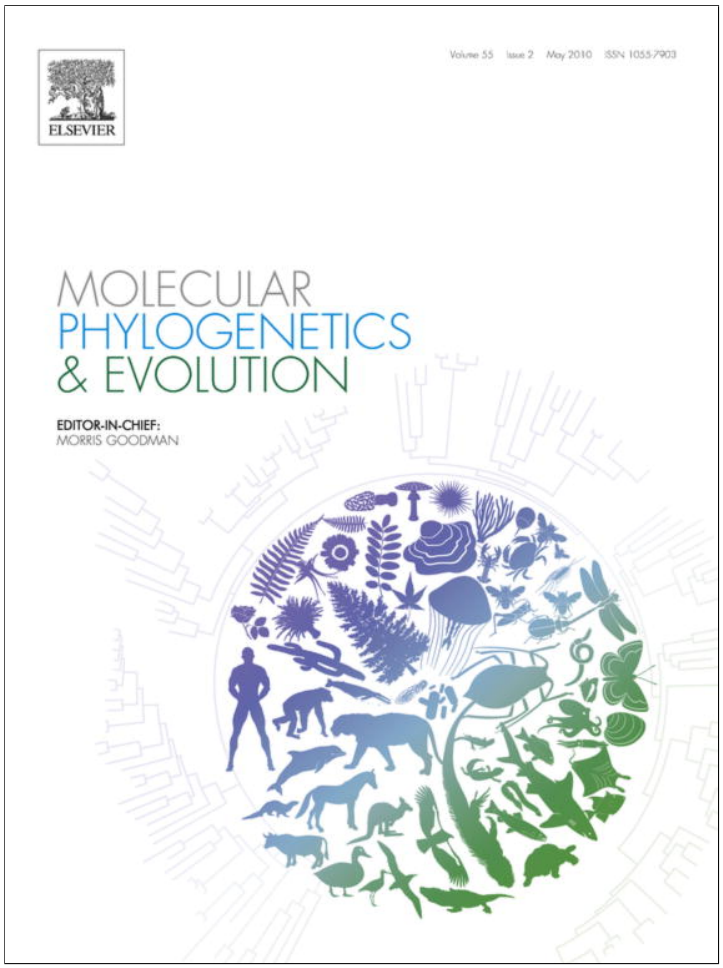

Molecular Phylogenetics and Evolution 55 (2010) 358–371

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

# Molecular phylogeny and Holarctic diversification of the subtribe Calathina (Coleoptera: Carabidae: Sphodrini)

Carlos Ruiz [a,*], Bjarte H. Jordal [b], Brent C. Emerson [c], Kipling W. Will [d], José Serrano [a]

[a] Departamento de Zoología y Antropología Física, Facultad de Veterinaria, Universidad de Murcia, Campus de Espinardo, 30071 Murcia, Spain
[b] Museum of Natural History, University of Bergen, Museplass 3, NO-5007 Bergen, Norway
[c] Centre for Ecology, Evolution and Conservation, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK
[d] ESPM Department, Organisms and Environment Division, University of California, 137 Mulford Hall, USA

ARTICLE INFO

*Article history:*
Received 25 November 2008
Revised 16 October 2009
Accepted 19 October 2009
Available online 10 November 2009

*Keywords:*
*Calathus*
Calathina
Carabidae
Sphodrini
Phylogeny
Holarctic region
Macaronesian archipelagos
Nearctic region
Speciation
cox1-cox2
Nuclear markers

ABSTRACT

A molecular phylogeny of the subtribe Calathina was inferred from DNA sequence data from the mitochondrial cox1-cox2 region and the nuclear genes 28S and EF-1α. All lineages within Calathina from the Holarctic region were represented except for the monotypic subgenus *Tachalus*. Maximum Parsimony and Bayesian analyses of the combined data set showed that the subtribe is a monophyletic lineage that includes a single genus *Calathus*, where other taxa currently ranked as independent genera (*Lindrothius*, *Synuchidius*, *Thermoscelis* and *Acalathus*) are nested within this genus. *Neocalathus* and *Lauricalathus*, both subgenera of *Calathus*, were found to be polyphyletic and in need of taxonomic revision. The subtribe appears to have originated in the Mediterranean Basin and thereafter expanded into most parts of the Palearctic region, the Macaronesian archipelagos (at least five independent colonisation events), the Ethiopian highlands and the Nearctic region (at least two independent events).



## 1. Introduction

According to the last Palearctic catalog (Hovorka and Sciaky, 2003), the subtribe Calathina includes four genera, the Holarctic *Calathus* Bonelli 1810, the monotypic genera *Thermoscelis* (Caucasus) and *Synuchidius* (Balkan Peninsula), and the Caucasian genus *Lindrothius*, (14 species). *Calathus* is the richest in species number (about 164), particularly in the Euromediterranean subregion. There are 19 species in the Nearctic region, 3 species in the Oriental region and 17 in the Afrotropical region. Members of *Calathus* are common in most temperate areas. Some lineages are composed of forest-dwelling and often flightless species, with low dispersal power and a restricted distribution. Other lineages include winged species with high dispersal power that are broadly distributed and often found in open and mesic to xeric habitats.

The systematics of the subtribe is based on morphological characters (Putzeys, 1873; Schatzmayr, 1937; Lindroth, 1956; Ball and Nègre, 1972; Perrault, 1977). However, characters commonly used to define species and to rank supraspecific taxa (pronotum shape, elytral vestiture, etc.) are variable (e.g., Ball and Nègre, 1972, p. 518) and thus less suitable for defining evolutionary lineages. A previous molecular study of the tribe Sphodrini (Ruiz et al., 2009) found Calathina to be a monophyletic clade on the basis of mitochondrial and nuclear markers and estimated its origin to be between 10 and 26 Mya (million years ago), thus suggesting that it is a relatively young and species rich group. In the present study we aim to reconstruct the phylogenetic history of this subtribe, to investigate evolutionary relationships among species, diversification rates, geographical origins of main lineages, and the colonisation of the Holarctic, with emphasis on the West-Palearctic region. Hypotheses to be investigated in this study are derived from taxonomical uncertainties. *Acalathus* is a genus that was treated as a subgenus of *Calathus* by Lindroth (1956) and Perrault (1977), as a genus closely related to *Calathus* (Casale, 1988), or more distantly related and placed in the subtribe Dolochina (Hovorka and Sciaky, 2003; Lorenz, 2005). Likewise Lindroth (1956) considered *Dolichus* as a species group within an inclusive genus *Calathus*, while other authors have placed this genus separate from *Calathus* in the subtribe Dolichina (Hovorka and Sciaky, 2003; Lorenz, 2005). The status and limits of certain taxa within Calathina is controversial and we assess hypotheses related to the taxonomic status of proposed

* Corresponding author. Fax: +34 968 364906.
*E-mail address:* caruiz@um.es (C. Ruiz).

1055-7903/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2009.10.026

genera that may be separate from *Calathus* (e.g., *Thermoscelis*, *Lindrothius*) or well-delimited subgenera within this genus (*Amphyginus*, *Bedelinus*, *Neocalathus*, *Lauricalathus*).

Several lineages are particularly species rich on the Macaronesian Islands west of the African coast. Previous studies on these lineages based solely on mitochondrial data have indicated that multiple colonisations have occurred on those islands (Emerson et al., 1999, 2000, see also Ruiz and Serrano, 2006). We further assess this hypothesis with a much broader selection of continental taxa. We have included representative species from nearly the entire geographical range of the subtribe Calathina, and analyzed their phylogenetic relationship using a mitochondrial fragment comprising cox1-tRNAleu-cox2 and fragments of the nuclear genes 28S and EF-1α. Because rDNA and intron sequences frequently include insertions and deletions in multiple alignments, we have explored the effects of different alignment strategies on the reconstruction of the phylogenetic relationships. Finally, we have explored various factors potentially causing a lack of resolution in species rich clades of late Mesozoic age.

## 2. Material and methods

### 2.1. Taxon sampling

Molecular sequence data were collected for 102 individuals of 64 species. All genera and subgenera within the subtribe Calathina have been sampled with the exception of the Mexican monotypic subgenus *Tachalus* (Table 1). Related subtribes Dolichina, Synuchina and Pristosiina were used as outgroups based on previous phylogenetic information (Ruiz et al., 2009). For a few species only dry pinned material was available, resulting in a partial sequence of the mitochondrial cox1-cox2 fragment (950–1300 bp).

### 2.2. PCR and sequencing

DNA was extracted with the Qiagen DNeasy tissue kit (Qiagen, Hilden, Germany). We sequenced a fragment of the mitochondrial cytochrome oxidase I and II (cox1-cox2) genes (1630 bp), which includes part of cox1 gene (865 bp), the complete tRNAleu (63 bp), and cox2 (702 bp). Within the ribosomal 28S rDNA we studied a partial sequence (D2–D4 region) of about 885 bp length. The amplified protein-coding Elongation Factor 1α gene (EF-1α C0) had no introns and consisted of 773 bp. Primers and PCR conditions used are described in Ruiz et al. (2009).

PCR products were purified with isopropanol and 5 M ammonium acetate, and sequencing was performed in both directions using the standard protocol for ABI BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems). Some polymorphic sites were found in a few sequences of EF-1α C0, and nucleotide variation from these sites was named according to International Union of Pure and Applied Chemistry (IUPAC) symbols.

### 2.3. Sequence alignment

Sequences of cox1-cox2 and EF-1α were unambiguously aligned, with insertions of 3 bp at the 3′ end of cox2 aligned after translation to amino acids. To align the 28S rDNA sequences five different strategies were followed: (i) A computer-based alignment with Clustal W version 1.6 (Thompson et al., 1994) in MEGA version 4 (Tamura et al., 2007) for the entire sequence (named here as CW). Transitions were weighted 0.5 related to transversions. Different weighting regimes of the gap opening (GOP) and gap extension penalties (GEP) were assayed, namely GOP/GEP: 1/0.25, 1/0.5, 2/0.25, 2/0.5, 2/1, 4/0.5, 4/2, 8/1, 8/4, and the default 15/6.6 often cited in many papers, resulting in 10 different matrices for each rDNA matrix. (ii) A manual alignment based on secondary structure (Kjer, 1995) following the model of Gillespie et al. (2004) (abbreviated as SS). To characterize the expansion segments we used the Mfold software (http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi) which folds RNA sequences based on minimum free energy (Zuker, 2003). Secondary structure was used to infer unambiguous or conserved regions (28Sc) and ambiguous or variable regions (28Sv), and ambiguous regions were thereafter excluded from analyses. (iii) A combination of the first two approaches (mixed approach or MX), in which the unambiguous and ambiguous regions were first defined based on secondary structure and the ambiguous regions (28Sv) were subsequently aligned in Clustal W using the parameters listed above. (iv) The ClustalW alignment of the ambiguous regions was also performed by anchoring to an unambiguously aligned flank of 10–15 bp on each side (named here as MXan). (v) A manual alignment using ClustalW as the starting point (abbreviated as MA).

Sensitivity analyses were carried out to evaluate different alignment parameters and strategies. Congruence of 28S alignments using different GOP/GEP and unambiguously aligned protein-coding genes (cox1-cox2, EF-1α) was selected as the optimality criterion for further combined analyses. Congruence was measured as the contribution of the gene partition to the Bremer Support (BS) (Bremer, 1994) of the combined analysis (Baker and DeSalle, 1997). This contribution was calculated with BS and Partitioned Bremer Support (PBS) in TreeRot v.3 (Sorenson and Franzosa, 2007). PBS values were standardized for partitions of different length by dividing the PBS value by the minimum possible length of each branch (DeBry, 2001). The relative PBS support for each node of each gene partition was also measured and these measures were related to the accumulated length from the root of the tree to the actual node in the combined data topology.

### 2.4. Phylogenetic inference

Each data partition was analyzed separately and in combination under Maximum Parsimony (MP), as implemented in PAUP* b4.0b10 (Swofford, 2002), and Bayesian analysis (BA) with Mr. Bayes v3.1. (Ronquist and Huelsenbeck, 2003). All data sets were analyzed with Modeltest (v.3.6) (Posada and Crandall, 1998) to find the best-fit model of sequence evolution. The Akaike information criterion (AIC) (Akaike, 1973) was preferred over hLRTs to test the various models as recommended by Posada and Buckley (2004). Parsimony analyses were conducted with 10,000 random addition replicates of heuristic searches and TBR branch swapping. Confidence in each node was assessed by bootstrapping 500 heuristic searches replicates with 20 random additions per replicate. Bayesian analyses were performed with 1,500,000 generations for separate analyses, and 3,000,000 generations for combined analyses, sampling trees every 100 generations. Likelihood values were observed with Tracer v1.4 (Rambaut and Drummond, 2005), discarding all trees before stability in likelihood values as a burn-in (first 3000 trees in separate analyses and 6500 trees in the combined ones). Stationarity was also reassessed using the convergence diagnostic: the average standard deviation of split frequencies and the potential scale reduction factor (PSRF).

To explore long-branch attraction (LBA) artifacts a number of strategies indicated by Bergsten (2005) were followed: analyses of unlinked markers under various tree-reconstruction methods with different sensitivity to LBA, different partition strategies under different models of molecular evolution, removal of a putatively wrongly placed taxon with a re-run of the analysis to check whether the remaining long branched taxa change position, and inclusion of all available sequences of main lineages of *Calathina* and *Dolichina* from GenBank.

**Table 1**
Accession numbers and locality for the taxa of the subtribe Calathina included in the study. AG: Algeria, BU: Bulgaria, CI: Canary Islands, ET: Ethiopia, GG: Georgia, IT: Italy, JA: Japan, MC: Macedonia, MEX: Mexico, MO: Morocco, MR: Madeira, NR: Norway, PL: Poland, PT: Portugal, SP: Spain, TR: Turkey, UK: United Kingdom, USA: United States of America.

| Species | Cod | Locality | cox1-cox2 | 28S | Ef-1α C0 |
|---|---|---|---|---|---|
| *C. (Amphyginus) rotundicollis* | 132 | Canalroya, Huesca (SP) | GU254302 | GU254392 | GU254475 |
| *C. (Amphyginus) rotundicollis* | 137 | Soria (SP) | AM410879 | GU254394 | ——————— |
| *C. (Amphyginus) rotundicollis* | 153 | Vinuesa (SP) | ——————— | FJ173073 | FJ173135 |
| *C. (Amphyginus) rotundicollis* | 203 | Monte Pollino, Calabria (IT) | GU254317 | GU254412 | GU254485 |
| *C. (Bedelinus) circumseptus* | 112 | Albacete (SP) | AM410880 | GU254385 | GU254470 |
| *C. (Bedelinus) circumseptus* | 228 | Córdoba (SP) | FJ173207 | FJ173086 | FJ173144 |
| *C. (Calathus) baeticus* | 143 | Sierra de Baza, Granada (SP) | AM410881 | GU254397 | GU254479 |
| *C. (Calathus) brevis* | 19 | Zamora (SP) | GU254292 | GU254349 | GU254437 |
| *C. (Calathus) fracassii* | 7 | Gran Sasso (IT) | AJ404989 | GU254338 | GU254427 |
| *C. (Calathus) fuscipes* | 13 | Gran Sasso (IT) | AJ404994 | GU254344 | GU254433 |
| *C. (Calathus) fuscipes* | 105 | Sierra de Urbión, Soria (SP) | AM410882 | GU254382 | GU254467 |
| *C. (Calathus) fuscipes algiricus* | 126 | Bab-Berret (MO) | AM410883 | FJ173072 | FJ173134 |
| *C. (Calathus) granatensis* | 9 | Cádiz (SP) | AJ404991 | GU254340 | GU254429 |
| *C. (Calathus) granatensis* | 124 | Igualeja, Málaga (SP) | AM410885 | GU254389 | GU254473 |
| *C. (Calathus) granatensis* | 133 | Soria (SP) | GU254303 | GU254393 | GU254476 |
| *C. (Calathus) hispanicus* | 14 | Sierra de Gredos, Ávila (SP) | AJ404995 | GU254345 | GU254434 |
| *C. (Calathus) hispanicus* | 15 | Pto. de Guadarrama (SP) | AJ404996 | GU254346 | ——————— |
| *C. (Calathus) muchei* | 156 | Samanli Dağlari (TR) | GU254306 | GU254400 | GU254481 |
| *C. (Calathus) opacus* | 8 | Moyen Atlas (MO) | AJ404990 | GU254339 | GU254428 |
| *C. (Calathus) opacus* | 150 | Ifrane (MO) | ——————— | GU254399 | GU254480 |
| *C. (Calathus) pirazzolii* | 6 | Gran Sasso (IT) | AJ404988 | GU254337 | GU254426 |
| *C. (Iberocalathus) rotundatus estrelensis* | 116 | Serra da Estrela (PT) | AM410888 | GU254388 | ——————— |
| *C. (Iberocalathus) rotundatus rotundatus* | 21 | Braganza (PT) | AM410886 | GU254351 | GU254439 |
| *C. (Iberocalathus) rotundatus rotundatus* | 115 | Cabezo de Manzaneda, Orense (SP) | AM410887 | GU254387 | GU254472 |
| *C. (Incertae sedis) aethiopicus* | 253 | Mount. Wachacha.(ET) | GU254329 | ——————— | GU254495 |
| *C. (Incertae sedis) colasianus* | 84 | Chao de Ribiera Seixal (MR) | AJ404998 | GU254374 | GU254461 |
| *C. (Incertae sedis) complanatus* | 83 | Rebecal (MR) | AJ404999 | GU254373 | GU254460 |
| *C. (Incertae sedis) vividus* | 86 | Rebecal (MR) | AJ405002 | GU254375 | GU254462 |
| *C. (Lauricalathus) abaxoides* | 69 | Tenerife (CI) | AJ236975 | GU254368 | GU254456 |
| *C. (Lauricalathus) angularis* | 44 | Gran Canaria (CI) | AJ236947 | GU254359 | GU254447 |
| *C. (Lauricalathus) angustulus* | 95 | Tenerife (CI) | AJ404973 | GU254380 | ——————— |
| *C. (Lauricalathus) appendiculatus* | 40 | Gran Canaria (CI) | AJ236956 | GU254358 | GU254446 |
| *C. (Lauricalathus) ascendens* | 93 | Tenerife (CI) | AJ404980 | GU254378 | GU254464 |
| *C. (Lauricalathus) auctus* | 97 | Tenerife (CI) | AJ404981 | GU254381 | GU254466 |
| *C. (Lauricalathus) canariensis* | 48 | Gran Canaria (CI) | AJ236955 | GU254360 | GU254448 |
| *C. (Lauricalathus) carinatus* | 87 | Tenerife (CI) | AJ404983 | GU254376 | ——————— |
| *C. (Lauricalathus) cognatus* | 52 | La Gomera (CI) | AJ236966 | GU254362 | GU254450 |
| *C. (Lauricalathus) depressus* | 89 | Tenerife (CI) | AJ404977 | GU254377 | GU254463 |
| *C. (Lauricalathus) freyi* | 94 | Tenerife (CI) | AJ404978 | GU254379 | GU254465 |
| *C. (Lauricalathus) gomerensis* | 55 | La Gomera (CI) | AJ236961 | GU254363 | GU254451 |
| *C. (Lauricalathus) laureticola* | 49 | La Gomera (CI) | AJ236968 | GU254361 | GU254449 |
| *C. (Lauricalathus) marcellae* | 61 | La Gomera (CI) | AJ236970 | GU254366 | GU254454 |
| *C. (Lauricalathus) rectus* | 71 | Tenerife (CI) | AJ236959 | GU254369 | GU254457 |
| *C. (Neocalathus) ambiguus* | 10 | Gran Sasso (IT) | AJ404992 | GU254341 | GU254430 |
| *C. (Neocalathus) ambiguus ambiguus* | 28 | Soria (SP) | GU254296 | GU254356 | GU254444 |
| *C. (Neocalathus) ambiguus rugicollis* | 172 | Sakaltutan Pass (TR) | GU254307 | GU254401 | ——————— |
| *C. (Neocalathus) ambiguus rugicollis* | 175 | Güven (TR) | GU254308 | GU254402 | GU254482 |
| *C. (Neocalathus) asturiensis* | 16 | Galicia (SP) | AJ236945 | GU254347 | GU254435 |
| *C. (Neocalathus) asturiensis* | 127 | Parque de Alvao (PT) | GU254300 | GU254390 | GU254474 |
| *C. (Neocalathus) aztec* | 220 | La Marquesa (MEX) | GU254320 | GU254415 | GU254488 |
| *C. (Neocalathus) aztec* | 222 | La Marquesa (MEX) | GU254322 | GU254417 | GU254490 |
| *C. (Neocalathus) aztec* | 267 | La Marquesa (MEX) | GU254333 | GU254425 | ——————— |
| *C. (Neocalathus) bolivari* | 257 | Tlamacas. Popo-Iztaccihuat (MEX) | GU254330 | GU254422 | GU254496 |
| *C. (Neocalathus) cinctus* | 17 | Norfolk Brecklands (UK) | AJ404997 | GU254348 | GU254436 |
| *C. (Neocalathus) cinctus* | 31 | Kizildag Pass (TR) | GU254297 | GU254357 | GU254445 |
| *C. (Neocalathus) cinctus* | 107 | Sierra de Urbión, Soria (SP) | AM410889 | GU254383 | GU254468 |
| *C. (Neocalathus) cinctus* | 141 | Aguelmam Afenourir Lake (MO) | GU254304 | GU254396 | GU254478 |
| *C. (Neocalathus) erratus* | 27 | Rila Mountains (BU) | AM410890 | GU254355 | GU254443 |
| *C. (Neocalathus) erratus* | 113 | Dzyanon village, Dolno (BU) | GU254299 | GU254386 | GU254471 |
| *C. (Neocalathus) gonzalezi* | 77 | Lanzarote (CI) | AJ404986 | GU254372 | GU254459 |
| *C. (Neocalathus) gregarius* | 179 | Pennsylvania (USA) | GU254309 | GU254403 | GU254483 |
| *C. (Neocalathus) leechi* | 221 | La Marquesa (MEX) | GU254321 | GU254416 | GU254489 |
| *C. (Neocalathus) leechi* | 223 | Nevado de Toluca (MEX) | GU254323 | GU254418 | GU254491 |
| *C. (Neocalathus) leechi* | 224 | Nevado de Toluca (MEX) | GU254324 | GU254419 | GU254492 |
| *C. (Neocalathus) leechi* | 226 | Tlamacas. Popo-Iztaccihuat (MEX) | GU254325 | GU254420 | GU254493 |
| *C. (Neocalathus) leechi* | 240 | Nevado de Toluca (MEX) | GU254327 | GU254421 | GU254494 |
| *C. (Neocalathus) leechi* | 258 | La Marquesa (MEX) | GU254331 | GU254423 | GU254497 |
| *C. (Neocalathus) marmoreus* | 218 | Mineral del Chico (MEX) | GU254318 | GU254413 | GU254486 |
| *C. (Neocalathus) melanocephalus* | 11 | Gran Sasso (IT) | AJ236944 | GU254342 | GU254431 |
| *C. (Neocalathus) melanocephalus* | 12 | Norfolk Brecklands (UK) | AJ404993 | GU254343 | GU254432 |
| *C. (Neocalathus) melanocephalus* | 25 | Rila Mountains (BU) | GU254295 | GU254353 | GU254441 |
| *C. (Neocalathus) melanocephalus* | 111 | Formigal, Huesca (SP) | GU254298 | GU254384 | GU254469 |
| *C. (Neocalathus) metallicus* | 26 | Pirin Mountains (BU) | AM410893 | GU254354 | GU254442 |
| *C. (Neocalathus) mexicanus* | 219 | La Malinche (MEX) | GU254319 | GU254414 | GU254487 |

**Table 1** (*continued*)

| Species | Cod | Locality | cox1-cox2 | 28S | Ef-1α C0 |
|---|---|---|---|---|---|
| *C. (Neocalathus) micropterus* | 20 | Poznan (PL) | GU254293 | GU254350 | GU254438 |
| *C. (Neocalathus) micropterus* | 146 | Trondheim (NR) | GU254305 | GU254398 | ——————— |
| *C. (Neocalathus) mollis* | 22 | Málaga (SP) | GU254294 | GU254352 | GU254440 |
| *C. (Neocalathus) mollis* | 120 | La Cuenca, Soria (SP) | FJ173193 | FJ173071 | FJ173133 |
| *C. (Neocalathus) opaculus* | 190 | Tallahasse, Florida (USA) | GU254315 | GU254410 | ——————— |
| *C. (Neocalathus) peropacus* | 188 | Tumacacori, Arizona (USA) | GU254313 | ——————— | ——————— |
| *C. (Neocalathus) ruficollis ruficollis* | 181 | Marin Co., California (USA) | GU254310 | GU254404 | ——————— |
| *C. (Neocalathus) ruficollis ruficollis* | 182 | Marin Co., California (USA) | ——————— | GU254405 | GU254484 |
| *C. (Neocalathus) ruficollis ruficollis* | 183 | Marin Co., California (USA) | ——————— | GU254406 | ——————— |
| *C. (Neocalathus) ruficollis ruficollis* | 185 | Marin Co., California (USA) | GU254311 | GU254407 | ——————— |
| *C. (Neocalathus) ruficollis ruficollis* | 189 | Sta Barbara Co., California (USA) | GU254314 | GU254409 | ——————— |
| *C. (Neocalathus) ruficollis ruficollis* | 193 | Sta Barbara Co., California (USA) | GU254316 | GU254411 | ——————— |
| *C. (Neocalathus) semisericeus* | 131 | Rif Mountains (MO) | GU254301 | GU254391 | ——————— |
| *C. (Neocalathus) semisericeus* | 139 | Rif Mountains (MO) | AM410892 | GU254395 | GU254477 |
| *C. (Neocalathus) simplicicollis* | 75 | Fuerteventura (CI) | AJ404985 | GU254371 | ——————— |
| *C. (Neocalathus) solieri* | 272 | Parc National Chréa (AG) | GU254334 | ——————— | ——————— |
| *C. (Neocalathus) sp* | 262 | Nevado de Toluca (MEX) | GU254332 | GU254424 | GU254498 |
| *C. (Neocalathus) subfuscus* | 72 | Pico Ariens (MR) | AJ405003 | GU254370 | GU254458 |
| *C. (Trichocalathus) obliteratus* | 58 | La Gomera (CI) | AJ236971 | GU254364 | GU254452 |
| *C. (Trichocalathus) pilosipennis* | 62 | La Gomera (CI) | AJ236964 | GU254367 | GU254455 |
| *C. (Trichocalathus) reflexímargo* | 60 | La Gomera (CI) | AJ236974 | GU254365 | GU254453 |
| *Lindrothius caucasicus orbicollis* | 271 | Martveli district, Lebarde (GG) | FJ173220 | FJ173100 | FJ173157 |
| *Lindrothius pseudopraestans* | 296 | Mount Fish (GG) | FJ173235 | ——————— | FJ173171 |
| *Lindrothius sp.* | 297 | Aibja, Abkhazia (GG) | FJ173236 | FJ173115 | FJ173172 |
| *Thermoscelis insignis (Calathina)* | 304 | Summelas, Trabzon (TR) | GU254335 | ——————— | ——————— |
| *Synuchidius ganglbaueri (Calathina)* | 306 | Pelister Mount. Veternica (MC) | GU254336 | ——————— | ——————— |
| *Acalathus (Procalathus) advena (Dolichina)* | 186 | San Bernardino Co., California (USA) | GU254312 | GU254408 | ——————— |
| *Acalathus (Procalathus) advena (Dolichina)* | 242 | Great Basin N.P., Nevada (USA) | GU254328 | ——————— | ——————— |
| *Anchomenidius astur (Dolichina)* | 277 | León (SP) | FJ173221 | FJ173101 | FJ173158 |
| *Dolichus halensis (Dolichina)* | 282 | Nishino, Noichi, Kōchi (JA) | FJ173226 | FJ173105 | FJ173162 |
| *Laemostenus complanatus (Sphodrina)* | 256 | Murcia (SP) | FJ173217 | FJ173097 | FJ173154 |
| *Pristosia aeneola (Pristosiina)* | 285 | Chino, Nagano (JA) | FJ173228 | FJ173108 | FJ173164 |
| *Synuchus nitudus (Synuchina)* | 289 | Jôza, Sakura, Chiba (JA) | FJ173230 | FJ173110 | FJ173166 |
| | | | 102 ind. | 100 ind. | 85 ind. |

### 2.5. Testing for hard polytomies

In order to explore the lack of resolution at deeper phylogenetic nodes, we evaluated whether such polytomies were 'hard', implying rapid radiation. For all tests we used ultrametric trees recovered with the BEAST software (Drummond and Rambaut, 2007). Lineage through time (LTT) plots were used to visualize changes in diversification rates. Two statistical analyses were used to test for significant departures from the constant speciation rate model (CR), the $\gamma$-statistic (Pybus and Harvey, 2000) and the birth–death likelihood (BDL) test (Rabosky, 2006a). Both methods have proven successful in reconstructing diversification rates in various organisms ($\gamma$-statistic: Barraclough and Vogler, 2002; Kadereit et al., 2004; Turgeon et al., 2005; Kozak et al., 2006; McKenna and Farrell, 2006; Weir, 2006; Roelants et al., 2007; Hines, 2008; Janssens et al., 2009; birth–death models: Rabosky et al., 2007; Egan and Crandall, 2008; Rabosky and Lovette, 2008). As incomplete taxon sampling can result in a spurious decline in diversification rates over time (Pybus and Harvey, 2000), we have explored the effects of missing species on our analysis. We assumed that our sample represents a fraction $f$ (from 0.2 to 0.6) of the true number of lineages, and generated sets of 5000 trees under the null hypothesis of constant rate pure birth, in which taxa were randomly pruned from these trees.

$\Gamma$-statistic analyses were conducted with the Ape Library in R (Paradis, 1997). This analysis indicates whether the diversification rate has increased ($\gamma > 0$), decreased ($\gamma > 0$) or remained constant ($\gamma = 0$) over time. To take into account incomplete sampling simulations were carried out using different numbers of missing taxa with the MCCR test (Pybus and Harvey, 2000) in LASER (Rabosky, 2006b). BDL analysis has been found to perform as well or better than the popular $\gamma$-statistic that only identifies temporal decreases in diversification (Rabosky, 2006a). BDL analyses were conducted using LASER library in R (Rabosky, 2006b) to test the null hypothesis of rate-constancy against rate-variable models using AIC. Significance of the change in AIC scores was determined by creating a distribution of AIC scores. This was done by simulating 5000 trees using yuleSim in LASER (Rabosky, 2006b), considering the same number of taxa and the same speciation rate as those estimated under the pure Birth model. Incomplete taxon sampling was con-

**Table 2**
Data statistics for each data set. Ti/Tv, transition/transversion ratio; Pi, parsimony informative sites.

| | A | C | G | Length | Variable | %Variable | Conserved | Pi | Divergence (%) | Ti/Tv |
|---|---|---|---|---|---|---|---|---|---|---|
| cox 1-cox 2 | 33.2 | 13.9 | 13.4 | 1630 | 720 | 44.2 | 908 | 583 | 7.3 | |
| cox1 | 31.1 | 14.3 | 13.9 | 865 | 348 | 40.2 | 517 | 289 | 4.8 | 1.0 |
| tRNAleu | 36.3 | 12.2 | 17.9 | 63 | 21 | 33.3 | 41 | 19 | 16.7 | 4.1 |
| cox2 | 35.5 | 13.6 | 12.4 | 702 | 351 | 50.0 | 350 | 275 | 11.8 | 1.0 |
| EF-1α | 25.4 | 27.0 | 25.0 | 773 | 153 | 19.8 | 620 | 102 | 2.2 | 3.4 |
| 28S CW 15/6.6 | 26.1 | 18.0 | 25.1 | 885 | 198 | 22.4 | 682 | 127 | 2.6 | 1.4 |

**Table 3**
Data sets and the estimated models of sequence evolution by Akaike information criterion (AIC). I = proportion of invariant sites; G = gamma-distributed rates across sites; 28Sc: conserved region of 28S.

| | AIC model | A | C | G | A–C | A–G | A–T | C–G | C–T | I | G |
|---|---|---|---|---|---|---|---|---|---|---|---|
| cox1-cox2 | GTR + I + G | 0.3553 | 0.081 | 0.1078 | 7.1325 | 20.4345 | 6.1987 | 4.5081 | 75.7883 | 0.4607 | 0.7057 |
| EF-1α | SYM + I + G | 0.25 | 0.25 | 0.25 | 0.8711 | 3.8231 | 1.8831 | 0.6169 | 11.8965 | 0.6142 | 0.9119 |
| 28Sc | GTR + I + G | 0.2746 | 0.1844 | 0.2562 | 0.5643 | 7.3847 | 4.1226 | 0 | 3.6522 | 0.7006 | 0.802 |

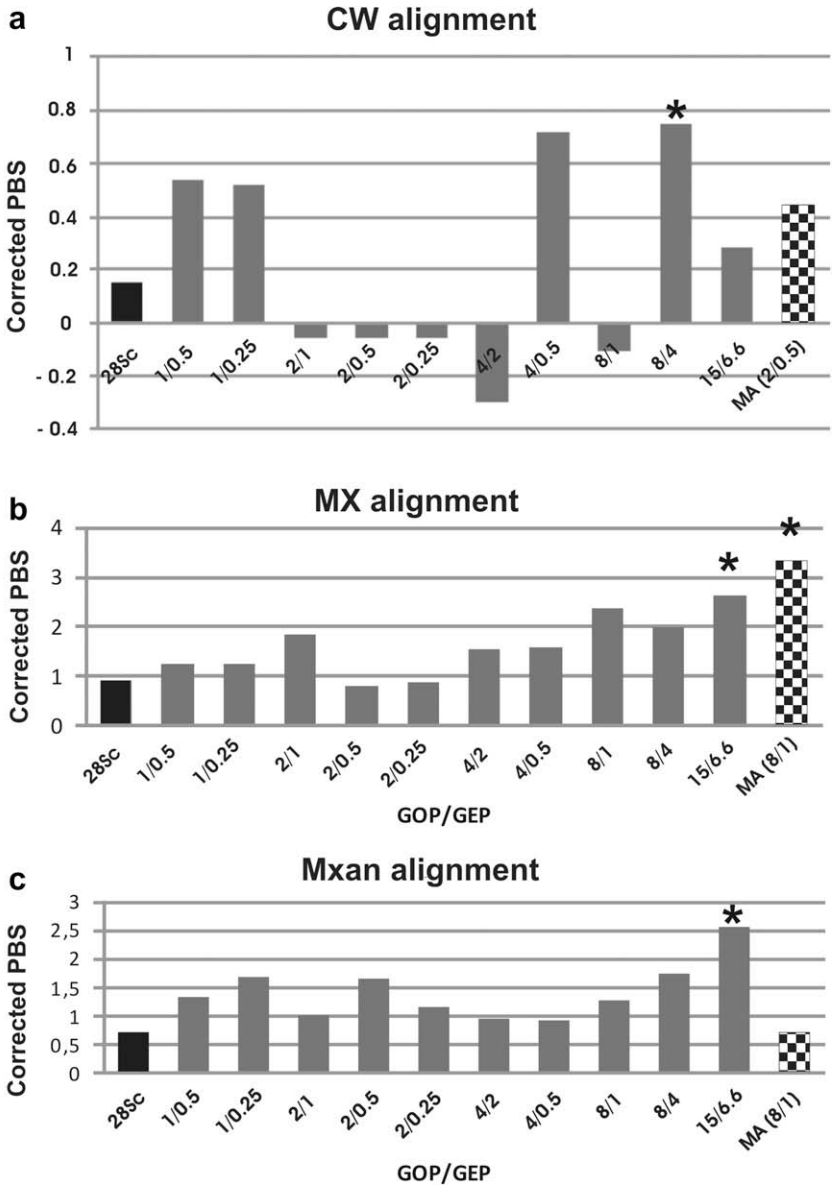


**Fig. 1.** Molecular phylogeny of the subtribe Calathina. Sum of the corrected PBS of all nodes for the Sensitivity Analysis of the whole 28S sequence, aligning manually (MA, checked bars, GOP/GEP of ClustalW as starting point indicated) and with 10 different parameter sets (Grey bars, GOP/GEP) for (a) the whole marker with Clustal W (CW approach); (b) considering only variable regions (MX approach); (c) considering only variable regions and anchoring both ends. (MXan approach). Asterisk indicates selected GOP/GEP values. Black bars represent unambiguously aligned regions (28Sc).

sidered by generating trees under a CR model with different levels of incomplete sampling (*f*) in LASER (Rabosky, 2006b).

The possibility that polytomies were caused by insufficient data ('soft' polytomies) was also evaluated. Simulated data were built with Mesquite v.2.01 (Maddison and Maddison, 2007) to explore the predicted increase in node support associated with an increased number of nucleotides (Philippe et al., 2005). Several previous analyses have used simulations to estimate the amount of data necessary to resolve difficult phylogenetic problems (Berbee et al., 2000; Fishbein et al., 2001; de Queiroz et al., 2002; Rokas et al., 2003; Zou et al., 2008; Spinks et al., 2009). Matrices of increasing size from 500 bp to 20 Kb were constructed taking into account the among-gene-tree variation frequently found in real data sets (Spinks et al., 2009). Two procedures were carried out: (i) simulated data were generated for each partition with a length that was proportional to that of

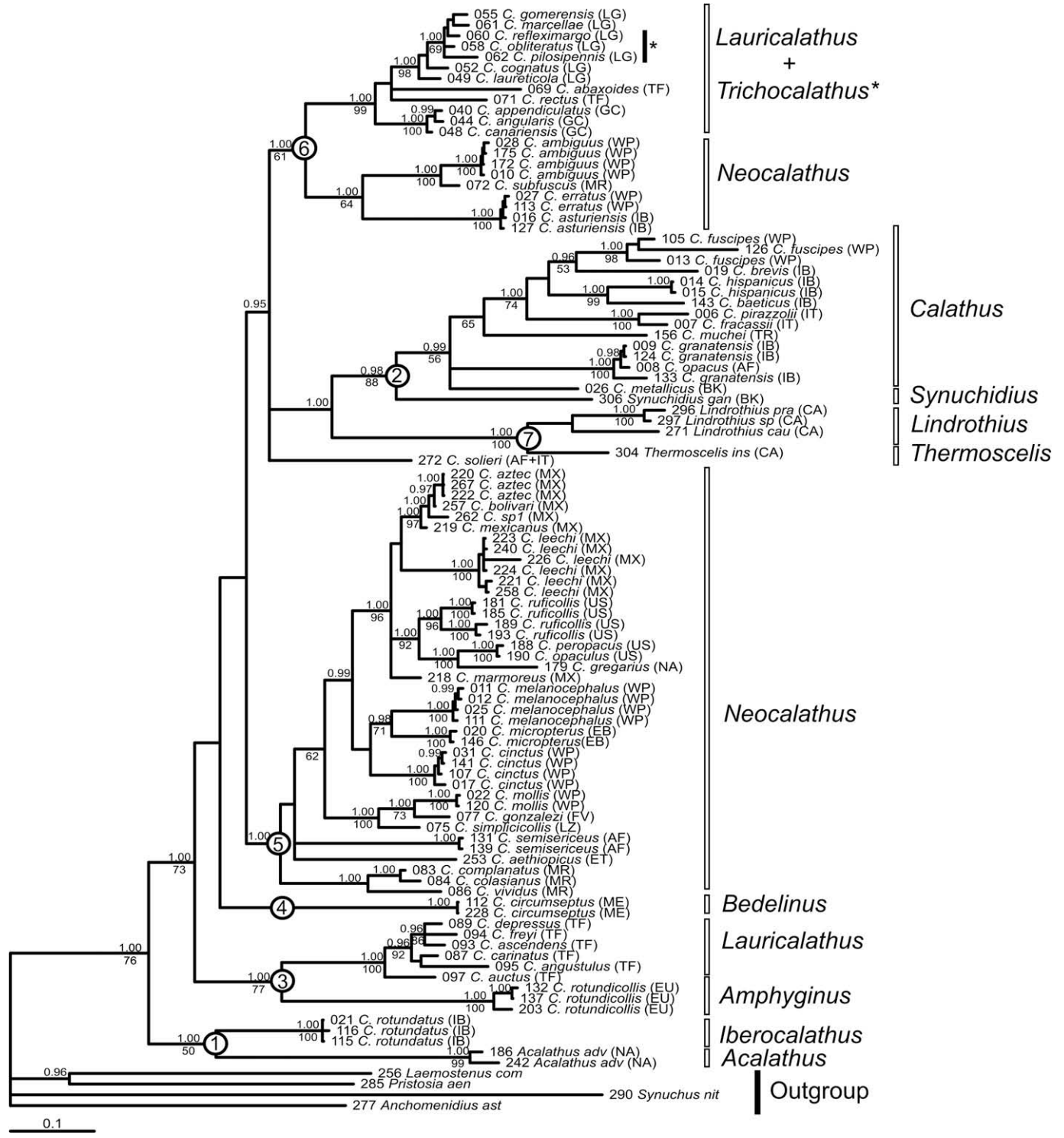


**Fig. 2.** The 50% majority consensus tree resulting from the BA of the cox1-cox2 data partitioned by codon (1st, 2nd and 3rd). Number above node shows posterior probability ⩾0.90. Bootstrap support values are shown below nodes. Circle numbers are discussed in the text. Name of taxa corresponds to those indicated in Table 1 (*C.* corresponds to *Calathus*). Capital letters in brackets denote the geographic region: WP, West-Palearctic; EB, Eurosiberian; EU, European; ME, Mediterranean; AF, North Africa; IB, Iberian peninsula; IT, Italy; BK, Balkan peninsula; TR, Turkey; CA, Caucasus; LZ, Lanzarote Island; FV, Fuerteventura Island; TF, Tenerife; LG: La Gomera Island; GC, Gran Canaria; MR, Madeira; NA, North America; MX, Mexico; US, United States of America; ET, Ethiopia.

the original partitioned dataset. Each simulated partition was generated using both the evolutionary model (Table 3) and the tree obtained from the original partition (Figs. 2–4). Polytomies of the trees from each original partition were resolved arbitrarily and the resulting new branches were assigned zero length with Mesquite v.2.01 (Maddison and Maddison, 2007). The simulated partitions were concatenated and then analyzed with MrBayes; (ii) a non-parametric bootstrap of the combined dataset was performed with Mesquite v.2.01 (Maddison and Maddison, 2007). Pseudoreplicates of variable size were obtained by re-sampling with replacement from the original data and analyzed with MrBayes. Bayesian analyses were carried out for each simulated

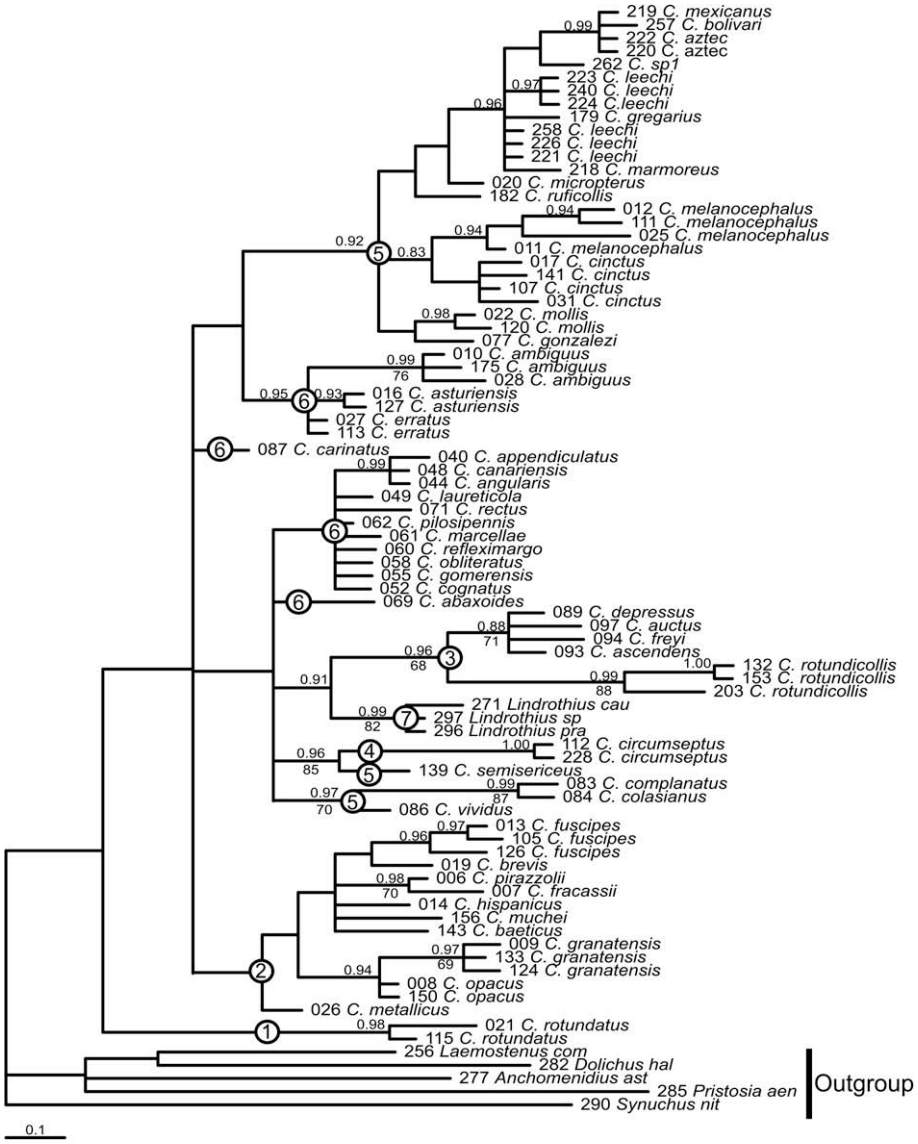


**Fig. 3.** The 50% majority consensus tree resulting from the BA of EF-1α C0 partitioned by codon (1st, 2nd and 3rd). Number above node shows posterior probability ⩾0.90. Bootstrap support values are shown below nodes. Circle numbers are discussed in the text. Name of taxa as in Fig. 2.

matrix and the proportion of nodes with posterior probabilities >0.95 was measured.

## 3. Results

### 3.1. Sequence alignment

The secondary structure inferred for 28S had four variable-length regions, comprising 11% of the total length. The CW approach resulted in generally longer alignments (885–933 base pairs) than the MX (886–910 base pairs) and MXan approaches (886–918 base pairs) (Fig. 1, Supplementary material). The unambiguously aligned regions (28Sc) comprised 67 parsimony informative sites across all taxa (8% of the conserved sites), while the remaining ambiguous regions of alignment (28Sv) exhibited 49–60 additional informative characters (37–60% of the variable regions) depending on the alignment approach and parameters.

Different alignments were evaluated based on their relative contribution (PBS) to the combined analysis of all data. The results of changing alignment parameters are shown in Fig. 1. The Clustal alignment of full length sequences (CW) showed two peaks of higher PBS values under GOP/GEP values 4/0.5 and 8/4. The gap penalties 2/1–0.5–0.25, 4/2 and 8/1 were found to be most incongruent with the combined tree topology. For the MX alignment, the highest PBS value was obtained with gap penalties set to 15/6.6 for the ambiguous regions. When the flanking sides of variable regions were anchored (MXan) the highest PBS value was also found with parameters set to 15/6.6, but the overall contribution using different GOP/GEP values was almost the same for each alignment (Fig. 1b and c). Manual correction for each of the alignment strategies showed different values of PBS (Fig. 1a–c).

After measuring the congruence between the variable regions of 28S and unambiguously aligned regions (mitochondrial plus EF-1α and conserved 28S), we decided to exclude the variable regions from the analysis because the different alignment strategies affected the resolution of the deepest nodes, even in the combined analysis (see below; e.g., relationships between clades 2 and 7 in Fig. 5, and Figs. 3–5 of Supplementary material). This could be due to alignment artefacts that become particularly relevant in cer-

tain clades defined by short internal branches (Kumar and Filipski, 2007).

### 3.2. Phylogenetic inference

#### 3.2.1. Separate analyses

Data statistics and models selected for the data partitions used in this study are shown in Tables 2 and 3. The three independent markers resulted in similar topologies for the subtribe Calathina. The mtDNA fragment generated a tree with high support for the main clades (Fig. 2). However, relationships between these clades were poorly supported in most cases, but slightly increased when each codon position was modelled independently in MrBayes.

The subtribe Calathina includes seven clades (following the notation by Ruiz and Serrano, 2006), which stand out as particularly stable and well supported:

- Clade 1 includes the endemic Iberian species *Calathus* (*Iberocalathus*) *rotundatus* and the North American species *Acalathus* (*Procalathus*) *advena*; this last genus is currently classified in a different subtribe (Dolichina *sensu* Hovorka and Sciaky, 2003 and Lorenz, 2005).
- Clade 2 is constituted by the Balkan *Synuchidius ganglbaueri*, three species of *Calathus* sometimes considered as members of the subgenus *Neocalathus* (*C. granatensis*, *C. opacus*, and *C. metallicus*), and all species of the subgenus *Calathus*.

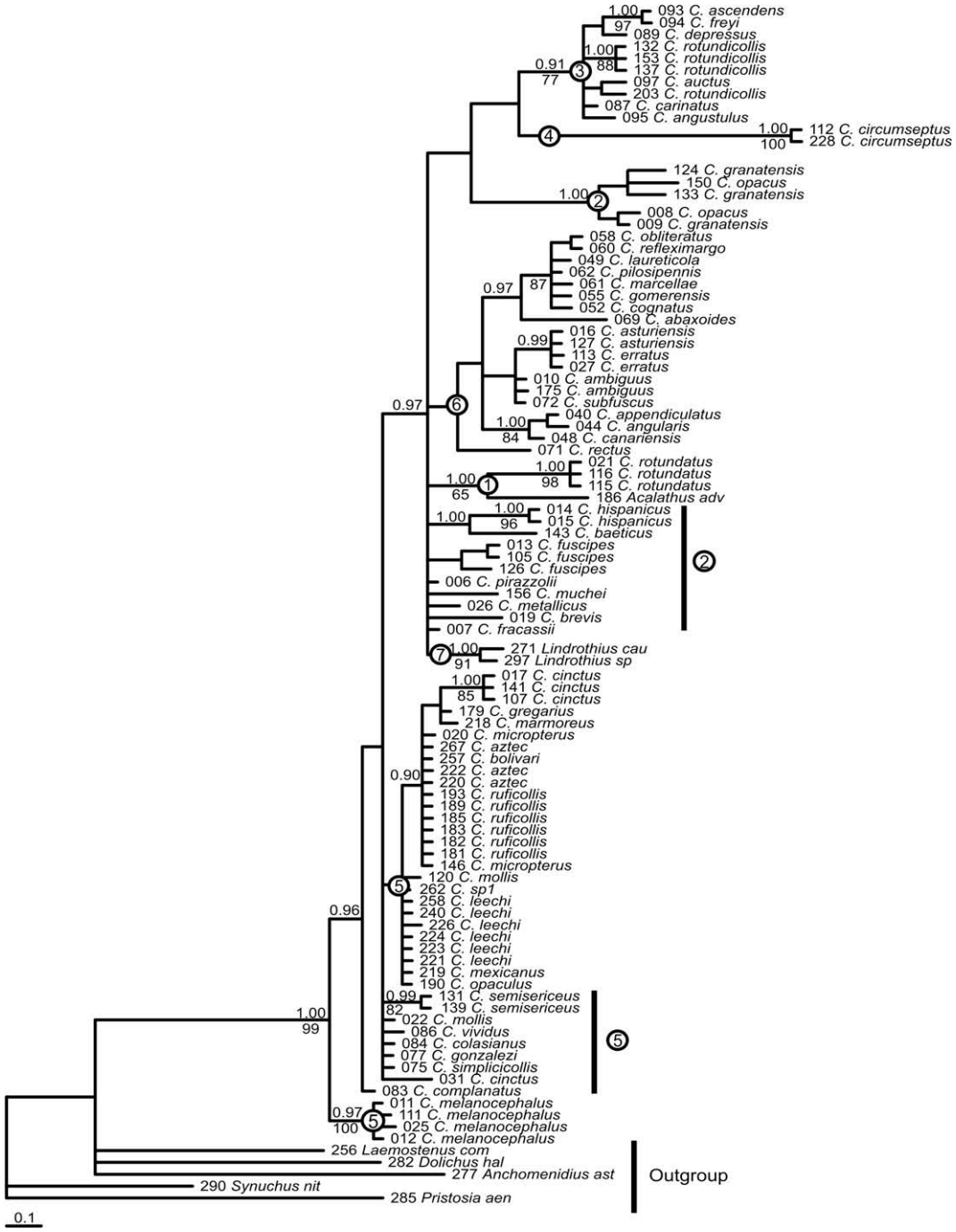


**Fig. 4.** The 50% majority consensus tree resulting from the BA of 28S aligned manually (MA). Number above node shows posterior probability ⩾0.90. Bootstrap support values are shown below nodes. Circle numbers are discussed in the text. Name of taxa as in Fig. 2.

- Clade 3 comprises the West-Palearctic *Calathus* (*Amphyginus*) *rotundicollis* and the Tenerifean *depressus* group (subgenus *Lauricalathus* in part).
- Clade 4 includes *Calathus circumseptus*, a Circummediterranean species of the monotypic subgenus *Bedelinus.*
- Clade 5 includes part of the subgenus *Neocalathus* (the Palearctic *melanocephalus* group plus all the Nearctic species of the subgenus), the North African *C. semisericeus*, and some taxa of uncertain affinity: the Madeiran *vividus* group and the Afrotropical *C. aethiopicus*.
- Clade 6 includes the remaining species of the subgenus *Neocalathus* (the *ambiguus* group) together with species from the Canary Islands (the species of the subgenus *Lauricalathus* other than those of clade 3, plus all *Trichocalathus*).
- Clade 7 comprises the Caucasian genera *Thermoscelis* and *Lindrothius*, which were considered independent genera in the last Palearctic catalog (Hovorka and Sciaky, 2003).

Clade 1 is the sister group of a well-supported clade made up by all other *Calathus* plus the related genera *Synuchidius*, *Thermoscelis* and *Lindrothius* (Fig. 2). Clades 2 and 7 grouped together with maximum support in the Bayesian analysis, and together with clade 6 and the North African *C. solieri* made up a relatively well-supported clade (pp: 0.95).

Nuclear EF-1α resulted in a topology that generally agreed with the mtDNA results (Fig. 3). The same main clades were recovered although the support for some of the deeper nodes was low. *Calathus circumseptus* (clade 4) was recovered with *C. semisericeus* with

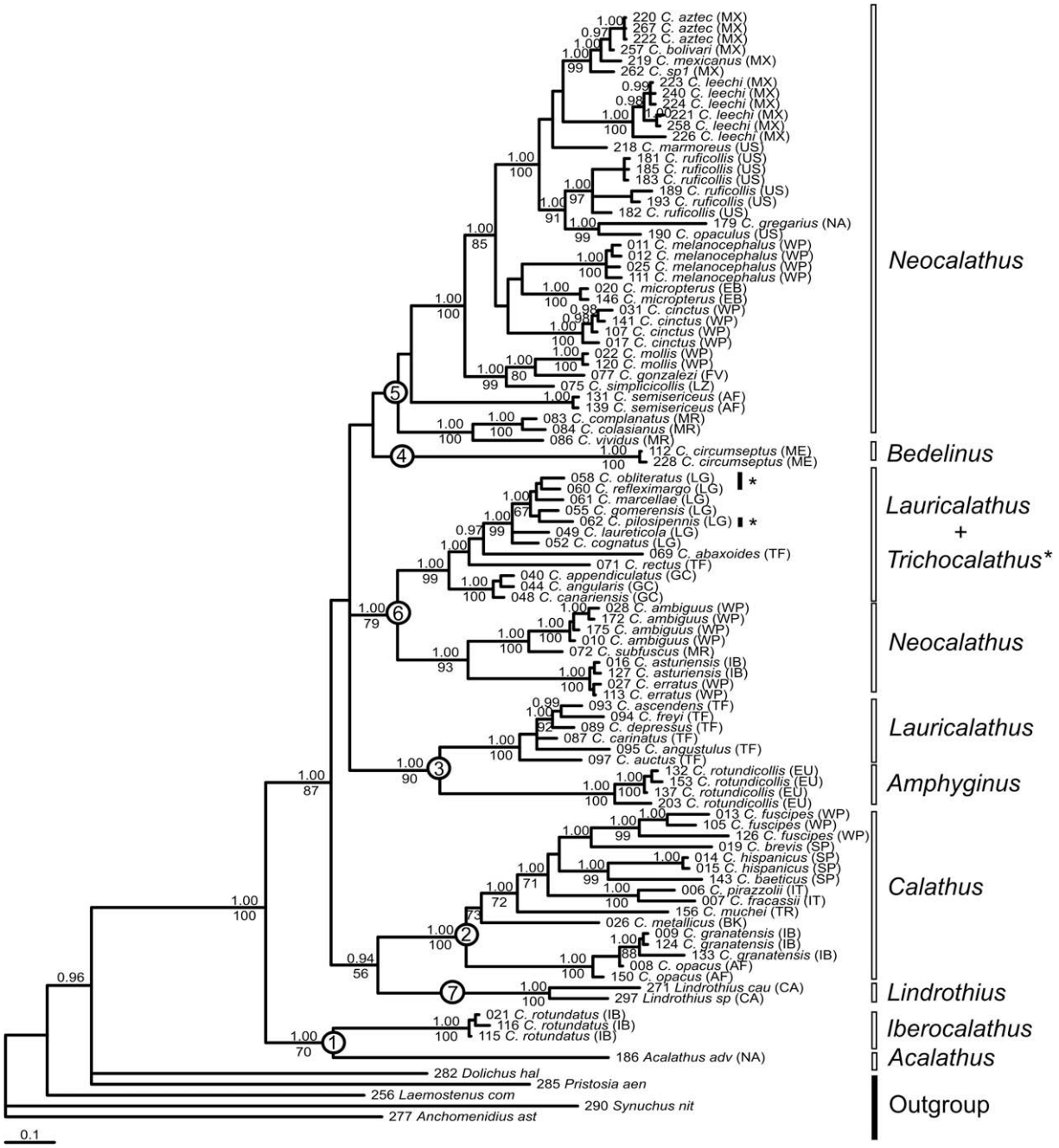


**Fig. 5.** The 50% majority consensus tree resulting from the BA of the combined data without variable region (SS approach). Number above node shows posterior probability ⩾0.90. Bootstrap support values are shown below nodes. Circle numbers are discussed in the text. Asterisk represents species belonging to subgenus *Trichocalathus.* Name of taxa and acronyms of geographic regions as in Fig. 2.

support (pp: 0.96). Clades 3 and 7 were recovered together but this group received varying support depending upon the partitioning strategy (with or without considering codon position, pp: 0.86–0.97).

Ribosomal 28S showed less resolution irrespective of the alignment selected (Fig. 4 and Figs. 3–5 of Supplementary material). Only the monophyly of *Calathus* was supported, while all other clade relationships were sensitive to alignment strategy.

### 3.2.2. Combined analysis

Combined analyses showed very similar results to separate analyses, in particular to the mtDNA analysis, but with increased nodal support. Clades 1–7 recovered by the mtDNA analyses were present in all combined analyses. Clade 1 formed the sister clade of clades 2–7, that showed unstable relationships sensitive to the 28S alignment strategy and reconstruction method. When only unambiguously aligned regions of 28S were used (28Sc), clades 2 and 7 grouped together but only with moderate support (0.94/56) (Fig. 5). When variable regions (28Sv) were included and aligned manually (MA), the same topology was recovered but with higher support (clade 2 and 7, pp: 1.00/57, in Fig. 3, Supplementary material). The other lineages of *Calathus* (clades 3–6) received moderate support (pp: 0.96). When the CW or MX approaches were used without manual correction, the position of clade 2 changed to become the sister group to a well-supported group of clades 3–7, irrespective of selected alignment gap penalties. Among these clades, only the sister relationship between clade 3 and 7 was resolved with support, albeit low (>0.90) (Figs. 4 and 5, Supplementary material).

PBS values were positive indicating moderate, but generally present, congruence between gene partitions (Fig. 6). These values also indicated that mtDNA data contributed more to the combined tree topology than the other molecular markers (about 4 times more than 28S and about 9 times more than EF-1α) (Fig. 7). Variable regions of the 28S data revealed twice as much contribution to the combined tree topology compared to the unambiguously aligned regions, but this contribution varied considerably according to the alignment strategy (Fig. 1). The PBS contribution from each partition in the phylogeny is shown in Fig. 6. No support was found for some of the deepest nodes (from a distance of 0 to 0.03 from the root, Fig. 6). MtDNA showed the highest contribution to the combined analysis in regions from median depth to the tips of the tree (distances from approximately 0.04–0.06). EF-1α contributed information at median depths (0.035–0.04) and showed low contribution at the tips of the combined data tree. Conserved regions of 28S (28Sc) showed contribution from median depth to the tips, and variable regions of 28S (28Sv) exhibited a wide range of PBS values depending upon gap penalties (mainly positive values, data not shown).

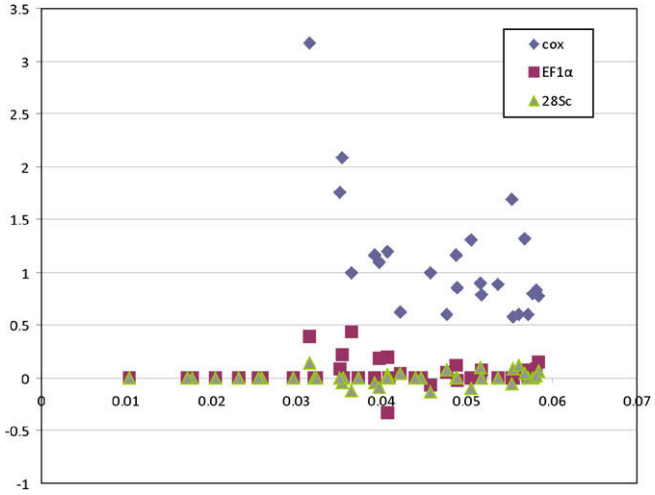


**Fig. 6.** Distribution of corrected Partitioned Bremer Support values of the different markers in relation to the distance of each node from the root (molecular clock enforced). Line at value 0 shows the transition between positive and negative contributions to the combined topology.

### 3.3. Testing for hard polytomies

The LTT plot exhibited a relatively constant slope. To test whether this observation indicates a constant diversification rate (CR), we evaluated our results with two diversification tests. The $\gamma$-statistics showed a positive value ($\gamma = 3.131573$) but failed to reject the model of a constant diversification rate (CR; $p = 0.999$). This result is robust to assumptions about missing taxa; the MCCR test of $\gamma$-statistics is not significant ($p = 1.0$) using different numbers of missing taxa. The BDL test (Table 4) showed a positive value of $\Delta AIC_{RC}$ (1.4166), which indicates that rate-variable models have the best fit to the data; however, the distribution of simulations with $N = 97$ was not significant and the null hypothesis of a constant diversification model at $\alpha = 0.05$ could not be rejected ($p = 0.2561$; Rabosky, 2006a). Simulations based on incomplete taxon sampling also failed to reject the null hypothesis under different levels of taxon inclusion ($p = 0.81$–$0.99$).

Matrix length analyses showed that the proportion of supported nodes increased with increasing matrix length both for new simulated data sets and for replicates of the original data set (Fig. 8). Plotting distance to the root against node support revealed that

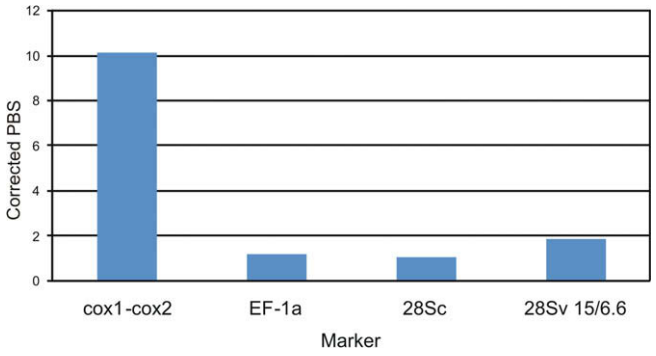


**Fig. 7.** Total corrected PBS contribution of each partition to the combined tree. Each bar represents the contribution of each marker: 28Sc (=conserved region of 28S), 28Sv 15/6 (=variable regions of 28S aligned with ClustalW with default GOP/GEP).

**Table 4**
Results of fitting diversification models to the subtribe Calathina using birth–death likelihood.

| Model | Pure birth | bd | DDL | DDX | yule2rate |
|---|---|---|---|---|---|
| Parameters | r1 = 56.45 | r1 = 28.87<br>a = 0.70 | r1 = 56.45<br>k = 1477084 | r1 = 17.96<br>x = −0.32 | r1 = 47.12<br>r2 = 118.80<br>st = 0.0026 |
| Ln (L) | 633.5398 | 638.7274 | 633.5393 | 636.5 | 640.4358 |
| AIC | −1265.08 | −1273.455 | −1263.079 | −1269 | −1274.872 |
| ΔAIC | −8.375 | 0 | −10.376 | −4.455 | 1.417 |

bd, birth–death model; DDL and DDX, logistic and exponential density-dependent speciation model; yule2rate, multi-rate variant of the Yule model; $r1 - r2$, net diversification rate (speciation event per million year); $a$, extinction fraction; st, time of rate shift (Mya); $k$, parameter in the logistic density-dependent model; $x$, parameter in the density-dependent exponential model; Ln (L), log-likelihood; AIC, Akaike information criterion ; ΔAIC, change in AIC between model and the overall best-fit model.

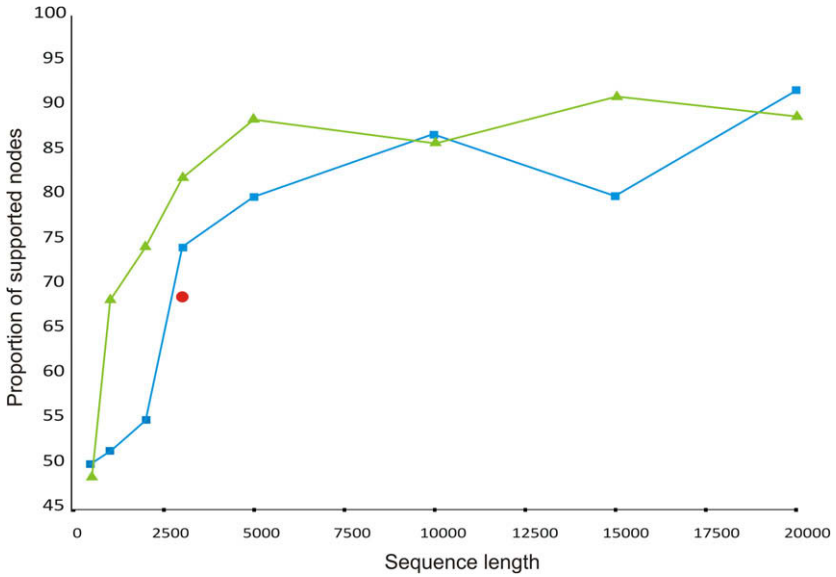


**Fig. 8.** Relationship between sequence length of simulated datasets and proportion of supported nodes (posterior probability). Simulated data sets were built up using new data sets (green line) or pseudoreplicates of the original data set (blue line). Red circle indicates proportion of supported nodes in the combined analysis of the original data set.

low support was randomly distributed within the tree, from deeper to shallower nodes (data not shown).

## 4. Discussion

### 4.1. Phylogenetic relationships within Calathina

Our analyses corroborated the monophyly of the subtribe Calathina and revealed seven well-supported clades that are inferred to represent main lineages within the subtribe. However, low support was obtained for some of the deepest nodes that were associated with short internal branches, and thus the relationships between these lineages could not be fully resolved. Lack of deep nodal resolution is frequently observed in molecular phylogenies (Whitfield and Lockhart, 2007) and may be due to a number of factors such as (i) inappropriate molecular data; (ii) inappropriate phylogenetic methods or substitution models; (iii) insufficient data (Walsh et al., 1999; Whitfield and Lockhart, 2007); (iv) incongruence between data partitions (Baker and DeSalle, 1997; Cognato and Vogler, 2001); or (v) rapid radiation (Rabosky and Lovette, 2008; Whitfield and Kjer, 2008) resulting in 'hard' polytomies (Maddison, 1989).

The first two factors are unlikely to explain our lack of basal resolution because we have used 812 parsimony informative characters from three independent markers that encompass a range of substitution rates, implementing different substitution models for each data partition together with phylogenetic methods that differ in their analytical strengths. The PBS analyses (Fig. 6) revealed that the low support associated with the deepest nodes was not due to incongruence, and a rapid radiation following the origin of Calathina is unlikely because there were no significant departures from a constant speciation rate model in the $\gamma$-statistic and the BDL tests. Although these tests lack power to detect increased diversification rates, they are more effective at detecting a decrease in diversification rates (Rabosky, 2006a), which is expected after rapid radiation events (Rabosky and Lovette, 2008). With regard to insufficient data, both simulations and length-variable bootstrap re-sampling showed that the proportion of supported nodes ($pp > 0.95$) increased with the size of the data matrix (Fig. 8), which makes this hypothesis the most likely explanation for the low resolution of the deepest nodes. More data, preferably from other genes, are therefore needed to resolve the relationship between main clades in Calathina.

Regardless of the weakly resolved basal relationships within Calathina, there was strong support for a paraphyletic *Calathus* that included the current genera *Synuchidius* (clade 2), *Thermoscelis* and *Lindrothius* (clade 7), and *Acalathus* (clade 1). This result agrees well with the ideas of Lindroth (1956) who divided *Calathus* into many different species groups, including the aforementioned genera. Thus, molecular data support the existence of a subtribe Calathina that includes a single inclusive genus *Calathus*, of which *Synuchidius*, *Lindrothius*, *Thermoscelis* and *Acalathus* should be ranked as subgenera. The inclusion of *Acalathus* within the subtribe Calathina instead of Dolichina has also been suggested by Ball and Nègre (1972), Perrault (1977) and Casale (1988).

The close relationship between the North American *Acalathus advena* and the Iberian species *Calathus* (*Iberocalathus*) *rotundatus* (clade 1) was somewhat surprising because of the geographical disjunction between these taxa. However, this relationship was supported independently by mtDNA as well as 28S data. Also, the relatively short branch lengths observed for *Calathus* and *Acalathus* in the 28S data indicate that this relationship is not an artefact due to long-branch attraction (Bergsten, 2005). The Bayesian analyses are particularly suited to handle erroneous relationships that can emerge as a consequence of elevated substitution rates, but none of the implemented evolutionary models that were fitted to different genomic data and different codon positions revealed alternative hypotheses (Figs. 4, 6 and 7, and Fig. 2 of Supplementary material). Spurious inclusion of *Acalathus* in the ingroup was also tested by sampling a much wider range of Calathina (all genera) and Dolichina (4 out of 7 genera) outgroups, but did not result in topological changes (data not shown). Long-branch attraction was also tested by removing each of these putatively wrongly placed taxa, one at a time, but none of them swapped position in the absence of the other. We are therefore confident that *Acalathus* belongs within the genus *Calathus*, a hypothesis that should be further tested by studying the extant species presently found in China. Molecular data also support the inclusion of the genus *Dolichus* within the subtribe Dolichina, as put forward by Casale (1988) and Ruiz et al. (2009).

Clade 2 showed an unexpected relationship between the Balkan *Synuchidius ganglbaueri*, *Calathus metallicus*, the Iberian *C. granatensis*, the North African *C. opacus* and all species of the subgenus *Calathus*. *Synuchidius ganglbaueri* is usually regarded as a separate genus from *Calathus* on morphological grounds (Apfelbeck, 1930; Guéorguiev, 2007). *Calathus metallicus* shows external characters similar to the subgenus *Neocalathus*, and *C. granatensis* and *C. opacus* have been placed either in the subgenera *Calathus* or *Neocalathus* because they show intermediate morphological characters between these subgenera. Nevertheless, at least one morphological character supports this clade by the presence of a tooth within the endophallus in all four taxa. A taxonomic reassessment is therefore required.

Clade 3 was made up by most but not all species of the subgenus *Lauricalathus* inhabiting the island of Tenerife (the *depressus* group), with the European *C.* (*Amphyginus*) *rotundicollis* as sister group (Fig. 9). Nuclear data added in this study thus corroborate a relationship reported by Emerson et al. (2000) and Ruiz and Serrano (2006) based on mitochondrial data only. Two additional species from Tenerife (*C. rectus* and *C. abaxoides*) were on the other hand related to taxa inhabiting other Canary Islands (Gran Canaria, La Gomera, El Hierro, clade 6), including the three La Gomeran species in the subgenus *Trichocalathus*. These Canarian taxa combined made up the sister group to a clade consisting primarily of western Palearctic species of the *ambiguus* group in the subgenus *Neocalathus*, including also one species from Madeira (Fig. 9). The nuclear data added in this study thus corroborate the lack of support for a monophyletic subgenus *Trichocalathus* (see also Emerson et al., 1999), as species in this subgenus from La Gomera are nested within the subgenus *Lauricalathus* on the same island. Generally short branches throughout the La Gomeran clade suggest that the ancestor of these species underwent rapid morphological evolution giving rise to the distinctive traits of *Trichocalathus*. A similar case of fast morphological evolution associated with little molecular divergence has been reported among carabids in the genus *Carabus* (see Osawa et al., 2004 and further references therein).

The lack of monophyly for species currently included in *Neocalathus* (clades 5 and 6) corroborates the recently suggested polyphyletic nature of this subgenus (Ruiz and Serrano, 2006). On morphological grounds *Neocalathus* is relatively uniform and does not currently pose any taxonomic controversies (e.g., Ball and Nègre, 1972; Toribio, 2006), but no cladistic analysis has been carried out on the whole subgenus, and a thorough study to corroborate the polyphyly of this subgenus is required. The Nearctic *Neocalathus* (clade 5) formed a monophyletic group divided into Mexican and North American sister clades (or nearly so, depending on the type of analysis), as proposed by Ball and Nègre (1972) based on morphological characters (Fig. 67, p. 519). We can therefore infer a single colonisation of the Nearctic region by *Neocalathus*, and only two independent colonizations when including *Acalathus*. Further molecular studies on Mexican *C.* (*Tachalus*) *ovipennis* and Asian taxa of *Procalathus* and *Acalathus* are nevertheless needed to achieve deeper insight into the evolutionary history of Nearctic *Calathus*.

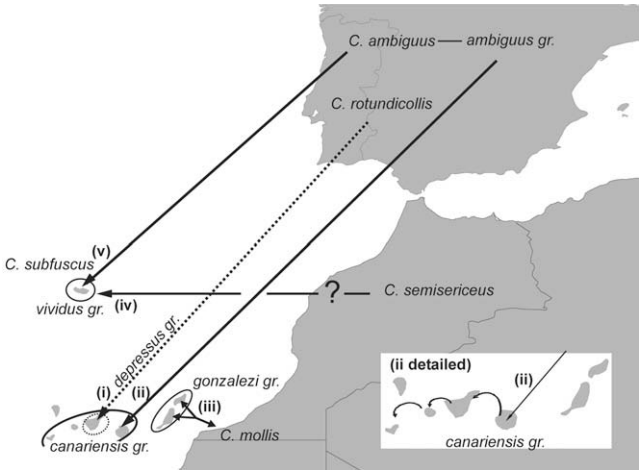


**Fig. 9.** Colonisation events of Macaronesia by subtribe Calathina. Roman numbers refer to colonisation events detailed in the text.

### 4.2. Diversification of the subtribe Calathina in the Mediterranean Basin, Caucasus and Europe

The distribution of taxa within the subtribe is concentrated around the Mediterranean Basin and the Caucasus. The most extensive diversification in this area is found within the subgenus *Calathus*, with about 50 flightless species most of which are endemic to the mountain massifs on the northern side of the Mediterranean Basin. This suggests that isolation in mountains has been a significant factor in driving speciation within this lineage where only the wing dimorphic and eurytopic *C. fuscipes* has a wide Palearctic distribution. The more distantly related species of this clade, *C. granatensis*, *C. opacus*, *C. metallicus* and *Synuchidius ganglbaueri* seem to be peripheral isolates around the Mediterranean Basin, apparently relics of an ancestor that once had a wider distribution. The same explanation holds for the Caucasian taxa *Lindrothius* and *Thermoscelis* (clade 7, Fig. 2). In contrast to the subgenus *Calathus*, the subgenera *Bedelinus* (*C. circumseptus*, clade 4, full winged) and *Amphyginus* (*C. rotundicollis*, clade 3, wing dimorphic) that have similar Mediterranean distributions are much less diverse lineages. *Calathus rotundicollis* has a discontinuous distribution that includes Atlantic Europe, Italy and Greece. The low sequence divergence between Italian and Iberian individuals (*p*-distance cox1-cox2: 0.014) supports a recent origin for this disjunction, possibly during the Pleistocene, as postulated for other organisms (e.g., Hewitt, 2001, 2004).

Members of the *melanocephalus* (Fig. 5, clade 5) and the *ambiguus* (Fig. 5, clade 6) groups in the subgenus *Neocalathus* show a wider geographical distribution, in agreement with their higher dispersal power (they are either winged or wing dimorphic species) and lack of ecological specialisation (they are generalist taxa). These species are not limited to the Mediterranean Basin but occur in temperate and boreal Europe, and have colonised the Macaronesian archipelagos and other areas out of the Palearctic region.

### 4.3. The colonisation of the Macaronesian subregion

The study of a large number of continental taxa and the use of multiple nuclear and mitochondrial markers allowed reconstruction of colonisation events and subsequent radiations in *Calathus* on the Macaronesian archipelagos. New data have not only corroborated previous conclusions by Emerson et al. (1999, 2000) but added interesting clues about the origin of multiple lineages endemic to these archipelagos. Based on these data, at least four or five independent colonisation events of the Macaronesian archipelagos have taken place (Fig. 9).

(i) The *depressus* group, which includes most *Lauricalathus* of Tenerife except for *C. abaxoides* and *C. rectus*, and the European *C. rotundicollis* share a common ancestor, suggesting a West-Palearctic origin of the Tenerifean clade. However, the genetic divergence between *C. rotundicollis* and the *C. depressus* group is relatively large (*p*-distance cox1-cox2: 0.07), and thus extinction of a more closely related taxa from another geographical region is a possibility.

(ii) The European *ambiguous* group consisting of *C. ambiguus*, *C. asturiensis*, *C. erratus* and the Madeiran *C. subfuscus* formed a sister group to a large clade of Canary island species (clade 6, the *canariensis* group). Although the sister relationship is not prima facie evidence for an ancestral European distribution, we note that most other distantly related species are also of western Palearctic origin. It therefore seems likely that the *canariensis* group diversified from a West-Palearctic ancestor. Further diversification within this group possibly occurred in an east to west direction as indicated by the nested structure among Gran Canaria, Tenerife and La Gomera populations and higher taxa. The nested structure is consistent with a stepping-stone model of colonisation from the island closest to the mainland to the more distant islands in the west, as reported for other insect taxa in the Canary Islands (Juan et al., 2000). The most recent colonisation event has possibly occurred between La Gomera and the youngest island of El Hierro (1 Myr old), where the endemic *C. spretus* is closely related to the La Gomera species *C. gomerensis* (not included, see Emerson et al., 2000).

(iii) The eastern Canary Islands lineage (*gonzalezi* group), represented by *C. gonzalezi* (Fuerteventura) and *C. simplicicollis* (Lanzarote), is paraphyletic with respect to the widespread continental species *C. mollis* that is distributed along North Africa and Europe. The paraphyletic distribution makes inference about their colonisation history equivocal. The nested position of the continental species (*C. mollis*) may indicate back colonisation from the eastern Canary Islands to the mainland, a scenario documented in other Macaronesian beetles (e.g., Jordal and Hewitt, 2004) and many other taxa (Bellemain and Ricklefs, 2008), or it may indicate a double colonisation of the Canary Islands by the continental ancestor (Emerson, 2002).

(iv) The origin of the *vividus* group in Madeira (*C. complanatus*, *C. colasianus*, *C. vividus*) is uncertain given its sister relationship with a clade consisting of species from many different geographical areas. The relatively deep genetic divergence of the *vividus* group points to a potentially ancient origin of this lineage.

(v) *C. subfuscus* from Madeira is closely related to the European *C. ambiguus*, which indicates an independent and recent colonisation of this island.

From this minimum estimate of five independent colonisations of the Macaronesian islands we note that a European mainland to Macaronesian island pattern is the most common pattern. This may appear at odds with the geographically closer location of the African mainland. However, the general biogeography of the Macaronesian islands indicates that many (but not all) groups of insects and other organisms have indeed colonised these islands from the mainland north of the Mediterranean Sea. These patterns are usually explained by way of prevailing winds and sea currents from the Iberian Peninsula (Juan et al., 2000).The inferred colonisation patterns in *Calathus* could also be an artefact caused by incomplete sampling, although this seems unlikely as we have studied most North African species of *Calathus* (7 out of 9 species), representing all extant lineages in that region (Antoine, 1957). It is however possible that now extinct ancestors of the present Macaronesian *Calathus* formerly inhabited North Africa, as has been suggested more generally for the flora and fauna of the region (Emerson, 2002).

## Acknowledgments

This work has been supported by projects 00595/PI/04 of the Fundación Séneca (Murcia) and CGL2006-06706/BOS of the Spanish Ministerio de Educación y Ciencia. Carlos Ruiz was granted with FPU predoctoral grant of the Ministerio de Educación y Ciencia. We heartily thank the many colleagues that helped us with collecting material, particularly Carmelo Andújar, Jose Fermín Sánchez-Gea, Esteban Jiménez, José Luís Lencina and Javier Ibáñez (University of Murcia), Boby Gueorguiev (University of Sophia). Roy Danielsson helped with loaned material (Museum of Zoology, University of Lund). Irene Muñoz and Lope Lorenzo helped with Fig. 9. George Ball (University of Alberta) helped with the identification of Nearctic material. Ignacio Ribera (Museo Ciencias Naturales, Madrid) helped with valuable bibliography.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2009.10.026.

## References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Caski, F. (Eds.), Proceedings of the Second International Symposium on Information Theory. Akadémiai Kiadó, Budapest., pp. 267–281.
Antoine, M., 1957. Coléoptères carabiques du Maroc (2éme partie). Mém. Soc. Sci. Nat. Phys. Maroc (Zool.) 3, 178–314.
Apfelbeck, V., 1930. Fauna insectorum balcanica VII. 2 teil. Neue Beiträge zur Systematischen Insektenkunde 4, 176.
Baker, R.H., DeSalle, R., 1997. Multiple sources of character information and the phylogeny of Hawaiian Drosophilids. Syst. Biol. 46, 654–673.
Ball, G.E., Nègre, J., 1972. The taxonomy of the Nearctic species of the genus *Calathus* Bonelli (Coleoptera: Carabidae: Agonini). Trans. Am. Entom. Soc. 98, 412–533.
Barraclough, T.G., Vogler, A.P., 2002. Recent diversification rates in North American tiger beetles estimated from a dated mtDNA phylogenetic tree. Mol. Biol. Evol. 19, 1706–1716.
Bellemain, E., Ricklefs, R.E., 2008. Are islands the end of the colonization road? Trends Ecol. Evol. 23, 461–468.
Berbee, M.L., Carmean, D.A., Winka, K., 2000. Ribosomal DNA and resolution of branching order among the Ascomycota: how many nucleotides are enough? Mol. Phyl. Evol. 17, 337–344.
Bergsten, J., 2005. A review of long-branch attraction. Cladistics 21, 163–193.
Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295–304.
Casale, A., 1988. Revisione degli Sphodrina (Coleoptera, Carabidae, Sphodrini). Monographie V. Museo Regionale di Scienze Naturali Torino, Torino.
Cognato, A.I., Vogler, A.P., 2001. Exploring data interaction and nucleotide alignment in a multiple gene analysis of *Ips* (Coleoptera: Scolytinae). Syst. Biol. 50, 758–780.
DeBry, R.W., 2001. Improving interpretation of the decay index for DNA sequence data. Syst. Biol. 50, 742–752.
de Queiroz, A., Lawson, R., Lemos-Espinal, J.A., 2002. Phylogenetic relationships of North American garter snakes (*Thamnophis*) based on four mitochondrial genes: how much DNA sequence is enough? Mol. Phyl. Evol. 22, 315–329.
Drummond, A., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 12–15.
Egan, A.N., Crandall, K.A., 2008. Divergence and diversification in North American Psoraleeae (Fabaceae) due to climate change. BMC Biol. 6, 55.
Emerson, B.C., Oromí, P., Hewitt, G.M., 1999. MtDNA phylogeography and recent intra-island diversification among Canary Island *Calathus* beetles. Mol. Phylogenet. Evol. 13, 149–158.
Emerson, B.C., Oromi, P., Hewitt, G.M., 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.
Emerson, B.C., 2002. Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. Mol. Ecol. 11, 951–966.
Fishbein, M., Hibsch-Jetter, C., Soltis, D.E., Hufford, L., 2001. Phylogeny of Saxifragales (Angiosperms, Eudicots): analysis of a rapid, ancient radiation. Syst. Biol. 50, 817–847.
Gillespie, J., Cannone, J., Gutell, R., Cognato, A., 2004. A secondary structural model of the 28S rRNA expansion segments D2 and D3 from rootworms and related leaf beetles (Coleoptera: Chrysomelidae; Galerucinae). Insect Mol. Biol. 13, 495–518.
Guéorguiev, B.V., 2007. Annotated Catalogue of the Carabid Beetles of Albania (Coleoptera: Carabidae). Pensoft, Sofia.
Hines, H.M., 2008. Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). Syst. Biol. 57, 58–75.
Hewitt, G.M., 2001. Speciation, hybrid zones and phylogeography – or seeing genes in space and time. Mol. Ecol. 10, 537–549.
Hewitt, G.M., 2004. Genetic consequences of climatic oscillations in the Quaternary. Phil. Trans. R. Soc. B (Biol. Sci.) 359, 183–195.
Hovorka, O., Sciaky, R., 2003. Tribe Sphodrini. In: Löbl, I., Smetana, A. (Eds.), Catalogue of Palearctic Coleoptera. Apollo Books, Stenstrup, pp. 521–547.

Janssens, S.B., Knox, E.B., Huysmans, S., Smets, E.F., Merckx, V.S.F.T., 2009. Rapid radiation of *Impatiens* (Balsaminaceae) during pliocene and pleistocene: result of a global climate change. Mol. Phyl. Evol. 52, 806–824.

Jordal, B.H., Hewitt, G.M., 2004. The origin and radiation of Macaronesian beetles breeding in *Euphorbia*: The relative importance of multiple data partitions and population sampling. Syst. Biol. 53, 711–734.

Juan, C., Emerson, B.C., Oromi, P., Hewitt, G.M., 2000. Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. Trends Ecol. Evol. 15, 104–109.

Kadereit, J.W., Griebeler, E.M., Comes, H.P., 2004. Quaternary diversification in European alpine plants: pattern and process. Phil. Trans. R. Soc. B (Biol. Sci.) 359, 265–274.

Kjer, K.M., 1995. Use of ribosomal-RNA secondary structure in phylogenetic studies to identify homologous positions – an example of alignment and data presentation from the frogs. Mol. Phylogenet. Evol. 4, 314–330.

Kozak, K.H., Weisrock, D.W., Larson, A., 2006. Rapid lineage accumulation in a non-adaptive radiation: phylogenetic analysis of diversification rates in eastern North American woodland salamanders (Plethodontidae: *Plethodon*). Proc. R. Soc. Lond. B. (Biol. Sci.) 273, 539–546.

Kumar, S., Filipski, A., 2007. Multiple sequence alignment: in pursuit of homologous DNA positions. Genome Res. 17, 127–135.

Lindroth, C.H., 1956. A revision of the genus *Synuchus* Gyllenhal (Coleoptera, Carabidae) in the widest sense, with notes on *Pristosia* and *Calathus*. Trans. R. Soc. Lond. 108, 485–585.

Lorenz, W., 2005. A Systematic List of Extant Ground Beetles of the World (Coleoptera "Geadephaga": Trachypachidae and Carabidae incl. Paussinae, Cicindelinae, Rhysodinae). Tutzing, Germany, 530p.

Maddison, W.P., Maddison, D.R., 2007. Mesquite: a modular system for evolutionary analysis. Version 2.01. Available from: <http://mesquiteproject.org>.

Maddison, W.P., 1989. Reconstructing character evolution on polytomous cladograms. Cladistics 5, 365–377.

McKenna, D.D., Farrell, B.D., 2006. Tropical forests are both evolutionary cradles and museums of leaf beetle diversity. Proc. Natl. Acad. Sci. USA 103, 10947–10951.

Osawa, S., Su, Z.-H., Imura, Y., 2004. Molecular Phylogeny and Evolution of Carabid Ground Beetles. Springer, Tokyo.

Paradis, E., 1997. Assessing temporal variations in diversification rates from phylogenies: estimation and hypothesis testing. Proc. R. Soc. Lond. B. (Biol. Sci.) 264, 1141–1147.

Perrault, G.G., 1977. Le sous-genre *Lindrothius* (Kurnakov) en Asie mineure Col. Carabidae – genus *Calathus* (Bon.). Bull. Mens. Soc. Linn. Lyon 46, 220–224.

Philippe, H., Delsuc, F., Brinkmann, H., Lartillot, N., 2005. Phylogenomics. Ann. Rev. Ecol. Evol. Syst. 36, 541–562.

Posada, D., Crandall, K.A., 1998. MODEL TEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Posada, D., Buckley, T., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.

Putzeys, J., 1873. Monographie des Calathides. Ann. Soc. Entomol. Belgique 16, 19–96.

Pybus, O.G., Harvey, P.H., 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. Proc. R. Soc. Lond. B. (Biol. Sci.) 267, 2267–2272.

Rabosky, D.L., 2006a. Likelihood methods for detecting temporal shifts in diversification rates. Evolution 60, 1152–1164.

Rabosky, D.L., 2006b. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. Evol. Bioinform Online 2, 247–250. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2674670>.

Rabosky, D.L., Donnellan, S.C., Talaba, A.L., Lovette, I.J., 2007. Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. Proc. R. Soc. Lond. B. (Biol. Sci.) 274, 2915–2923.

Rabosky, D.L., Lovette, I.J., 2008. Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? Evolution 62, 1866–1875.

Rambaut, A., Drummond, A., 2005. Tracer. MCMC Trace File Analyser. Version 1.3. University of Oxford, Oxford.

Roelants, K., Gower, D.J., Wilkinson, M., Loader, S.P., Biju, S.D., Guillaume, K., Moriau, L., Bossuyt, F., 2007. Global patterns of diversification in the history of modern amphibians. Proc. Natl. Acad. Sci. USA 104, 887–892.

Rokas, A., Williams, B.L., King, N., Carrol, S.B., 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425, 798–804.

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

Ruiz, C., Serrano, J., 2006. Molecular phylogeny and systematics of *Calathus* Bonelli (Coleoptera : Carabidae : Sphodrini) based on mitochondrial cox1-cox2 sequences. Entomol. Fenn. 17, 214–220.

Ruiz, C., Jordal, B., Serrano, J., 2009. Molecular phylogeny of the tribe Sphodrini (Coleoptera: Carabidae) based on mitochondrial and nuclear markers. Mol. Phylogenet. Evol. 50, 44–58.

Schatzmayr, A., 1937. I *Calathus* d'Europa. Pubblicazioni del Museo Entomologico "Pietro Rosi". Duino 2, 1–49.

Spinks, P., Thomson, R., Lovely, G., Shaffer, H.B., 2009. Assessing what is needed to resolve a molecular phylogeny: simulations and empirical data from Emydid turtles. BMC Evol. Biol. 9, 56.

Sorenson, M.D., Franzosa, E.A., 2007. TreeRot, Version 3. Boston University, Boston.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Ver. 4. Sinauer, Sunderland.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Toribio, M., 2006. Notas sobre el género *Calathus* Bonelli, 1810, en la Península Ibérica (Coleoptera, Carabidae). Bull. Soc. Ent. France 111, 51–57.

Turgeon, J., Stoks, R., Thum, R.A., Brown, J.M., Mcpeek, M.A., 2005. Simultaneous Quaternary radiations of three damselfly clades across the Holarctic. Am. Nat. 165, E78–E107.

Walsh, H.E., Kidd, M.G., Moum, T., Friesen, V.L., 1999. Polytomies and the power of phylogenetic inference. Evolution 53, 932–937.

Weir, J.T., 2006. Divergent timing and patterns of species accumulation in lowland and highland neotropical birds. Evolution 60, 842–855.

Whitfield, J.B., Kjer, K.M., 2008. Ancient rapid radiations of insects: challenges for phylogenetic analysis. Ann. Rev. Ent. 53, 449–472.

Whitfield, J.B., Lockhart, P.J., 2007. Deciphering ancient rapid radiations. Trends Ecol. Evol. 22, 258–265.

Zou, X.H., Zhang, F.M., Zhang, J.G., Zang, L.L., Tang, L., Wang, J., Sang, T., Ge, S., 2008. Analysis of 142 genes resolves the rapid diversification of the rice genus. Genome Biol. 9 (3), R49.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406–3415.