear genes. However, a LogDet distance analysis of the nuclear DNA partitions did not remove *A. orientalis* from the *Coleobothrus* clade. One exception to the within-clade correlation in base composition was found for the two basal shrub feeding taxa in the EF-1α analysis, *A. affine* and *A. euphorbiae*, which had EF-1α base composition similar to the succulent taxa and the outgroups. LogDet analysis of these data placed *A. bicinctum* as the most basal taxon in this clade, consistent with most other analyses.

We applied the consistency index (CI) and the retention index (RI) as measures of homoplasy (Table 2). These two correlated well with each other, and also reflected the different distributions of substitution frequencies in the four genes (Fig. 4) such that those genes with the fewest substitutions per character also had the highest consistency and synapomorphy (RI) indices.

We also used the partitioned Bremer support index (PBS) to measure the relative support of each partition to the various depths of the phylogeny (Baker and DeSalle, 1997). All four genes contributed positively to the combined data tree, although COI contributed consistently less at higher level nodes (Fig. 5, Table 5). This trend was particularly clear when PBS was corrected for minimum tree lengths for each partition.

Separate Analyses—Maximum Likelihood

All models selected by Modeltest were derivatives of a general time reversible model, where all assumed at least two different transition rates and two different transversion rates (Table 6). Major differences were found between the nuclear and mitochondrial genes in the guanine base frequencies and guanine-thymine substitution frequencies. Also, the shape of the gamma-distributed substitution rates was higher in the *Enolase* partition, indicating a more uniform distribution of character changes for variable sites (cf. Fig. 4). When all sites were considered (including invariable sites), all gene partitions had nearly identical shape parameters (Table 4).

Overall, the ML analyses confirmed our results from the parsimony analyses, with a few noteworthy exceptions (Fig. 6). In both of the *Enolase* and 16S partitions, *A. canescens* was moved to the basal position in clade D, and the position of *A. orientalis* was changed. Compared to the combined data topology, several improvements occurred in the COI partition, with monophyly of *A. euphorbiae* and *A. affine*, and of all *Coleobothrus* species (including *A. mairei*). *Aphanarthrum armatum* moved closer to the base of the COI tree, and moved to the 'correct' basal position in the *Enolase* tree. However, apparently worse results also occurred, with paraphyly of *A. c. canariense* (EF-1α), *A. wollastoni* (*Enolase*), and the otherwise monophyletic *A. wollastoni* plus *A. aeoni* (COI). Overall, the number of topological differences compared to the likelihood analysis of all data combined was higher than for the equivalent parsimony analyses, with 6, 4, 6, and 12 incongruent nodes found in the EF-1α, *Enolase*, 16S, and COI partitions, respectively.

Dynamics of Data Combination

Parsimony.—Combination of characters from the same genome resulted in two consensus trees which only differed from the combined data tree by two (nucDNA) or three (mtDNA) topological positions (Fig. 7). All nodes supported by more than 70% bootstrap support in each of the two genomic partitions were fully congruent with each other and with the combined data tree. Bootstrap support in the nucDNA partition incorporated complementary support from each of the *Enolase* and EF-1α partitions, whereas some support in the mtDNA tree was lost, particularly for clades B and C, compared to the separate 16S analysis. The combination of all molecular partitions (Fig. 8) resulted in an overall increased bootstrap support for the most basal nodes (A to C), as well as some of the more derived clades (E to G). Adding morphological characters did not influence tree topology, or node support, at all.

None of the separate or combined data partitions were incongruent when forced upon the combined data tree (Table 7). Only the COI data were incongruent on the nucDNA topology, whereas all nucDNA partitions were incongruent upon the mtDNA topology. EF-1α had the most congruent topology with only two partitions (mtDNA, COI) being significantly incongruent.

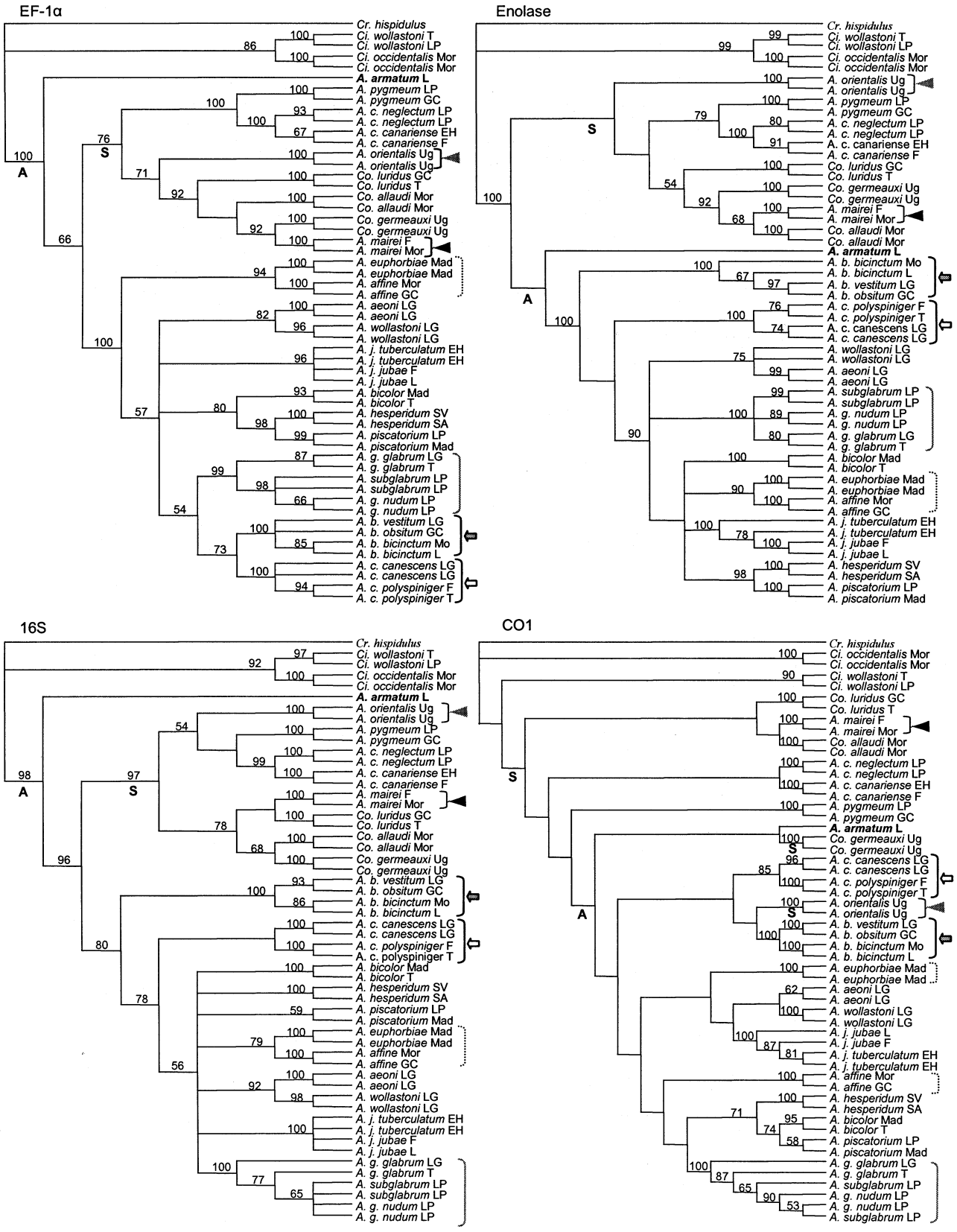

FIGURE 3. Strict consensus trees resulting from the parsimony analyses of each gene partition. Bootstrap support values are shown above the nodes, transitions in habitat selection under ACCTRAN optimisation is marked below nodes (A, arborescent; S, succulent). See Table 2 for further details. Selected taxa that varied in their position across analyses of different partitions are indicated by different symbols and fonts.

TABLE 4. Summary of sequence statistics for the various molecular character partitions. Tests on deviation from base stationarity were performed on all sites, or informative sites only, and Bonferroni corrected for 11 consecutive and nonoverlapping data partitions (critical $P = 0.0045$, underlined). The transition to transversion rate (Ti:Tv) and the shape of the Γ parameter (α) were estimated under a HKY+Γ model of substitution, with invariant sites artificially set to zero for comparisons between partitions.

	EF-1α					Enolase				16S		COI			
	1st	2nd	3rd	Intron	All	1st	2nd	3rd	All	No gaps	All	1st	2nd	3rd	All
Number positions	280	280	280	87	927	229	229	229	687	385	454	164	164	164	492
Variable positions	28	13	183	76	300	51	24	212	287	109	164	41	15	161	217
Informative positions	22	12	169	69	272	36	14	204	254	100	154	40	12	161	213
Base homogeneity (P)	1.00	1.00	0.70	0.11	1.00	1.00	1.00	<u>0.001</u>	<u>0.001</u>	1.00	1.00	1.00	1.00	<u>0.000</u>	0.99
π_A	0.29	0.30	0.26	0.33	0.29	0.32	0.32	0.30	0.32	0.35	0.36	0.28	0.17	0.45	0.30
π_C	0.17	0.24	0.26	0.12	0.22	0.11	0.23	0.24	0.20	0.20	0.19	0.24	0.25	0.19	0.23
π_G	0.39	0.16	0.12	0.13	0.22	0.41	0.15	0.17	0.24	0.11	0.10	0.27	0.16	0.07	0.17
π_T	0.15	0.29	0.36	0.42	0.27	0.16	0.29	0.29	0.24	0.34	0.34	0.21	0.41	0.29	0.30
Base homogeneity (inf. only)	0.99	1.00	<u>0.001</u>	0.03	0.60	0.89	1.00	<u>0.000</u>	<u>0.000</u>	0.99	0.99	1.00	1.00	<u>0.000</u>	<u>0.000</u>
π_A	0.09	0.28	0.19	0.35	0.22	0.24	0.30	0.30	0.29	0.34	0.37	0.11	0.06	0.44	0.35
π_C	0.38	0.38	0.32	0.13	0.29	0.18	0.43	0.25	0.25	0.17	0.16	0.43	0.36	0.19	0.25
π_G	0.23	0.26	0.09	0.10	0.11	0.18	0.13	0.15	0.15	0.09	0.07	0.21	0.17	0.08	0.11
π_T	0.30	0.08	0.40	0.42	0.38	0.40	0.14	0.30	0.31	0.40	0.40	0.24	0.41	0.29	0.29
Ti:Tv ratio	2.66	1.02	4.83	0.85	2.42	3.48	0.55	2.59	2.44	1.75	1.70	17.14	1.22	4.97	2.87
α shape	0.01	0.01	0.73	2.24	0.18	0.15	0.01	1.76	0.21	0.15	0.17	0.11	0.01	1.44	0.25

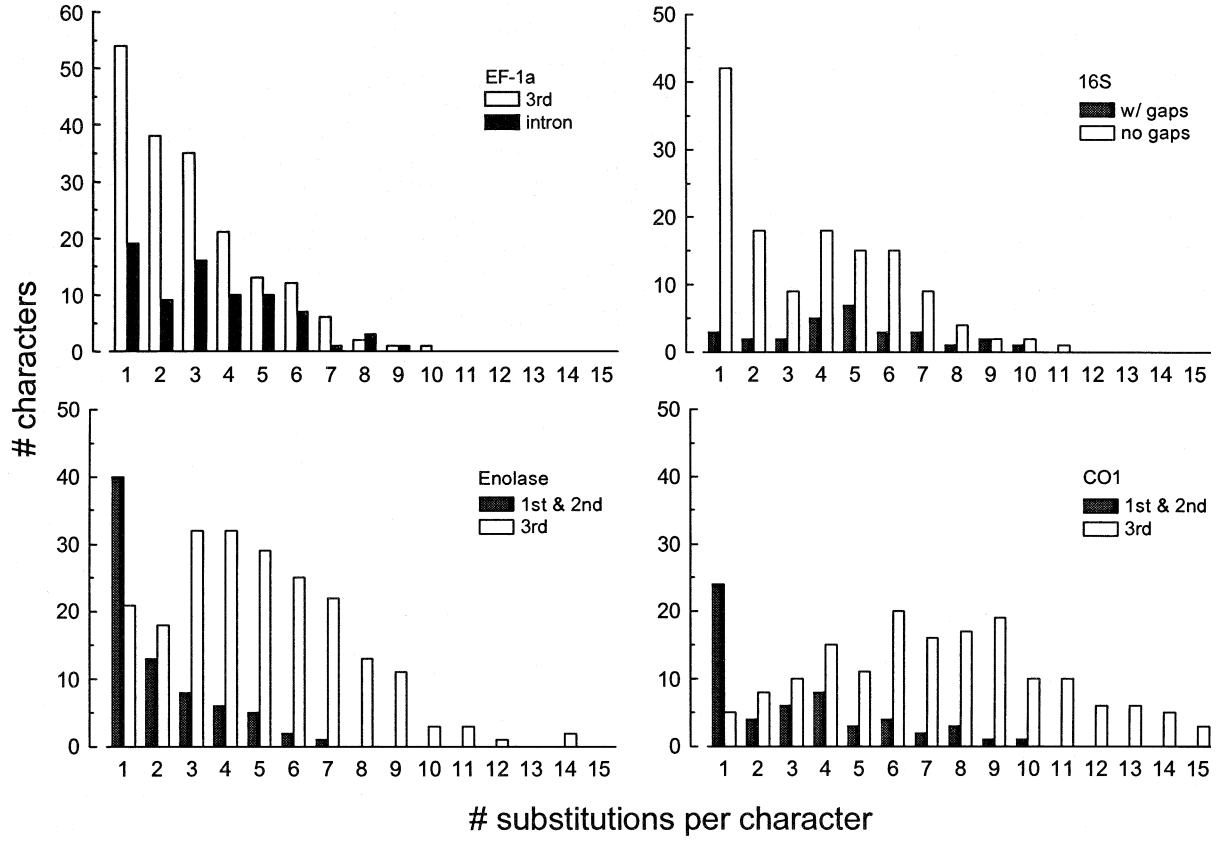

FIGURE 4. Frequency distributions of character changes in the four molecular partitions. Invariable characters are not shown on the graphs. In 16S we distinguished between sites in alignment ambiguous and unambiguous regions.

Conversely, COI was the most incongruent topology, only congruent with the mtDNA data. The two mitochondrial partitions were also less congruent than the two nuclear partitions as measured by the ILD test (COI versus 16S, $P = 0.02$; EF-1α versus *Enolase*, $P = 0.08$; mtDNA versus nucDNA, $P = 0.001$).

Maximum Likelihood.—The combination of all molecular data resulted in one tree which only differed in two node positions compared to the parsimony analysis of combined data; *A. orientalis* was the sister group to *Coleobothrus*, with *Co. allaudi* as the sister species of *A. mairei* (Fig. 8). These nodes had very short internal branches in the likelihood analyses, a possible explanation for the unstable positions of these taxa.

Shimodaira-Hasegawa tests of all molecular partitions (as listed in Table 7) produced quite different results from the Wilcoxon sign rank tests, with six new comparisons significantly incongruent, whereas six comparisons that were incongruent in the parsimony analyses were insignificantly different under likelihood optimization. The most striking difference was observed in comparisons involving 16S data that were congruent with all partitions except COI. With parameters optimized for the COI data, this partition was incongruent when forced upon all other topologies, except mtDNA. Furthermore, the EF-1α topology was also much more incongruent in the SH analyses, now incongruent with all data combined, mtDNA, *Enolase* and COI, but not 16S and nucDNA.

Bayesian analysis.—The implementation of a mixed model that included separate estimation of parameters for each coding position, intron, rDNA, and morphology (12 partitions) produced an identical topology to the ML (molecular data only) and Bayesian analyses under a uniform model (Fig. 8). When models were fitted to each gene and morphological data (5 partitions), three

TABLE 5. Partitioned Bremer support for each of the data partitions, at interspecific, specific, and subspecific levels. PBS was also corrected by the minimum possible length (mL) of each separate partion.

	EF-1α		Enolase		16S		COI		Morphology	
	PBS	/mL	PBS	/mL	PBS	/mL	PBS	/mL	PBS	/mL
Interspecies	102.4	0.119	128.9	0.109	61.6	0.100	10.3	0.008	5.8	0.105
Species	267.0	0.311	299.8	0.253	194.5	0.315	387.0	0.293	7.7	0.140
Subspecies	9.5	0.011	33.5	0.028	39.5	0.064	80.0	0.061	−1.5	−0.027
Total	378.9	0.442	462.2	0.390	295.6	0.478	477.3	0.362	12.0	0.218

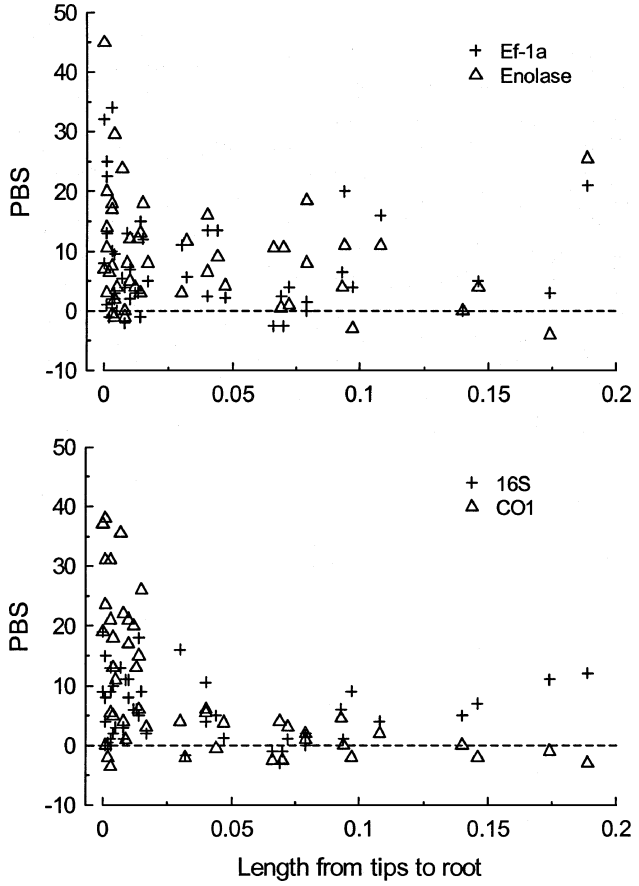

FIGURE 5. Distribution of partitioned Bremer support values in relation to the distance measured from the node tips to the root (all molecular data, clock enforced). Stippled line marks the transition between positive and negative contribution to the combined data topology.

shallow nodes shifted their position within clade G (see Fig. 1), moving *A. affine* and *A. euphorbiae* to the most basal position in that clade, followed by *A. jubae*.

Speciation Rates

Likelihood ratio tests showed a significant departure from a clock-like substitution rate in the combined molecular data as well as for the mtDNA data for 54 taxa. With the outgroups *Cisurgus* and *Crypturgus* pruned from the mtDNA matrix, however, a clock-like rate could not be rejected at the 1% significance level ($2\ln LR = 73.02$, $df = 49$, $P = 0.023$). Further deletion of the three putative COI pseudogene sequences (*A. aeoni* [1], *Co. allaudi* [2]) resulted in even more clock-like rates ($2\ln LR = 64.66$, $df = 46$, $P = 0.04$). Hence, we used the mtDNA data with clock-like rates to estimate distances from the tips to the root of the similarly pruned combined data parsimony tree.

We plotted the cumulative number of lineages of *Aphanarthrum* and *Coleobothrus* on the relative time elapsed since the most recent common ancestor for all species of these genera (Fig. 9). This plot shows a more or less straight line through time, indicating a nearly constant speciation to extinction rate.

Geographical Patterns and Host Plant Use

Island colonization.—Within *Aphanarthrum* and *Coleobothrus* combined, only one unambiguous colonization of the islands was found using the "chart state changes" option in MacClade (accelerated and delayed transformations indifferent). Five unambiguous back colorizations to Africa were found on the MP topology (Fig. 10), and three on the ML topology. The alternative placements of *A. orientalis* and *A. mairei* did not influence these results. However, when *A. armatum* was forced to be monomorphic African, implying recent dispersal to the Canary islands in this species, the numbers of unambiguous back colorizations were reversed to one for each topology, with up to five possible island colorizations (one unambiguous). A maximum of seven transitions were required by the current data, at least three of which were recent intraspecific dispersals.

Likelihood estimates of ancestral distributions (African mainland versus islands) supported the general trends found in the parsimony optimizations. The ancestor of *Aphanarthrum* was equally likely African or island distributed. However, all remaining supraspecies nodes supported island ancestry, involving at least six back colorizations to the mainland (Fig. 11). When predominantly African distributed species were assumed

TABLE 6. Maximum-likelihood parameters estimated by Modeltest and selected by the Akaike information criterion (AIC). All base frequencies are empirical.

Parameter	EF-1α	Enolase	nDNA	16S	COI	mtDNA	Molecular
Model	GTR+I+Γ	GTR+I+Γ	TrN+I+Γ	TIM+I+Γ	TIM+I+Γ	TIM+I+Γ	GTR+I+Γ
−ln	5562.961	6458.344	12209.649	3376.833	5991.715	9543.586	22211.869
*p*A	0.288	0.316	0.300	0.363	0.302	0.331	0.311
*p*C	0.217	0.197	0.209	0.189	0.227	0.209	0.209
*p*G	0.218	0.240	0.227	0.104	0.168	0.138	0.194
*p*T	0.276	0.247	0.264	0.344	0.303	0.322	0.286
r-AC	1.457	1.345	1.000	1.000	1.000	1.000	1.381
r-AG	4.846	3.776	3.680	17.724	22.760	19.417	6.098
r-AT	2.478	0.793	1.000	1.003	1.621	1.564	1.608
r-CG	1.125	1.228	1.000	1.003	1.621	1.564	1.334
r-CT	10.539	6.710	6.031	5.545	13.698	9.858	8.229
r-GT	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Pinv	0.414	0.518	0.534	0.515	0.491	0.526	0.540
α	0.490	1.245	0.947	0.501	0.534	0.600	0.814

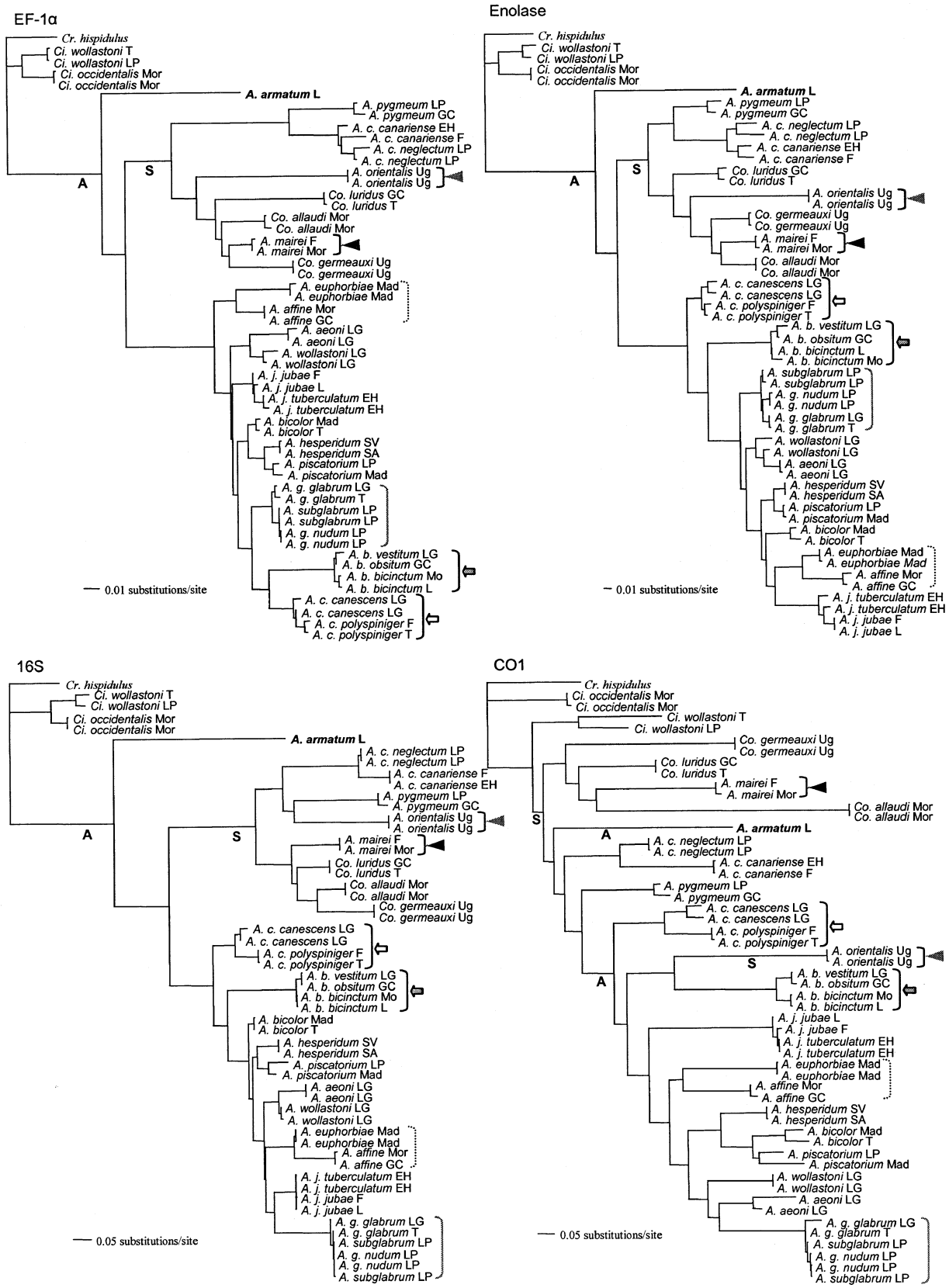

FIGURE 6. The most likely tree topologies estimated from each gene partition, using the best model selected by Modeltest (Table 6). See Figure 3 for further details.

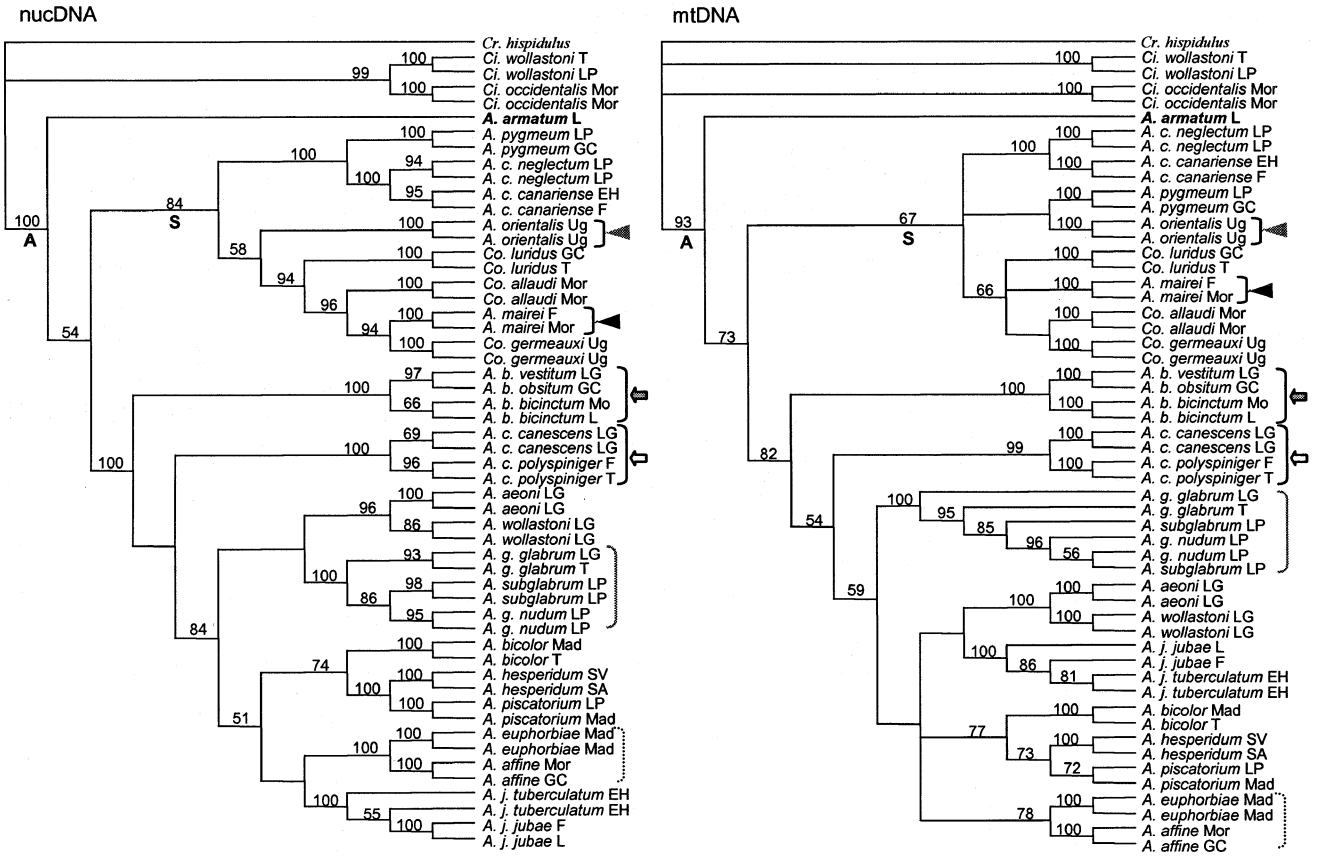

FIGURE 7. Comparison of the most parsimonious mtDNA and nucDNA consensus trees. Bootstrap support higher than 50% is given above each node. See Table 2 and Figure 3 for further details.

originally African (excluding island distributions), the result indicated a putative minimum of two island colorizations and three back colorizations. Under all analyses, the island-mainland distributions observed in *A. bicinctum* and *A. affine* are most likely due to back colonization to the continent.

Host switching.—The transition to breeding in *Euphorbia* occurred twice in Crypturgini (Fig. 10). Within *Aphanarthrum* and *Coleobothrus*, a single origin of breeding in succulents occurred after the transition to *Euphorbia*. Likelihood estimation of the ancestral character state in *Aphanarthrum* suggested that feeding on arborescent shrubs was more than twice as likely as feeding on succulent euphorbs (Fig. 10). The number of transitions to *E. balsamifera* (one to three) and *E. lamarckii* complex (one to two) was more uncertain, with one and three origins of *E. lamarckii* and *E. balsamifera* feeding, respectively, under ACCTRAN optimization. Likelihood estimation of ancestral states in the shrub feeding clade also supported three separate origins of *E. balsamifera*, with all supraspecies nodes obtaining at least twice the probability for *E. lamarckii* as the ancestral host in the shrub feeding clade.

Speciation modes.—Comparisons of sister subspecies, species, and higher groups identified five cases of allopatric splits in five subspecies pairs, four allopatric and three sympatric splits in seven species pairs, and three allopatric and nine sympatric splits among higher level

TABLE 7. Percentage increase in parsimony tree lengths when optimized on competing topologies from other partitions or combinations thereof. Bold numbers mark significant difference as measured by the Wilcoxon signed rank test (Bonferroni corrected significance levels for 49 comparisons, 0.001).

Topology→ Data↓	All	mtDNA	nDNA	EF-1α	Enolase	16S	COI	Morph
All	—	1.65	0.32	0.92	0.78	**1.41**	**6.83**	**35.86**
mtDNA	1.02	—	1.88	**2.34**	**2.44**	0.71	1.93	**39.92**
nDNA	0.19	**4.13**	—	0.68	0.39	**3.21**	**12.20**	**35.15**
EF-1α	1.17	**5.24**	1.05	—	2.91	2.80	**16.20**	**39.51**
Enolase	0.93	**4.82**	0.68	2.62	—	**4.74**	**10.91**	**33.93**
16S	1.78	0.49	2.91	**4.53**	**4.21**	—	**11.17**	**42.56**
COI	3.03	1.67	**3.79**	3.64	**4.02**	3.41	—	**40.45**
Morph	20.00	29.09	20.00	16.36	20.00	23.64	**49.09**	—

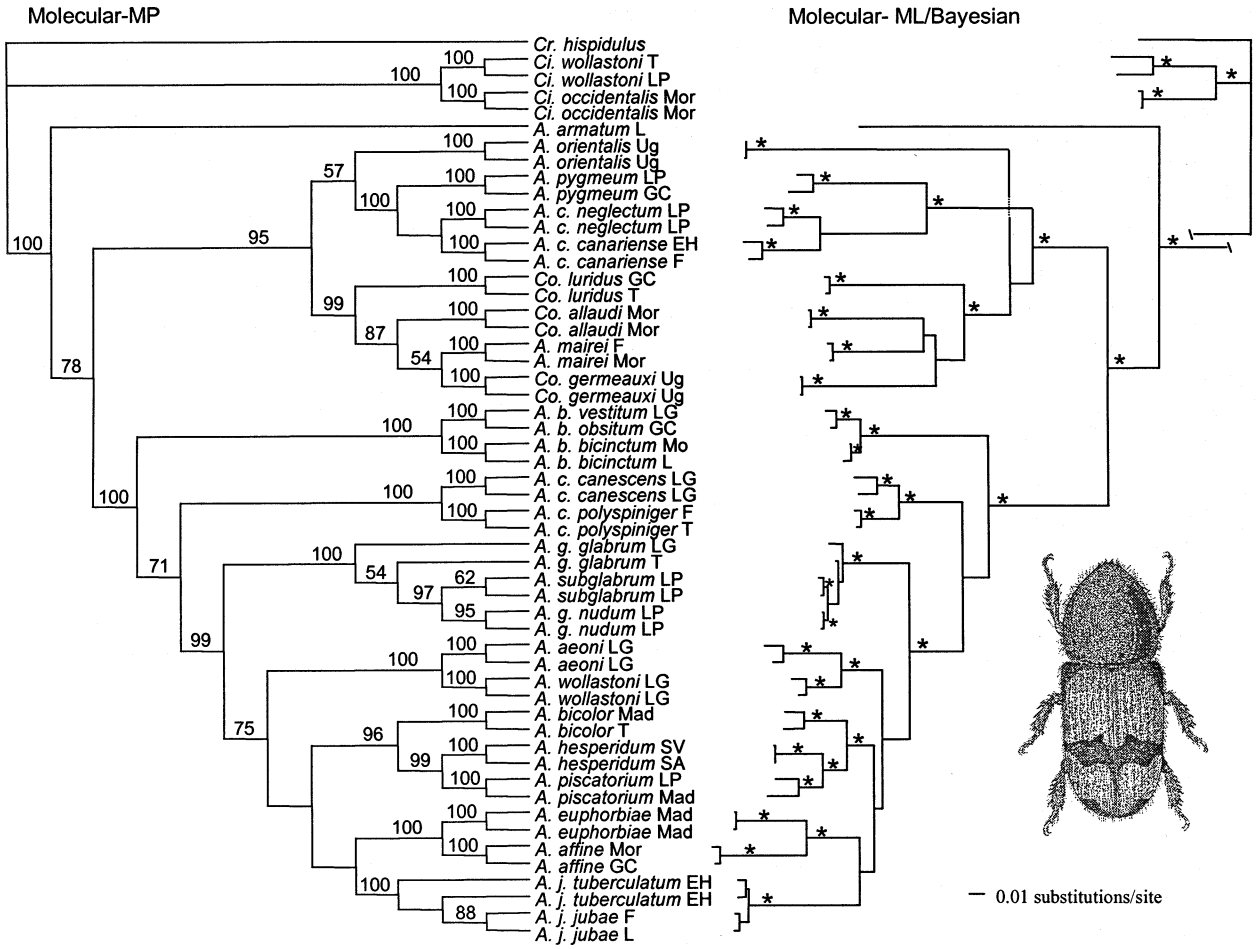

FIGURE 8. Comparison of the most parsimonious (left) and Bayesian/ML (right) tree topologies resulting from the analyses of all molecular data (12-parameter mixed model and uniform model indifferent). Addition of morphological data did not change these topologies in the parsimony or mixed-model Bayesian analyses. Posterior probabilities above 95% are indicated by a star in the Bayesian topology (12 parameters). The line drawing illustrates *Aphanarthrum armatum*.

clades (Table 8). Only one shift in host plant group was found in the sister species or subspecies comparisons, but this event occurred in allopatry (*A. euphorbiae*).

DISCUSSION

Phylogenetic Analyses

Data combinability.—The simultaneous analysis of five different data partitions resulted in a well-resolved tree with most nodes receiving higher branch support than in any of the separate analyses. Some of the improvement in branch support was likely due to the increased number of informative characters in the combined data matrix. Variable-length bootstrap resampling of each partition showed an overall (but modest) correlation between the number of resampled characters and branch support for nodes congruent with the combined data (except COI), a pattern observed in numerous other studies (e.g., Cummings et al., 1995; Springer et al., 2001; Koepfli and Wayne, 2003; Rokas et al., 2003). However, the choice of character sets sampled is not trivial in resolving conflicting nodes, due to apparent idiosyncrasies of single genes. Location dependent processes in the genome, including biased base frequencies as observed in the COI and *Enolase* genes (see Table 4), may allow individual partitions to converge towards wrong topologies (e.g., Cummings et al., 1995; Mitchell et al., 2000). Sampling characters from many independent sources will most likely reduce the negative effect of nonindependence when location specific idiosyncrasies are diluted in combined analyses. This could be a possible explanation for

TABLE 8. Summary of putative speciation modes associated with cladogenetic events among subspecies, species and higher clades.

	Geography			
	Allopatry		Sympatry	
Host group	Same	Shift	Same	Shift
Subspecies	5	0	0	0
Species	3	1	3	0
Clades	3	0	5	4
Subtotal	11	1	8	4
Total	12		12	

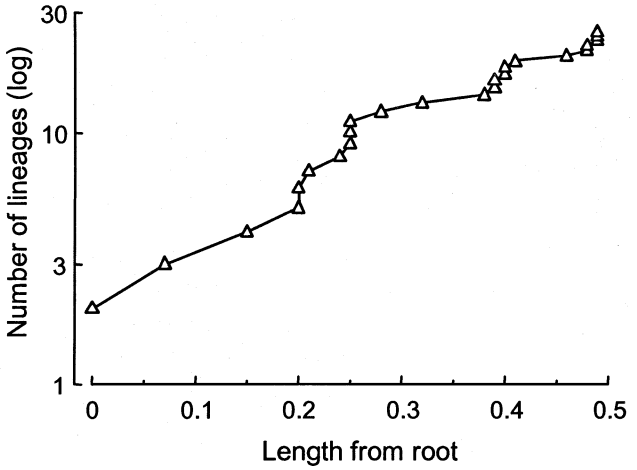

FIGURE 9. Lineage through time plot showing nearly constant speciation rate in *Aphanarthrum*. Splits between subspecies were included, but not splits between other populations. Length from the root was calculated as the number of mitochondrial substitutions per site, using the best mtDNA likelihood model with clock enforced.

the largely increased support in the combined data observed for the *Coleobothrus* clade and the clades F and G (cf. Fig. 1) that were incongruent in some of the separate analyses. Empirical evidence so far have unequivocally shown that an equal number of characters sampled from smaller fragments throughout the genomes resolve the species tree more quickly than in a gene by gene procedure (e.g. Cummings, 1995; Rokas et al., 2003). Our sample of four gene fragments and morphological data may not yet constitute a balanced representation of the entire genome, explaining why all nodes are not yet fully resolved. Nevertheless, our results illustrate well how the process of combining multiple sources of data leads to complementary support in various hierarchical levels of the tree topology, and falls in line with several other studies which have emphasized the importance of using multiple and independent sources of phylogenetic characters (Baker and DeSalle, 1997; Mitchell et al., 2000; Baker et al., 2001; Cameron and Mardulyn, 2001; Cognato and Vogler, 2001; Koepfli and Wayne, 2003; Rokas et al., 2003).

Although topological incongruence with the combined data ranged from 1 (16S) to 14 (COI) conflicting nodes in the separate parsimony analyses, these conflicting nodes had low, if any, branch support. Hence, the analyses of separate data did not help to clarify any of the less resolved relationships from the combined analyses; neither did any of the separate gene trees suggest any strongly supported alternative hypotheses. Still, measures of incongruence (ILD, Wilcoxon sign rank test, SH test) indicated significant differences among partitions, especially between nuclear and mitochondrial partitions. A general weakness of these tests is that they do not consider the nature of incongruence, such that one rogue taxon may cause the entire data set to be incongruent. Also, the likelihood- and parsimony-based incongruence tests produced very different results, further reducing confidence in such methods. The use of partitioned branch support (partitioned Bremer support *sensu* Baker, 1997), on the other hand, is a very useful tool in the localization and quantification of node conflicts. Although the significance of Bremer's (1994) node decay index is dependent on the branch length (DeBry, 2001), the way that the PBS measure is used here only considers the relative contribution (or conflict) to each node separately. The largest cost of any data partitions enforced on the combined topology was four extra steps in one *Enolase* node (see Fig. 5), which is nevertheless insignificant when branch length is taken into account (cf. Table 3 in DeBry, 2001). The heterogeneity measured between the nuclear and mitochondrial partitions was hardly visible in the PBS analyses, with only three nodes having a negative contribution from both mtDNA partitions (clades including *A. affine, A. euphorbiae, A. jubae*; *A. hesperidum, A. piscatorium*; all of these plus *A. bicolor*). However, these three nodes were characterized by very short ML branch lengths in separate and combined analyses (cf. Figs. 6, 8), and the apparent conflict may reflect random sorting events during a short period of elevated speciation rates (cf. Fig. 9).

Potential negative effects of combining 'statistically' incongruent partitions involve less resolved or less supported tree topologies. However, exclusion or sequential down weighting of the most incongruent partitions, COI (entire gene fragment or third positions only) or morphology, did not change the combined data topologies in neither of the parsimony or Bayesian analyses. Further exclusion of alignment ambiguous regions still resulted in the same topology, with node support remaining essentially the same, or slightly lower, through all these analyses. Moreover, the application of a mixed model in the Bayesian analyses did not produce different results from the uniform Bayesian analysis. Hence, the Bayesian results suggest that the effect of fitting separate models to subpartitions become less critical with an increased number of independent character sources.

Overall, our data have demonstrated how significant measures of incongruence can mislead when conflict in fact is very low (Baker et al., 2001). The incongruence observed between the separate data partitions then possibly reflects the smaller number of informative characters in each of them, which makes tree estimation more susceptible to homoplasious changes, in particular those partitions experiencing the highest substitution rates.

Phylogenetic relationships.—This study clearly demonstrated the paraphyly of *Aphanarthrum* with respect to *Coleobothrus*. The basal position of *A. armatum* was supported by all data sets except COI, in either parsimony or likelihood analyses, or both. COI performed much worse than any other partition, and had the highest level of homoplasy, significant nonstationarity in base composition, the fewest supported nodes, and the lowest branch support per informative character. The quite different properties (and congruence with the combined data tree) of the completely linked 16S data emphasize that the poor performance of the COI data is not genome specific, and confirm previous findings that COI is not very

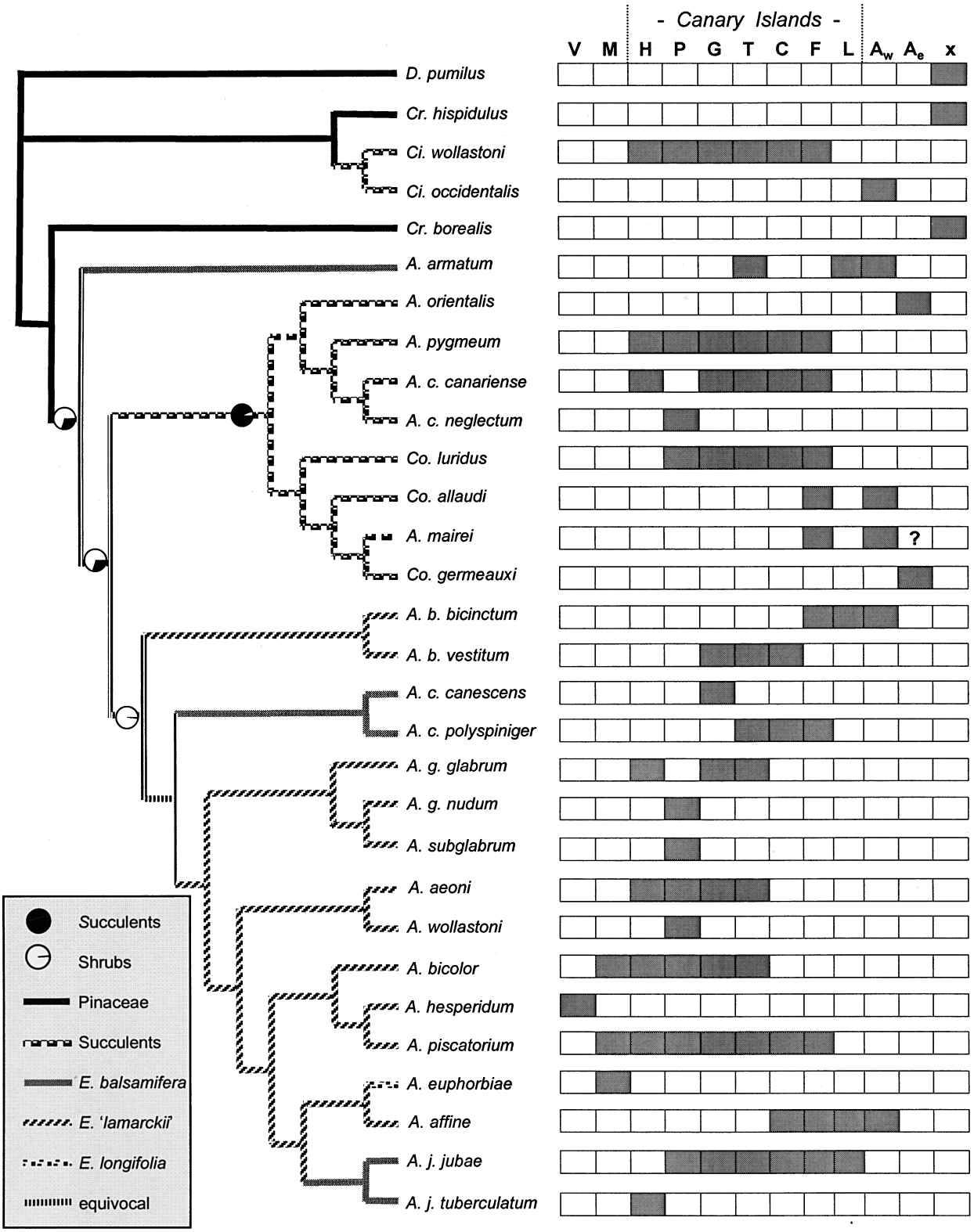

FIGURE 10. Parsimony (branches) and likelihood (pies) reconstruction of host plant use on the combined parsimony topology. Succulents include *E. canariense, E. handiense, E. echinus, E. teke*; the '*E. lamarckii*' complex includes *E. regis-jubae, E. lamarckii, E. artropurpurea, E. tuckeyana, E. piscatoria*. Pies show the probability of ancestral host plant groups (succulents versus arborescent shrubs). The right hand matrix displays the geographical distribution of each species or subspecies: V, Cape Verde archipelago; M, Madeira archipelago; H, El Hierro; P, La Palma; G, La Gomera; T, Tenerife; C, Gran Canaria; F, Fuerteventura; L, Lanzarote (all Canary Islands); A_w, western Africa (Morocco, Senegal, and Gambia); A_e, eastern Africa (Uganda, Kenya, Sudan, Ethiopia); X, Holarctic region.

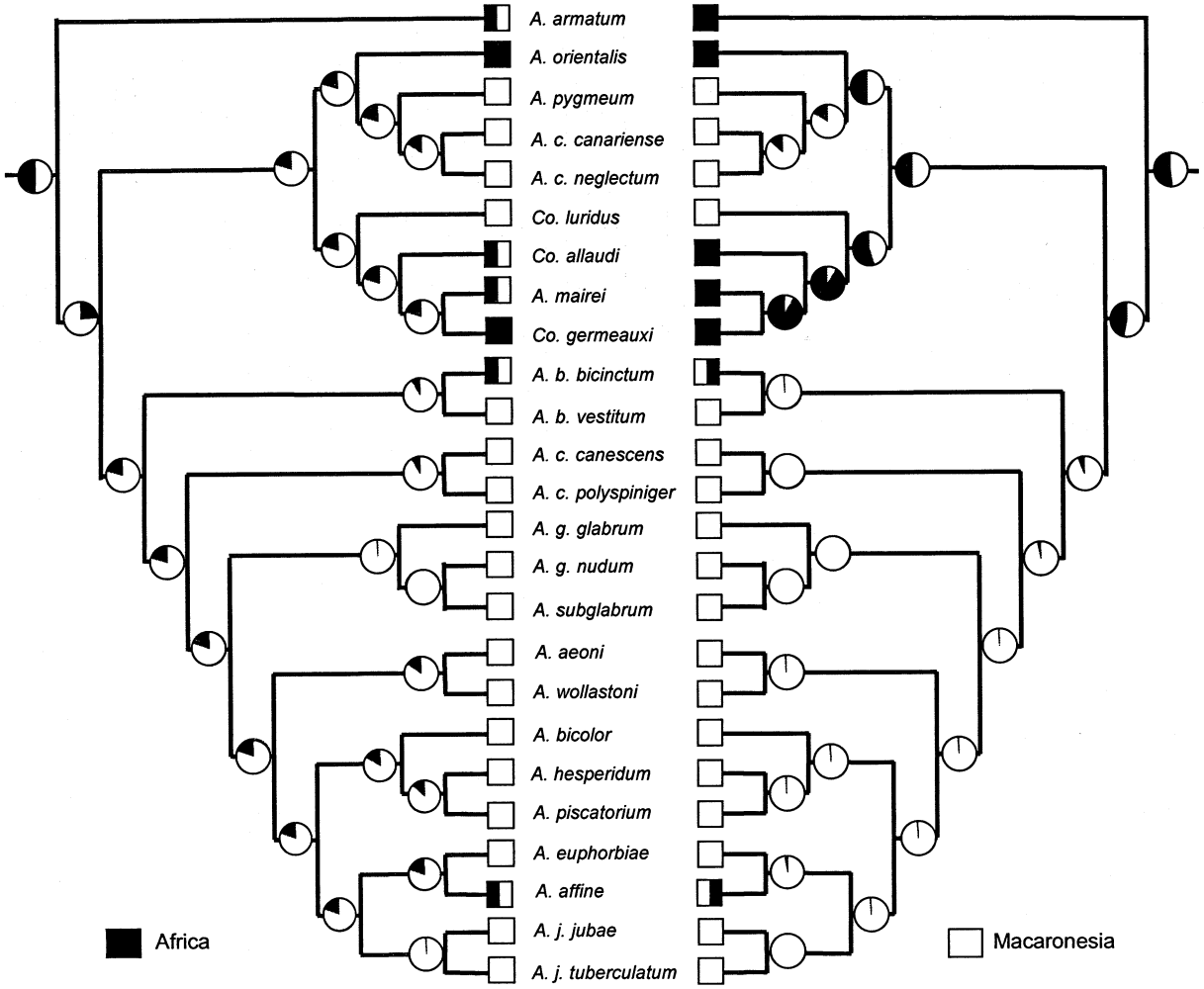

FIGURE 11. Likelihood estimation of the geographical distribution of ancestors (mainland, black; Macaronesian islands, white). Terminal boxes indicate the current distribution. The relative proportion in each pie indicates the probability of each ancestral character state. The alternative reconstruction shown on the right assumes recent African to Macaronesian dispersal in some species, including the basal *A. armatum*.

useful as a stand alone data source for higher level insect phylogenetics (Durando et al., 2000; Baker et al., 2001; Jordal et al., 2002; Danforth et al., 2003). However, it is also possible that our COI data included conserved mitochondrial pseudogenes integrated in the nuclear genome. Although none of these sequences contained stop codons, several recent studies have detected multiple integrations of indel-free mitochondrial pseudogenes (Bensasson et al., 2001, but see Williams and Knowlton, 2001, for an extreme example). Until more conclusive evidence is obtained in favor or disfavor of the pseudogene hypothesis, we caution against any strong interpretations of the COI data. We note, however, that exclusion of the COI data did not change the topology, or the average bootstrap support, rendering these data harmless in the face of potential paralogy.

The inclusion of *A. mairei* in the *Coleobothrus* clade, and the basal position of *A. orientalis* (ML/Bayesian) in this clade, reinforced the evidence against the monophyly of *Coleobothrus*. Although the two *Aphanarthrum* species had base compositions more similar to the *Coleobothrus* species than to the other *Aphanarthrum* species, LogDet distance analyses of the nuclear DNA data, or all molecular data, did not shift the position of these species. Thus, biased base frequencies can not alone explain these relationships (DeBry, 2003). Also several suits of morphological characters associate *A. orientalis* and *A. mairei*, with *Coleobothrus* (Fig. 2). Male genitalia are nearly identical in these five species, all with a hair-pinned ventral sclerite and asymmetrical tegmen (Appendices 1, 2). Consequently, *Coleobothrus* must be treated as a synonym of *Aphanarthrum*, thus rejecting Menier's proposed generic status of *Coleobothrus* (Menier, 1973). In the other taxa of the succulent breeders, *A. canariense* and *A. pygmeum*, genitalia are highly specialized, or reduced, showing how rapid genitalic characters can diverge in secondarily sympatric species. However, molecular data left no doubt about the monophyly of all species breeding in succulent euphorbs.

The 12 remaining Macaronesian *Aphanarthrum*, which all breed in shrubs, made up one strongly supported clade. The support for *A. bicinctum* and *A. canescens* as

the two most basal taxa was also strong. Even though the separate analyses disagreed upon which of these taxa should be most basal, most of them, as well as the combined parsimony, likelihood and Bayesian analyses, placed *A. bicinctum* in that position. All 12 species are externally very similar to each other, but the shape of the transverse elytral fasciae distinguish the two basal taxa from the remaining ten. Whereas most nodes in this clade were strongly supported, two short internodes had the lowest possible branch support (two basal nodes in *A. glabrum*'s sister group; Fig. 1). The two nodes were also those that most frequently changed under various combinations of partitions (and various optimality criteria) and we consider these unresolved until more data are collected. However, it seems a difficult task to achieve well supported resolutions for these nodes, based on the low branch support provided by all five partitions. The low resolution was also reflected in short internal branch lengths in each of the underlying gene trees (cf. Fig. 6), corresponding to a small but sudden increase in speciation rate (cf. Fig. 9). Under such circumstances, only large amounts of data would provide a sufficient number of informative characters needed to improve phylogenetic accuracy and confidence.

Given the extensive overlap and general lack of good diagnostic external characters, the strongly supported monophyly of each species was strikingly clear. With only one exception, in the paraphyletic *glabrum* complex, these species have existed as separate evolutionary entities for millions of years (all interspecific COI divergences above 10%). The deep coalescence between ecologically and morphologically similar species thus refute the hypotheses that hybridization or recent diversification act as possible confusing factors in the evolution of morphological variation.

Subspecies Formation

The paraphyly of *A. glabrum* with respect to *A. subglabrum*, and the very little genetic divergence between the two species, illustrate a case of very recent speciation. The paraphyletic signature is a predicted one under circumstances where new species originate by peripheral isolation, and does not preclude species status of the two species involved (Harrison, 1998). Some speciation events must necessarily be recent and too little time has elapsed for complete and reciprocal monophyly to evolve for every single gene tree. Population sizes in these species, as in most species of *Aphanarthrum*, are enormous and probably require millions of generations to obtain exclusive groups. Nevertheless, the *Enolase* gene strongly supported reciprocal monophyly of each of the two *glabrum* subspecies and *A. subglabrum*, suggesting that selection on certain nuclear genes may follow the assumed reproductive isolation between the two species on La Palma. More samples are needed to explore the exact nature of this incipient speciation process, and a more detailed investigation is currently under way. What this example has clearly demonstrated though is the need for sampling multiple populations and markers to discover the paraphyletic nature of species.

Sampling of multiple populations also provided information on the reality of proposed subspecies (Israelson, 1972). Overall, most species showed clear geographical divergences corresponding to the suggested subspecific classification. Although we did not manage to sample one of the *A. canescens* subspecies, and *A. bicinctum obsitum* from Gran Canaria and *A. jubae tuberculatum* could not be judged by the limited sample included, all other subspecies were clearly distinct and each of the gene trees showed reciprocal monophyly of *A. canescens* and *A. canariense* subspecies, whereas *Enolase* and the combined data showed the same for *A. bicinctum*. Very deep coalescence characterized the first two subspecies splits, which were even deeper than some of the splits between morphologically distinct species (e.g., *A. bicolor* versus *A. hesperidum* versus *A. piscatorium*, or *A. wollastoni* versus *A. aeoni*). Taken together with the almost straight line LTT plot (Fig. 9), the recent formation of subspecific lineages fits the general picture of a group undergoing active evolution and continuous speciation.

Factors Associated with Speciation

Host plant switching has only played a minor role in recent speciation events in crypturgine beetles. Among the total of 24 cladogenetic events in the *Aphanarthrum* clade (Fig. 10), only 6 of these were associated with major changes in host plant use, significantly lower than expected from a random distribution of character states (8 to 12, average 10.3). The use of succulent euphorbs was particularly conservative, with no reversals to other host groups within *Aphanarthrum*. Given that this lineage has existed for a long time, and that both succulents and shrub-like euphorbs coexist in most Canary islands and in Morocco, stasis in host plant use may be interpreted as strong constraints on host switching. Switching between the various shrub-like euphorb groups was more frequent, but even the most recent of these (to *E. longifolia*) involved considerable sequence divergence (12.8% to 13.5% COI, uncorrected), and occurred in allopatry.

Many recent sister lineages are allopatrically distributed, suggesting that these diverged by geographical isolation on different islands. The degree of sympatry increases with increasing phylogenetic depth, however, suggestive of secondary range expansion (Barraclough et al., 1998; Barraclough and Vogler, 2000). Although allopatric differentiation may be the most universal speciation mode in island archipelagos, as has also been shown in some other insect groups (Jordan et al., 2003; Percy, 2003), several studies have reported a more even contribution (Kambysellis and Craddock, 1997; Joy and Conn, 2001), or even favored ecological specialization as the most prominent isolating factor on islands (Shaw, 2002).

Colonization Patterns

Several lines of evidence suggest that the relatively short distance between the African continent and the

Atlantic islands has not been an important barrier to dispersal in this group of insects. Contrary to the single-island colorizations seen in many flightless beetles (Juan et al., 2000; Gillespie and Roderick, 2002), the high frequency of geographic interchange in *Aphanarthrum* beetles seems connected to their flight capacity. Recent exchange has taken place within *A. mairei*, *A. affine*, and *A. bicinctum*, and seems equally likely in two more species, *A. armatum* and *A. allaudi*, which have previously been collected in western Africa and Fuerteventura, respectively (Fig. 10). Although the most parsimonious or likely solution of African-Macaronesian exchange suggests a Macaronesian origin of *Aphanarthrum* and five back colonisations to the mainland, this is dependent on the ancestral distribution of the basal *A. armatum*. Given the frequent interchange and high dispersal capacity in these beetles, *A. armatum* could have originally been restricted to Africa only, implying up to three ancient colonisations of Macaronesia. A reasonable consensus area of origin may include the Moroccan coast and the oldest (eastern) Canary islands. After all, the most basal taxon is currently found in this area, and the ancestor of at least one of the eastern African species, *Co. germeauxi*, most likely occurred in the same area.

Whereas the geographical origin of *Aphanarthrum* remains uncertain, the monophyletic clade of Macaronesian shrub breeding species implies a single common ancestor from the Canary Islands. Furthermore, the two species endemic to the Cape Verde and Madeira islands are nested within this clade, related to species endemic to a restricted number of islands. *Aphanarthrum hesperidum* is related to *A. piscatorium* and *A. bicolor*, a group of mainly western Canary Islands species. Interestingly, the host plant in Cape Verde, *E. tuckeyana*, is closely related to species in the *lamarckii* complex in the Canary Islands as well (Molero et al., 2002), and both host plant and beetle species seem to have followed the same dispersal route. A similar hypothesis exists for one clade of *Liparthrum* beetles that has the same specialized association with *Euphorbia* (Jordal et al., 2003). Little more is known about the origin of the Cape Verde flora and fauna, but at least in *Echium* plants there is strong evidence in favour of an eastern Canary Islands to Cape Verde dispersal route. The species endemic to Madeira, *A. euphorbiae*, is related to an eastern Canary Island plus Moroccan species, but is nested within a clade of Canary Island endemics, suggesting that these islands served as the ancestral area. Recent dispersal in *A. piscatorium* and *A. bicolor* from the western islands, in particular Tenerife (phylogeographical studies in progress, but see Fig. 1), to several Madeiran islands further supports our hypothesis for a Canary Island to Madeira dispersal route. Relatively more is known about the origin of the Madeiran flora and fauna, which have been derived from a variety of geographical sources. Among these, only plants seem to have colonized from the western islands (Bohle et al., 1996; Panero et al., 1999; Barber et al., 2002), whereas several beetle groups had a more eastern origin (Emerson et al., 2000; Rees et al., 2001), and examples of trees, butterflies, and reptiles indicate an Iberian or Moroccan continental origin (Brunton and Hurst, 1998; Nogales et al., 1998; Hess et al., 2000).

Conclusion

This phylogenetic study has demonstrated the importance of thorough sampling in two different ways. First, it shows how samples from multiple populations can validate or reject the existence of genetically distinct subspecific lineages, which can further provide sufficient evidence to assess the phyletic status of species. Secondly, it highlights the many pitfalls associated with limited sampling of characters. Whereas most character partitions produced many identical results, no single partition had the capacity of resolving all nodes with confidence, or achieve congruence with a topology based on all available data. Because node-specific conflict was low among the partitions, incongruence was probably due to different levels of homoplasy, and not necessarily to different genealogical histories, further supporting the strategy of combining all data to estimate a new hypothesis on *Aphanarthrum* phylogeny.

The relatively low frequency of host switching renders geographical isolation a much more important factor in the recent diversification of *Aphanarthrum* beetles. However, rare transitions into novel resources have nevertheless opened avenues of new opportunities for these beetles, by providing unexploited ecological niches for further diversification. Thus, the contribution from one major historical host switch, e.g., to succulent growth forms, has provided the *Aphanarthrum* lineage with 10 additional species. This demonstrates in the simplest possible way how important host switching may be to the total diversity, even though other factors seem currently much more influential.

Acknowledgements

The potential of Macaronesian *Aphanarthrum* as a useful evolutionary model system was first pointed out by Lawrence Kirkendall and Harald Breilid. Together with Kjetil Harkestad they also provided specimens during the initial stage of the project. We would further like to thank Pedro Oromi and Roberto Jardim for help with logistics and permits to collect in some of the national parks; many thanks also to each of the 'Cabildos' in the Canary Islands for issuing collecting permits. Sharon Shute and Martin Baehr kindly arranged access to type material in the Natural History Museum, London, and Zoologische Staatssammlung, Munich. Finally, we wish to thank Chris Simon, Patrick Mardulyn, and Mike Caterino for helpful comments on a previous version of this paper. This project was funded by a Marie Curie Fellowship HPMF-CT2001-01323 to BHJ.

References

Baker, R. H., and R. DeSalle. 1997. Multiple sources of character information and the phylogeny of Hawaiian drosophilids. Syst. Biol. 46:645–673.

Baker, R. H., G. S. Wilkinson, and R. DeSalle. 2001. Phylogenetic utility of different types of molecular data used to infer evolutionary relationships among stalk-eyed flies (Diopsidae). Syst. Biol. 50:87–105.

Barber, J. C., J. Francisco-Ortega, A. Santos-Guerra, K. G. Turner, and R. K. Jansen. 2002. Origin of Macaronesian *Sideritis* L. (Lamioideae: Lamiaceae) inferred from nuclear and chloroplast sequence datasets. Mol. Phylogenet. Evol. 23:293–306.

Barker, F. K., and F. M. Lutzoni. 2002. The utility of the Incongruence Length Difference test. Syst. Biol. 51:625–637.
Barraclough, T. G., and S. Nee. 2001. Phylogenetics and speciation. Trends Ecol. Evol. 16:391–399.
Barraclough, T. G., and A. P. Vogler. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. Am. Nat. 155:419–434.
Barraclough, T. G., A. P. Vogler, and P. H. Harvey. 1998. Revealing the factors that promote speciation. Phil. Trans. R. Soc. Lond. B 353:241–249.
Bensasson, D., D. Zhang, D. L. Hartl, and G. M. Hewitt. 2001. Mitochondrial pseudogenes: Evolution's misplaced witnesses. Trends Ecol. Evol. 16:314–321.
Bohle, U., H. H. Hilger, and W. F. Martin. 1996. Island colonization and evolution of the insular woody habit in *Echium* L. (Boraginaceae). Proc. Natel. Acad. Sci. U.S.A. 93:11740–11745.
Bremer, K. 1994. Branch support and tree stability. Cladistics 10:295–304.
Brown, R. P., N. M. Suarez, A. Smith, and J. Pestano. 2001. Phylogeography of Cape Verde Island skinks (*Mabuya*). Mol. Ecol. 10:1593–1597.
Brunton, C. F. A., and G. D. D. Hurst. 1998. Mitochondrial DNA phylogeny of Brimstone butterflies from the Canary Islands and Madeira. Biol. J. Linn. Soc. 63:69–79.
Buckley, T. R., C. Simon, and G. K. Chambers. 2001. Exploring among-site rate variation models in a maximum likelihood framework using empirical data: Effects of model assumptions on estimates of topology, branch lengths, and bootstrap support. Syst. Biol. 50:67–86.
Buckley, T. R., C. Simon, P. K. Flook, and B. Misof. 2000. Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. Insect Mol. Biol. 9:565–580.
Cameron, S. A., and P. Mardulyn. 2001. Multiple molecular data sets suggest independent origins of highly eusocial behaviour in bees (Hymenoptera: Apinae). Syst. Biol. 50:194–214.
Carter, S., and A. R. Smith. 1988. Euphorbiaceae (part 2). Flora of Tropical East Africa (R. M. Polhill, ed.). A. A. Balkema, Rotterdam.
Cognato, A. I., and A. P. Vogler. 2001. Exploring data interaction and nucleotide alignment in a multiple gene analysis of *Ips* (Coleoptera: Scolytinae). Syst. Biol. 50:758–780.
Cummings, M. P., S. P. Otto, and J. Wakeley. 1995. Sampling properties of DNA sequence data in phylogenetic analysis. Mol. Biol. Evol. 12:814–822.
Danforth, B. N., L. Conway, and S. Ji. 2003. Phylogeny of eusocial *Lasioglossum* reveals multiple losses of eusociality within a primitively eusocial clade of bees (Hymenoptera: Halictidae). Syst. Biol. 52:23–36.
DeBry, R.W., 2001. Improving interpretation of the Decay Index for DNA sequence data. Syst. Biol. 50:742–752.
DeBry, R. W. 2003. Identifying conflicting signal in a multigene analysis reveals a highly resolved tree: The phylogeny of rodentia (Mammalia). Syst. Biol. 52: 604–617.
Despres, L., E. Pettex, V. Plaisance, and F. Pompanon. 2002. Speciation in the Globeflower fly *Chiastocheta* spp. (Diptera: Anthomyiidae) in relation to host plant species, biogeography, and morphology. Mol. Phylogenet. Evol. 22:258–268.
Durando, C. M., R. H. Baker, W. J. Etges, W. B. Heed, M. Wasserman, and R. DeSalle. 2000. Phylogenetic analysis of the *repleta* group of the genus *Drosophila* using multiple sources of characters. Mol. Phylogenet. Evol. 16:296–307.
Emerson, B. C. 2002. Evolution on oceanic islands: Molecular phylogenetic approaches to understanding pattern and process. Mol. Ecol. 11:951–966.
Emerson, B. C., P. Oromi, and G. M. Hewitt. 2000. Interpreting colonization of the *Calathus* (Coleoptera: Carabidae) on the Canary Islands and Madeira through the application of the parametric bootstrap. Evolution 54:2081–2090.
Farrell, B. D., A. Sequeira, B. O'Meara, B. B. Normark, J. Chung, and B. Jordal. 2001. The evolution of agriculture in beetles (Curclionidae: Scolytinae and Platypodinae). Evolution 55:2011–2027.
Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1995. Constructing a significance test for incongruence. Syst. Biol. 44:570–572.
Gillespie, R. G., and G. K. Roderick. 2002. Arthropods on islands: Colonization, speciation, and conservation. Ann. Rev. Ent. 47:595–632.
Harrison, R. H. 1998. Linking evolutionary pattern and process: The relevance of species concepts for the study of speciation. Pages 19–31 *in* Endless forms: Species and speciation (D. J. Howard and S. H. Berlocher, eds.). Oxford University Press, Oxford.
Hess, J., J. W. Kadereit, and P. Vargas. 2000. The colonization history of *Olea europaea* L. in Macaronesia based on internal transcribed spacer 1 (ITS-1) sequences, randomly amplified polymorphic DNAs (RAPD), and intersimple sequence repeats (ISSR). Mol. Ecol. 9:857–868.
Israelson, G. 1972. Male copulatory organs of Macaronesian species of *Aphanarthrum* Wollaston. With designations of lectotypes and descriptions of new taxa (Col. Scolytidae). Ent. Scand. 3:249–257.
Israelson, G. 1976. Redescription of *Deropria elongata* (Eggers), with notes on some species of *Aphanarthrum* Wollaston (Coleoptera, Scolytidae). Zool. Mededelingen 50:39–44.
Israelson, G. 1979. Taxonomical and nomenclatural notes on some Canarian Coleoptera. Vieraea 9:183–210.
Jordal, B. H. 2002. Elongation Factor 1α resolves the monophyly of the haplodiploid ambrosia beetles Xyleborini (Coleoptera: Curculionidae). Insect Mol. Biol. 11:453–465.
Jordal, B. H., L. R. Kirkendall, and K. Harkestad. 2004. Phylogeny of a Macaronesian radiation: Host-plant use and possible cryptic speciation in *Liparthrum* bark beetles. Mol. Phylogenet. Evol. 31:554–571.
Jordal, B. H., B. B. Normark, B. D. Farrell, and L. R. Kirkendall. 2002. Extraordinary haplotype diversity in haplodiploid inbreeders: Phylogenetics and evolution of the sib-mating bark beetle genus *Coccotrypes*. Mol. Phylogenet. Evol. 23:171–188.
Jordan, S., C. Simon, and D. Polhemus. 2003. Molecular systematics and adaptive radiation of Hawaii's endemic damselfly genus *Megalagrion* (Odonata: Coenagrionidae). Syst. Biol. 52:89–109.
Joy, D. A., and J. E. Conn. 2001. Molecular and morphological phylogenetic analysis of an insular radiation in Pacific black flies (*Simulium*). Syst. Biol. 50:18–38.
Juan, C., B. C. Emerson, P. Oromi, and G. M. Hewitt. 2000. Colonization and diversification: Towards a phylogeographic synthesis for the Canary Islands. Trends Ecol. Evol. 15:104–109.
Kambysellis, M. P., and E. M. Craddock. 1997. Ecological and reproductive shifts in the diversification of the endemic Hawaiian *Drosophila*. Pages 475–509 *in* Molecular evolution and adaptive radiation (T. J. Givnish and K. J. Sytsma, eds.). Cambridge University Press, Cambridge.
Kelley, S. T., and B. D. Farrell. 1998. Is specialization a dead end? The phylogeny of host use in *Dendroctonus* bark beetles (Scolytidae). Evolution 52: 1731–1743.
Koepfli, K. and R. K. Wayne. 2003. Type I STS markers are more informative than Cytochrome *b* in phylogenetic reconstruction of the Mustelidae (Mammalia: Carnivora). Syst. Biol. 52: 571–593.
Lee, M. S. Y. 2001. Unalignable sequences and molecular evolution. Trends Ecol. Evol. 16:681–685.
Lockhart, P. J., M.A. Steel, M.D. Hendy, and D. Penny. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 11: 605–612.
Losos, J. B., and R. E. Glor. 2003. Phylogenetic comparative methods and the geography of speciation. Trends Ecol. Evol. 18:2003.
Machado, C. A., E. Jousselin, F. Kjellberg, S. G. Compton, and E. A. Herre. 2000. Phylogenetic relationships, historical biogeography and character evolution of fig-pollinating wasps. Proc. R. Soc. London Ser. B 268:685–694.
Maddison, W. P., and D. R. Maddison, 1992. MacClade. Version 3.04. Sinauer Associates, Sunderlands, Massachusetts.
Menier, J. J. 1973. Rehabilitation du genre *Coleobothrus* Enderlein et description d'une espece nouvelle de l'est africain: *C. germeauxi*. Bull. Soc. Entomol. France 78:205–209.

Mitchell, A., C. Mitter, and J. C., Regier. 2000. More taxa or more characters revisited: Combining data from nuclear protein-encoding genes for phylogenetic analyses of Noctuoidea (Insecta: Lepidoptera). Syst. Biol. 49:202–224.
Molero, J., T. Garnatje, A. Rovira, N. Garcia-Jacas, and A. Susanna. 2002. Karyological evolution and molecular phylogeny in Macaronesian dendroid spurges (*Euphorbia* subsect. *Pachycladae*). Plant Syst. Evol. 231:109–132.
Nixon, K. C., and J. M. Carpenter. 1996. On simultaneous analysis. Cladistics 12:221–241.
Nogales, M., M. Lopez, J. Jimenez-Asensio, J. M. Larruga, M. Hernandez, and P. Gonzalez. 1998. Evolution and biogeography of the genus *Tarentola* (Sauria: Gekkonidae) in the Canary Islands, inferred from mitochondrial DNA sequences. J. Evol. Biol. 11:481–494.
Normark, B. B., B. H. Jordal, and B. D. Farrell. 1999. Origin of a haplodiploid beetle lineage. Proc. R. Soc. Lond. B 266:2253–2259.
Pagel, M. 1994. Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters. Proc. R. Soc. London Ser. B 255:37–45.
Pagel, M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. Syst. Biol. 48:612–622.
Panero, J. L., F. Francisco-Ortega, R. K. Jansen, and A. Santos-Guerra. 1999. Molecular evidence for multiple origins of woodiness and a New World biogeographic connection of the Macaronesian Island endemic *Pericallis* (Asteraceae: Senecioneae). Proc. Natl. Acad. Sci. U. S. A. 96:13886–13891.
Pellmyr, O., and J. Leebens-Mack. 2000. Reversal of mutualism for adaptive radiation in Yucca moths. Am. Nat. 156:S62–S76.
Percy, D. M. 2003. Radiation, diversity and host plant interactions among island and continental legume-feeding psyllids. Evolution 57:2540–2556.
Poe, S., and D. L. Swofford. 1999. Taxon sampling revisited. Nature 398:299–300.
Pollock, D. D., D. J. Zwickl, J. A. McGuire, and D. M. Hillis. 2002. Increased taxon sampling is advantageous for phylogenetic inference. Syst. Biol. 51:664–671.
Posada, D., and K. A. Crandall. 1998. Modeltest: Testing the model of DNA substitution. Bioinformatics 14:817–818.
Rees, D. J., B. C. Emerson, P. Oromi, and G. M. Hewitt. 2001. The diversification of the genus *Nesotes* (Coleoptera: Tenebrionidae) in the Canary Islands: Evidence from mtDNA. Mol. Phylogenet. Evol. 21:321–326.
Rokas, A., B.L Williams, N. King and S. B. Carroll. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425:798–804.
Ronquist, F., and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.
Rosenberg, M. S., and S. Kumar. 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. Proc. Natel. Acad. Sci. U.S.A. 98:10751–10756.
Scheffer, S. J., and B. M. Wiegmann. 2000. Molecular phylogenetics of the holly leafminers (Diptera: Agromyzidae: *Phytomyza*): Species limits, speciation and dietary specialization. Mol. Phylogenet. Evol. 17:244–255.
Seigler, D. S. 1994. Phytochemistry and systematics of the Euphorbiaceae. Ann. Missouri Bot. Garden 81:380–401.
Shaw, K. L. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proc. Natl. Acad. Sci. U.S.A. 99:16122–16127.
Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16:1114–1116.
Sorenson, M. D. 1999. TreeRot, version 2. Boston University, Boston.
Springer, M. S., R.W. DeBry, C. Douady, H. M. Amrine, O. Madsen, W. W. deJong, and M. J. Stanhope. 2001. Mitochondrial versus nuclear gene sequences in deep-level mammalian phylogeny reconstruction. Mol. Biol. Evol. 18:132–143.
Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24:4876–4882.
Tracy, M. R., and S. B. Hedges. 2000. Evolutionary history of the *Enolase* gene family. Gene 259:129–138.
Webster, G. L. 1994. Classification of the Euphorbiaceae. Ann. Missouri Bot. Garden 81:3–32.
Wheeler, W. C., J. Gatesy, and R. DeSalle. 1995. Elision: A method for accomodating multiple molecular sequence alignments with alignment-ambiguous sites. Mol. Phylogenet. Evol. 4:1–9.
Williams, S. T., and N. Knowlton. 2001. Mitochondrial pseudogenes are pervasive and often isidious in the snapping shrimp genus *Alphaeus*. Mol. Biol. Evol. 18:1484–1493.
Wood, S. L. 1986. A reclassification of the genera of Scolytidae (Coleoptera). Gt. Basin Nat. Mem. 10:126.
Wood, S. L. 1988. Nomenclatural changes and new species of Scolytidae (Coleoptera). Gt. Basin Nat. 48:188–195.
Yoder, A. D., J. A. Irwin, and B. A. Payseur. 2001. Failure of the ILD to determine data combinability for Slow Loris phylogeny. Syst. Biol. 50:408–424.

First submitted 1 June 2003; reviews returned 27 October 2003; final acceptance 14 March 2003
Associate Editor: James Whitfield

APPENDIX 1

Morphological characters and description of character states. After each description follows the maximum possible steps for the character, and the consistency (CI) and retention (RI) indices as measured over the combined data parsimony tree (Fig. 1), but including only one terminal per species or subspecies. Characters with a maximum of one *step apply to outgroup* apomorphies only.

Male Genitalia

1. Anterior part of spiculum gastrale: (0) simple rod; (1) with tiny tooth or nod; (2) Y-shaped fork; (3) L-shaped; (4) with subapical tooth. Max steps 12, CI = 0.75, RI = 0.89.
2. Apophyses: (0) long and slender, longer than tube; (1) as short as tube. Max steps 11, CI = 0.5, RI = 0.90.
3. Tegmen: (0) symmetric; (1) asymmetric. Max steps 5, CI = 1.0, RI = 1.0.
4. Manubrium: (0) long, narrow; (1) small, acute; (2) absent. Max steps 7, CI = 0.67, RI = 0.80.
5. Flagellum: (0) short; (1) long, coiled. Max steps 4, CI = 0.50, RI = 0.67.
6. End plates: (0) absent; (1) present. Max steps 7, CI = 0.50, RI = 0.83.
7. Dorsal face of end plates: (0) smooth; (1) fringed. Max steps 2, CI = 1.0, RI = 1.0.
8. Spines on end plates: (0) absent; (1) few and scattered; (2) many. Max steps 3, CI = 0.67, RI = 0.00.
9. Lamina (ventral sclerite): (0) absent; (1) present. Max steps 5, CI = 1.0, RI = 1.0.
10. Terminal tip of lamina: (0) unfolded; (1) folded. Max steps 3, CI = 0.50, RI = 0.50.
11. Anterior lamina: (0) broad; (1) thick, U-shaped; (2) thin, U-shaped; (3) spoon-shaped; (4) narrowly reduced; (5) slender hair-pin; (6) composite. Max steps 19, CI = 0.75, RI = 0.85.

Head

12. Lower frons: (0) smooth; (1) reticulate. Max steps 12, CI = 0.33, RI = 0.82
13. Frons: (0) flat; (1) concave; (2) convex. Max steps 3, CI = 0.50, RI = 0.50.
14. Distance between eyes: (0) 3 × width of eye; (1) 2 × width of eye; (2) 1 × width of eye. Max steps 6, CI = 0.67, RI = 0.75.
15. Eyes: (0) weakly sinuate; (1) emerginate. Max steps 1 (CI, RI n/a).

16. Funicle: (0) 3-segmented; (1) 2-segmented. Max steps 1 (CI, RI n/a).
17. Sutures on antennal club: (0) 2, recurved; (1) 2, constricted; (2) 2, compressed at tip; (3) absent. Max steps 5, CI = 1.0, RI = 1.0.
18. Septum of antennal club: (0) absent; (1) single septum; (2) pair at suture 1. Max steps 5, CI = 0.67, RI =0̂.67.

Pronotum

19. Pronotal surface: (0) smooth; (1) reticulate. Max steps 7, CI = 0.25, RI = 0.50.
20. Pronotal disc: (0) evenly rounded; (1) summit at basal third; (2) flattened. Max steps 7, CI = 0.67, RI = 0.80.
21. Pronotal shape: (0) broadly rounded; (1) narrowly rounded; (2) constricted at anterior fourth. Max steps 13, CI = 0.67, RI = 0.91.
22. Pronotal tubercles at anterior margin: (0) absent; (1) 2–4; (2) 2–4, margin recurvec; (3) 8–10. Max steps 11, CI = 0.33; RI = 0.56.
23. Pronotal asperities: (0) absent ; (1) fine; (2) coarse, extends to posterior parts. Max steps 3, CI = 1.0, RI = 1.0.

Elytra

24. Fasciae on disc: (0) absent; (1) transverse only; (2) with median rings; (3) median rings squared, open anteriorly; (4) median rings filled; (5) median area V-shaped. Max steps 18, CI = 1.0, RI = 1.0.
25. Fasciae on declivity: (0) absent; (1) present. Max steps 13, CI = 0.50, RI = 0.92.
26. Declivity: (0) rounded; (1) concave. Max steps 3, CI = 0.50, RI = 0.50.

Legs

27. Protibial teeth: (0) 10; (1) 6; (2) 5; (3) 4. Max steps 7, CI = 0.33, RI = 0.20.
28. Tarsal segments 1–4: (0) longer than 5; (1) shorter than 5. Max steps 1 (CI, RI, n/a).

APPENDIX 2

Morphological character matrix for the 28 characters described in Appendix 1.

Dolurgus pumilus	000200--0-	-120100010	00000010
Deropria elongata	400001100??	?1??01301?	03200001
Crypturgus borealis	000010--0-	-120112112	20000020
Cr. hispidulus	000010--0-	-020112102	20000020
Cisurgus wollastoni	000010--0-	-001112002	20000030
Ci. occidentalis	000010--0-	-001112002	20000030
Coleobothrus luridus	1012010010	5001111200	20000120
Co. allaudi	1012010010	5001111200	20000120
Co. germeauxi	1012010010	5001111210	20000130
Aphanarthrum armatum	000000--10	5001111210	11011020
A. mairei	1012010010	5001111210	20000030
A. orientalis	1012010010	5001111210	20000030
A. pygmeum	010200--10	4002111211	22150020
A. canariense canariense	2101010010	6002111211	22150030
A. canariense neglectum	2101010010	6002111211	22150030
A. bicinctum bicinctum	1102010010	1101111210	10031030
A. bicinctum obsitum	1102010010	1101111210	10031030
A. bicinctum vestitum	1102010010	1101111210	10041030
A. canescens canescens	1102011010	1101111210	10041030
A. canescens polyspiniger	1102011010	1101111210	10041030
A. subglabrum	3102010210	3101111210	10021030
A. glabrum nudum	3102010010	3101111210	10021030
A. glabrum glabrum	3102010210	3101111210	10021030
A. euphorbiae	1102010010	4101111210	12021030
A. affine	1102010010	4101111210	12021030
A. jubae jubae	1102010010	1101111200	12021030
A. jubae tuberculatum	1102010010	1101111200	11021030
A. aeoni	1102010011	2101111210	11021030
A. wollastoni	1102010011	2101111210	12021030
A. bicolor	1102010110	0101111210	12021030
A. hesperidum	1102010010	2101111210	10021030
A. piscatorium	1102010011	2101111210	10021030