Molecular systematics and evolution of the subgenus *Mesocarabus* Thomson, 1875 (Coleoptera: Carabidae: *Carabus*), based on mitochondrial and nuclear DNA

CARMELO ANDÚJAR[1]*, JESÚS GÓMEZ-ZURITA[2], JEAN-YVES RASPLUS[3] and JOSÉ SERRANO[1]

[1]*Departamento de Zoología y Antropología Física, Facultad de Veterinaria, Universidad de Murcia, 30071 Murcia, Spain*
[2]*Institut de Biologia Evolutiva (CSIC-UPF), Pg. Marítim de la Barceloneta 37, 08003 Barcelona, Spain*
[3]*Montpellier SupAgro, CIRAD, IRD, INRA,CBGP,UMR 1062, F-34988 Montferrier Sur Lez, France*

The subgenus *Mesocarabus* Thomson, 1875 is a western Palaearctic group that currently includes five species: four of them inhabiting western Europe (*Carabus lusitanicus* Fabricius, 1801, *Carabus problematicus* Herbst, 1786, *Carabus dufourii* Dejean & Boisduval, 1829, and *Carabus macrocephalus* Dejean, 1826) and one found in the Rif Mountains in northern Morocco (*Carabus riffensis* Fairmaire, 1872). Representatives of *Mesocarabus* have been included in previous molecular phylogenetic studies, but taxon- or gene-sampling limitations yielded inconclusive results regarding its monophyly and sister relationship. Here we perform molecular phylogenetic analyses based on five mitochondrial (3625 nt) and eight nuclear (5970 nt) genes sequenced in many *Mesocarabus* populations, and in related western Palaearctic *Carabus* Linnaeus, 1758. We conducted parsimony, maximum-likelihood, and Bayesian analyses and found a well-supported sister relationship between a monophyletic *Mesocarabus* with Iberian species of the subgenus *Oreocarabus* Géhin, 1876. Within *Mesocarabus*, the European species form a monophyletic lineage sister to Moroccan *C. riffensis*. A time-calibrated phylogeny suggests the split between *Mesocarabus* and *Oreocarabus* occurred at 11.8 Mya (95% highest posterior density, HPD, 8.7–15.3 Mya), and the divergence between *C. riffensis* and European *Mesocarabus* at 9.5 Mya (95% HPD 7.0–12.5 Mya). The early diversification of *Mesocarabus* and related subgenera during the Miocene, and alternative hypotheses concerning the origin of *Mesocarabus* in the Iberian Peninsula and the Betic-Riffian plate are discussed using calibration data and dispersal–vicariance biogeographic analyses. Finally, we found instances of incongruence between mitochondrial DNA and nuclear-based phylogenies of *Mesocarabus*, which are hypothesized to be the result of introgressive hybridization.

INTRODUCTION

Mesocarabus Thomson, 1875 (Coleoptera: Carabidae: Carabini) is a well-delimited subgenus among the highly diverse *Carabus* Linnaeus, 1758, based on morphological characters of the adult specimens (Breuning, 1932–1937; Turin, Penev & Casale, 2003; Deuve, 2004), and currently comprises five western Palaearctic species: four in western Europe (*Carabus lusitanicus* Fabricius, 1801; *Carabus problematicus* Herbst, 1786, *Carabus dufourii* Dejean & Boisduval, 1829, and *Carabus macrocephalus* Dejean, 1826) and

*Corresponding author. E-mail: candujar@um.es

one in North Africa (*Carabus riffensis* Fairmaire, 1872). The systematic placement of *Mesocarabus* among *Carabus* has been controversial. Bengtsson (1927) placed *Mesocarabus* within the division Metacarabi based on larval morphology, Ishikawa (1978) and Deuve (1994) included it within the Lobifera division based on endophallic morphology, and it was considered as an independent section, the Mesocarabigenici, based on the *nd5* mitochondrial gene (Imura, 2002). Significant advances in the systematics of *Carabus* have resulted from recent molecular studies (Prüser, 1996; Imura, 2002; Su *et al*., 2003; Osawa, Su & Imura, 2004; Sota & Ishikawa, 2004; Andújar, Serrano & Gómez-Zurita, 2012; Deuve *et al*., 2012). The most extensive approach to date was based on a phylogeny of the mitochondrial *nd5* gene, compiled in Osawa *et al*. (2004). These authors revealed an early explosive radiation of *Carabus* associated with a basal polytomy in the *nd5* phylogeny (Su, Imura & Osawa, 2001), and supported the splitting of the genus into 137 subgenera grouped in 29 sections of uncertain relationships; *Mesocarabus* was the single member of section Mesocarabogenicici (Imura, 2002; Osawa *et al*., 2004). Sota & Ishikawa (2004) in turn obtained a well-resolved phylogeny of *Carabus* based on the study of two nuclear genes, and provided fair resolution to the lineage splits in the early evolution of the genus. More recently, Deuve *et al*. (2012) analysed several genes and a worldwide representative *Carabus* sampling, and obtained a well-resolved phylogeny, with important inconsistencies between mitochondrial and nuclear gene-based trees. Despite these attempts, which included representatives of *Mesocarabus* and potentially related taxa, neither the monophyly of *Mesocarabus* nor its sister relationship were fully resolved. In an early study by Prüser (1996), *Mesocarabus* was retrieved as monophyletic, but hypothetic closest relatives based on morphological evidence were missing from his analyses. *Orinocarabus* Kraatz, 1878 was retrieved as sister taxon of *Mesocarabus* using the mitochondrial *nd5* gene, albeit with low support (Su *et al*., 2003). Sota & Ishikawa (2004) also found this sister relationship, but *Mesocarabus* was recovered as paraphyletic, to include *Carabus (Oreocarabus) amplipennis* Lapouge, 1924, the same taxon that has been found as sister to *Mesocarabus* in Deuve *et al*. (2012).

In order to clarify the systematic placement of *Mesocarabus* in the broader context of the diversification of *Carabus*, we have conducted a phylogenetic approach based on five mitochondrial and eight nuclear gene fragments, including at least one representative of each of the eight main divisions of the genus *Carabus* proposed by Deuve (2004). Several hypotheses for the systematics of *Mesocarabus* are tested: (1) the subgenus *Mesocarabus* is monophyl-

etic; (2) the subgenus *Orinocarabus* is the sister taxon of *Mesocarabus*, as found by Su *et al*. (2003) and Sota & Ishikawa (2004); (3) the Moroccan *C. (Mesocarabus) riffensis* is the sister taxon to European *Mesocarabus*, as postulated by Prüser (1996). Finally, phylogenetic inference is used to investigate the age of the group using Bayesian methods and to explore biogeographic patterns in its early diversification using dispersal–vicariance biogeographic analyses.

MATERIAL AND METHODS
SPECIES AND GENE SAMPLING

The sampling available for this study is shown in Table 1. It includes 22 *Mesocarabus* populations ranging from the Rif Mountains to northern European localities (Fig. 1), representing the five currently valid species in the subgenus (Serrano, 2003; Deuve, 2004), as well as several taxa belonging to other subgenera postulated to be related to *Mesocarabus* based on the analysis of morphology and/or previous molecular phylogenies: two species of *Orinocarabus* (Su *et al*., 2003; Sota & Ishikawa, 2004); three species of Iberian *Oreocarabus* Géhin, 1876, a subgenus closely related to *Orinocarabus*; and species of western European lineages more distantly related (Fig. 2; Table 1). A deeper insight into the relationships of *Mesocarabus* within the whole genus *Carabus* was explored by including at least one representative of each of the eight main divisions proposed by Deuve (2004), as well as three species of the genus *Calosoma* Weber, 1801 to root the trees (Table 1). Thirty-two specimens were extracted for this study using the Dneasy Blood and Tissue kit (Qiagen, Hilden, Germany) and Invisorb Spin Tissue Mini Kit (Invitek, Berlin, Germany), following the manufacturers' instructions. These data were completed with those of the twenty-eight specimens included in the study by Andújar *et al*. (2012) and supplemented with data of 11 taxa retrieved from public sequence databases.

The data matrix included sequences from 11 DNA fragments belonging to nine different ribosomal and protein coding genes from mitochondrial (*nd5*, *cox1-a*, *cox1-b*, *cob*, and *rrnL*) and nuclear (*SSU*, *LSU-a*, *LSU-b*, *HUWE1*, *ITS2*, and *TOP*) genomes, with a total aligned length of 8525 nt. Polymerase chain reactions (PCRs) were made using PuReTaq Ready-To-Go PCR beads (GE Healthcare, UK) or Qiagen Taq Polymerase, with 39 cycles using 50–52 °C as the annealing temperature. The primers used for each gene fragment are given in Table S1. Both strands of the PCR products were sequenced with the same primers used for PCR by Macrogen Inc. (Seoul, Korea) and the 'Centre National de Séquençage' (Genoscope

Table 1. Species of *Carabus* and *Calosoma* (out-group) investigated, with data on specimen collection locality and voucher reference

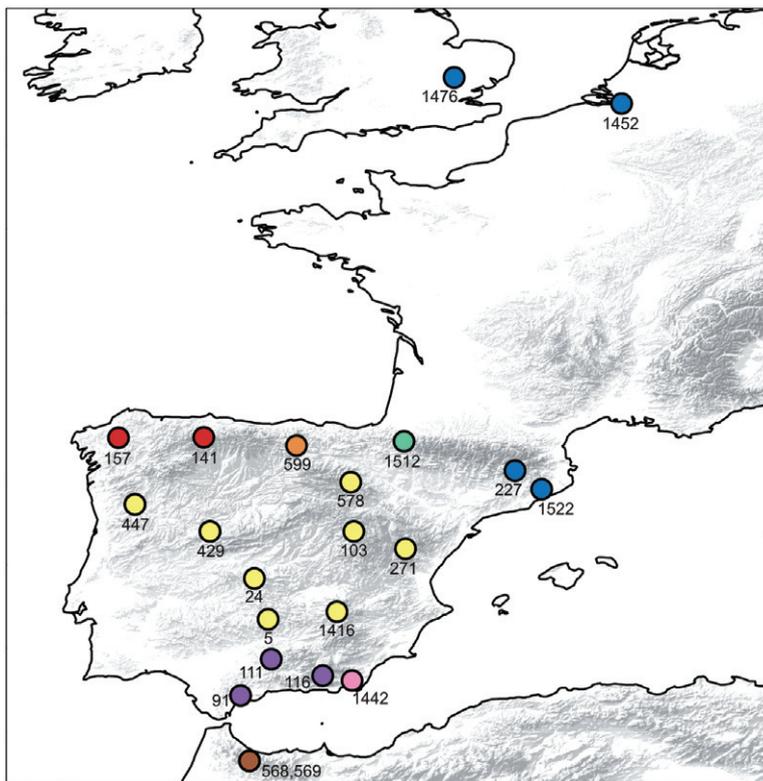| Voucher	Species	Locality	Lat.	Long.
1603-TURQ	*Calosoma (Callisthenes) breviusculus*	Susuz, Kars, Turkey	40.86	43.03
1590-CALO	*Calosoma (Calosoma) sycophanta*	Arroyo de Santiago, Nerpio, Albacete, Spain	38.07	−2.50
1601-TURQ	*Calosoma (Campalita) auropunctatum*	Susuz, Kars, Turkey	40.86	43.03
323-PEMA	*Carabus (Archicarabus) steuartii*	Penamá, Allariz, Orense, Spain	42.16	−7.81
1537-SAOU	*Carabus (Archicarabus) nemoralis*	Foret de Saou, Drome, France	44.66	5.12
1549-RONC	*Carabus (Archicarabus) nemoralis*	Roncesvalles, Navarra, Spain	43.03	−1.30
1548-SENY	*Carabus (Chrysocarabus) rutilans*	Montseny, Barcelona, Spain	41.75	2.43
1615-ROMA	*Carabus (Chrysocarabus) auronitens*	Resita, Romania	45.32	21.85
1553-GALI	*Carabus (Eucarabus) arvensis deyrollei*	Fuentes del Miño, Lugo, Spain	43.24	−7.31
1606-MORR	*Carabus (Eurycarabus) faminii*	Bab Berret, Morocco	34.99	−4.85
1625-EURY	*Carabus (Eurycarabus) faminii*	El Alia, Tunisia	37.18	−10.03
1614-TURQ	*Carabus (Procrustes) coriaceus*	Oysu, Altintas, Turkey	38.96	29.89
1600-TURQ	*Carabus (Limnocarabaus) clatratus*	Susuz, Kars, Turkey	40.86	43.03
1584-MAZA	*Carabus (Macrothorax) morbillosus*	Mazarrón, Murcia, Spain	37.63	−1.19
1585-MAZA	*Carabus (Macrothorax) morbillosus*	Mazarrón, Murcia, Spain	37.63	−1.19
1599-SMAR	*Carabus (Macrothorax) rugosus*	Facinas, Cádiz, Spain	36.15	−5.61
1609-KSAR	*Carabus (Macrothorax) rugosus*	Ksar-el-Kebir, Morocco	34.9	−5.80
1623-TUNB	*Carabus (Macrothorax) morbillosus*	Oued El Bragate, Bazina, Tunisia	36.92	9.37
1624-TUNC	*Carabus (Macrothorax) morbillosus*	El Alia, Tunisia	37.18	−10.03
91-ROSA	*Carabus (Mesocarabus) dufourii*	San pedro de Alcantara, Málaga, Spain	36.62	−5.08
111-ZUHE	*Carabus (Mesocarabus) dufourii*	Zuheros, Córdoba, Spain	37.53	−4.31
116-PRAG	*Carabus (Mesocarabus) dufourii*	P. de la Ragua, Bayarcal, Almería, Spain	37.11	−3.03
578-PONC	*Carabus (Mesocarabus) lusitanicus aragonicus*	Puerto de Oncala, Oncala, Soria, Spain	41.95	−2.33
1442-ALHA	*Carabus (Mesocarabus) lusitanicus baguenai*	Sierra Alhamila, Níjar, Almería, Spain	36.99	−2.30
103-VALC	*Carabus (Mesocarabus) lusitanicus helluo*	Villanueva de Alcorón, Guadalajara, Spain	40.72	−2.25
271-JAVA	*Carabus (Mesocarabus) lusitanicus helluo*	Camarena de la Sierra, Teruel, Spain	40.1	−1.01
5-VESC	*Carabus (Mesocarabus) lusitanicus latus*	Fuencaliente, Ciudad Real, Spain	38.52	−4.39
24-ELVI	*Carabus (Mesocarabus) lusitanicus latus*	El Viezo, Los Navalucillos, Toledo, Spain	39.54	−4.73
1416-VIFU	*Carabus (Mesocarabus) lusitanicus latus*	Villanueva de la Fuente, Ciudad Real, Spain	38.71	−2.67
429-LVEG	*Carabus (Mesocarabus) lusitanicus lusitanicus*	Las Veguillas, Salamanca, Spain	40.72	−5.84
447-VILA	*Carabus (Mesocarabus) lusitanicus lusitanicus*	Vila Real, Portugal, Spain	41.39	−7.71
599-PLUN	*Carabus (Mesocarabus) macrocephalus barcelecoanus*	Portillo de Lunada, Burgos, Spain	43.17	−3.65
157-FORO	*Carabus (Mesocarabus) macrocephalus cantabricus*	Foro, La Coruña, Spain	43.05	−8.12
141-PVEN	*Carabus (Mesocarabus) macrocephalus macrocephalus*	P. de Ventana, Asturias, Spain	43.06	−6.00
227-PTRA	*Carabus (Mesocarabus) problematicus*	S. del Serrat, Vallcebre, Barcelona, Spain	42.23	1.77
1452-KALM	*Carabus (Mesocarabus) problematicus*	Kalmthout, Antwerp, Belgium	51.4	4.43
1476-ENGL	*Carabus (Mesocarabus) problematicus*	Saffron Walder, England	52.05	0.25
1512-OCHA	*Carabus (Mesocarabus) problematicus*	Ochagavía, Navarra, Spain	42.97	−1.00
1522-SENY	*Carabus (Mesocarabus) problematicus*	Montseny, Barcelona, Spain	41.77	2.44
568-KETA	*Carabus (Mesocarabus) riffensis*	Bab Berret, Morocco	34.99	−4.85
569-KETA	*Carabus (Mesocarabus) riffensis*	Bab Berret, Morocco	34.99	−4.85
432-PSAH	*Carabus (Morphocarabus) monilis*	Puerto de Sahun, Huesca, Spain	42.57	0.41
1538-SAOU	*Carabus (Morphocarabus) monilis*	Foret de Saou, Drome, Francia	44.66	5.12
44-BABA	*Carabus (Nesaeocarabus) abbreviatus*	Barranco de Badajoz, Tenerife, Spain	28.30	−16.43
1588-TENE	*Carabus (Nesaeocarabus) abbreviatus*	San José de los Llanos, Tenerife, Spain	28.33	−16.78
35-SSVI	*Carabus (Oreocarabus) guadarramus*	S. San Vicente, Navamorcuende, Toledo, Spain	40.15	−4.74
81-NAVA	*Carabus (Oreocarabus) ghiliani*	Navacerrada, Madrid, Spain	40.79	−4.01
835-PBAR	*Carabus (Oreocarabus) guadarramus*	El Barrancazo, Alcaraz, Albacete, Spain	38.57	−2.38
1200-PIVI	*Carabus (Oreocarabus) guadarramus*	La Vidriera, Huescar, Granada, Spain	38.07	−2.52
1228-SGUI	*Carabus (Oreocarabus) guadarramus*	Puerto del Pinar, Huescar, Granada, Spain	38.04	−2.56
1527-SIES	*Carabus (Oreocarabus) amplipennis*	Serra de Estrella, Portugal	40.38	−7.63
1528-MONC	*Carabus (Oreocarabus) guadarramus*	Moncayo, Lituénigo, Zaragoza, Spain	41.79	−1.81
1618-ORIN	*Carabus (Orinocarabus) concolor*	Val d'Aoste, Italy	45.73	7.40
1619-ORIN	*Carabus (Orinocarabus) fairmairei*	Piedmont, Italy	−	−
1616-ROMA	*Carabus (Platycarabus) irregularis*	Resita, Romania	45.32	21.85
37-ROBU	*Carabus (Rhabdotocarabus) melancholicus*	Los Navalucillos, Toledo, Spain	39.57	−4.71
1593-TIDI	*Carabus (Rhabdotocarabus) melancholicus*	Ketama, Morocco	34.91	−4.57
1597-EALM	*Carabus (Rhabdotocarabus) melancholicus*	Facinas, Cádiz, Spain	36.16	−5.65
1617-ROMA	*Carabus (Tachypus) cancellatus*	Resita, Romania	45.32	21.85
1621-TACH	*Carabus (Tachypus) cancellatus*	Ariège, France	−	−

Figure 1. Sampling localities of *Mesocarabus* specimens used in this study and identified by voucher number, as listed in Table 1. Colour code: brown, *Carabus riffensis*; red, *Carabus macrocephalus*; orange, *Carabus macrocephalus barcelecoanus*; purple, *Carabus dufourii*; yellow, *Carabus lusitanicus*; pink, *Carabus lusitanicus baguenai*; blue, *Carabus problematicus*; and green, *Carabus problematicus*, from Ochagavía.
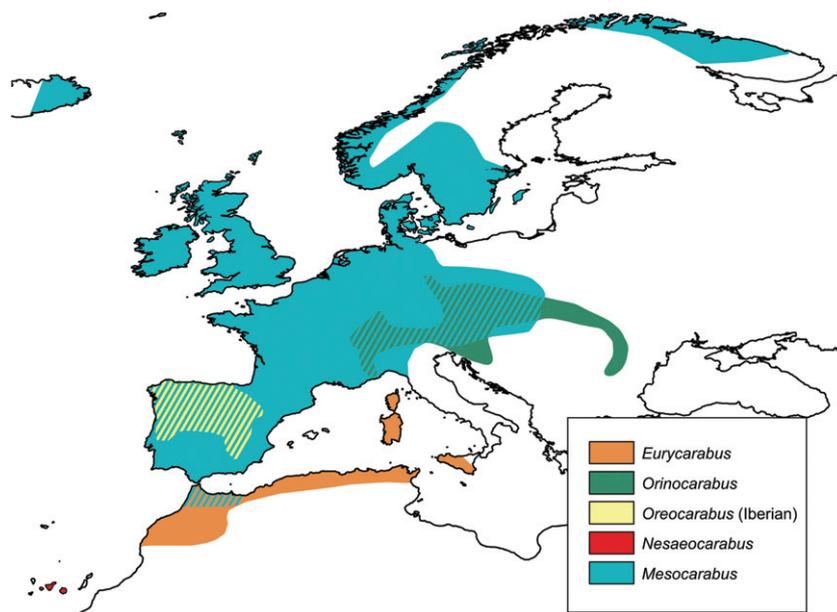

Figure 2. Distribution map of some *Carabus* lineages within the Metacarabi in the western Palaearctic region.

project, France). Additionally, some sequences of *nd5*, *cox1-a*, *HUWE1*, *WINGLESS*, and *PECPK* were obtained from public sequence databases (Table S2).

DNA SEQUENCE ALIGNMENT

Sequences were aligned using the online version of MAFFT 6.240 (Katoh *et al.*, 2002; Katoh, Asimenos & Toh, 2009), with the L-INS-i algorithm for the coding protein genes and Q-INS-i for ribosomal fragments (Katoh & Toh, 2008), a structural-aided alignment algorithm shown to outperform non-structural methods (Letsch *et al.*, 2010). The correct translation to amino acids for protein coding genes was checked in MEGA 4 (Tamura *et al.*, 2007). Heterozygous positions in individual nuclear sequences were coded with IUPAC ambiguity symbols. Concatenated matrices were obtained by combining: (1) five mitochondrial gene fragments (*MIT*: 60 taxa, 3625 nt); (2) six nuclear fragments (*NUC*: 60 taxa, 4900 nt); and (3) all sequenced gene regions (*ALL-A*: 60 taxa, 8525 nt). We generated an additional data set including 11 taxa, with *nd5*, *cox1-a*, *HUWE1*, *WG*, and *PEPCK* sequences retrieved from GenBank (*ALL-B*: 71 taxa, 9595 nt). The latter data set included some combined conspecific sequences from different studies, except for *Calosoma*, which required the combination of data from different taxa within the same subgenus.

PHYLOGENETIC ANALYSES

Data matrices were analysed with parsimony (MP), maximum-likelihood (ML), and Bayesian inference (BI) phylogenetic methods. MP searches were performed with TNT 1.1 (Goloboff, Farris & Nixon, 2003), based on routine searches with 10 000 replicates of random sequence additions, using the tree bisection reconnection (TBR) branch-swapping algorithm, and saving up to 500 trees per replicate. The strict consensus of all most-parsimonious trees was selected as the best phylogenetic hypothesis. Support values were calculated with 10 000 bootstrap pseudoreplicates, each one with a routine search including ten replicates of random additions of taxa, the TBR algorithm, and saving the 50 best trees per replicate. MP analyses were conducted with combined data sets without specifying partitions. ML trees were obtained using RAxML 7.0.4 (Stamatakis, 2006). Combined data sets were partitioned by gene, and protein-coding genes were additionally partitioned, considering first and second codon positions together, and third codon position as an independent partition. An independent GTR + I + G model was applied to each data partition. The best scoring ML tree was selected from 100 inferences on the original alignment with different randomized MP starting trees, as conducted with the rapid hill-climbing algorithm (–d option; Stamatakis

et al., 2007). Support values were obtained with 1000 bootstrap replicates (–i and –b options; Felsenstein, 1985). MP and ML analyses were conducted only for combined data sets. BI was run in MrBayes 3.1 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003) for each individual gene and for combined data sets. Combined data were partitioned by gene, and protein-coding genes, considered individually and in combined matrices, were analysed considering two partitions as before. For each partition the optimal substitution model (Table 2) was selected using the Akaike information criterion in jModelTest (Posada & Buckley, 2004; Posada, 2008). BI consisted of two independent runs, each with three hot and one cold chain, for 10 million generations for individual gene fragments and 20 million generations for the combined data sets, whereby trees were sampled every 500 generations. The standard deviation of split frequencies was checked to assess the convergence of results, as well as the mean and effective sampled size (ESS) of likelihood values computed with TRACER 1.5 (Rambaut & Drummond, 2007). The 50% majority rule and strict consensus trees were calculated, excluding 10% of the initial trees, after the plateau in tree likelihood values had been reached. Trees were visualized using FigTree 1.1.2 (Rambaut, 2008), and node posterior probabilities were interpreted as support values. BI is known to occasionally produce incorrect long-branch estimates, at least for partitioned data sets, because of rate heterogeneity among partitions (Marshall, Simon & Buckley, 2006; Brown *et al.*, 2010; Marshall, 2010). These flawed estimates are stable across independent runs with a set of fixed parameters, but also when Bayesian priors are modified, hindering their detection. To minimize this analytical problem, BI was conducted with the 'ratepr = variable' command in MrBayes to accommodate among-partition rate variation, as recommended by Marshall *et al.* (2006). We also checked for suspicious long branches derived from Bayesian analyses (MrBayes and BEAST) by estimating the 95% highest posterior density (HPD) depth interval for the node of the most recent common ancestor (MRCA) of *Carabus*, confirming whether this interval included the branch length resulting from the best RAxML tree (Spinks & Shaffer, 2009).

We have conducted a partition homogeneity test (PHT; Farris *et al.*, 1995; Swofford, 2003) between all gene pairs studied, and also for the *MIT*, *NUC*, and *ALL-B* combined data set, partitioning by gene, and for the *ALL-B* data set, partitioning by mitochondrial and nuclear genome. PHT analyses were run in PAUP* 4.0 (Swofford, 2003) with 1000 replicates, each of them with ten parsimony searches with the initial random addition of taxa, excluding invariant positions, and saving a single optimal tree per repli-

Table 2. Information on individual gene, codon partition, and combined data sets from aligned sequence data in *Carabus*

Dataset	N	Length	Gapped positions (%)	%GC	Const. sites	Inf. sites	T_s/T_v	MODEL
mtDNA (*MIT*)	60	3625	0.33	27.7	2337	1091	1.4	HKY+I+G
Protein coding								
nd5	69	891	0	23	491	299	1.3	GTR+I+G
nd5-12p		594	0	29.1	462	74	1.8	HKY+I+G
nd5-3p		297	0	10.9	29	225	1.2	GTR+G
cox1-a	62	575	0	33.9	355	201	1.3	HKY+I+G
cox1-a-12p		384	0	43.8	343	37	n/a	SYM+I
cox1-a-3p		191	0	13.9	12	164	1.1	HKY+G
cox1-b	59	758	0	30.4	474	252	1.4	HKY+I+G
cox1-b-12p		506	0	39.7	456	41	6.9	GTR+I
cox1-b-3p		252	0	11.7	18	211	1.2	HKY+G
cob	59	667	0	29.5	384	256	1.6	HKY+I+G
cob-12p		445	0	38.3	372	57	6.9	HKY+I+G
cob-3p		222	0	12.3	12	199	1.2	GTR+I+G
Ribosomal								
rrnl	62	734 (726–732)	1.6	23.8	592	110	1.2	GTR+I+G
nuDNA (*NUC*)	60	4900	22.8	50.3	3172	937	1.4	GTR+G
Protein coding								
HUWE1	53	698 (616–633)	20.1	44.3	436	162	1.7	HKY+G
TP	32	615	0	51.7	413	161	2	GTR+I+G
TP-12p		410	0	41.5	369	24	1.6	HKY+I+G
TP-3p		205	0	72.4	44	137	2	GTR+I+G
PEPCK	26	630	0	48.9	442	124	1.4	SYM+I+G
PEPCK-12p		420	0	48.4	383	22	1.2	GTR+I+G
PEPCK-3p		210	0	50	102	49	1.4	SYM+G
WG	26	440	0	49.9	310	84	2.2	GTR+I+G
WG-12p		293	0	48.5	274	10	2	HKY+G
WG-3p		147	0	52.4	36	74	2.4	HKY+G
Ribosomal								
LSU-a	60	976 (938–951)	5.6	57.6	821	92	3.6	GTR+G
LSU-b	59	1092 (805–938)	35.3	50	633	258	1.1	GTR+G
ITS2	59	927 (481–724)	57.7	43.7	286	258	1.3	GTR+G
SSU	25	592	0	49.7	583	6	1	K80+G
ALL-A	60	8525	13.2	38.9	5509	2028	1.4	GTR+I+G
ALL-B	71	9595	11.7	39.1	6220	2263	1.4	GTR+I+G

Const. sites: Constant sites; Inf. sites: Parsimony informative sites; T_s/T_v: Transition/Transversion ratio.

cate. Additionally, any incongruence of combined mitochondrial (*MIT*) and nuclear (*NUC*) data sets with respect to their 50% majority rule consensus trees, as obtained from MrBayes analyses, were reciprocally assessed by SH testing (Shimodaira & Hasegawa, 1999) in PAUP* based on 1000 RELL bootstrap pseudoreplicates.

CALIBRATION ANALYSES

Calibration analyses were conducted in BEAST 1.6.1 (Drummond & Rambaut, 2007) for *MIT*, *NUC*, *ALL-A* and *ALL-B* data sets, excluding out-groups and partitioning by gene and codon positions for protein-

coding genes, as explained above. Two independent runs of 50 million generations sampled every 2000th generation were performed for each analysis, using a Yule tree prior and the evolutionary model best fitting each of the partitions, considering ten gamma categories when this rate-variation parameter was included in the selected model. Analyses were performed twice, under strict-clock (SC) and relaxed-clock assumptions, with the latter analysis using an uncorrelated lognormal (ULN) model to fit across-branch rate variation, in order to select for the most appropriate clock model explaining the data. Clock model selection was based on Bayes factor (BF) comparisons (Kass & Raftery, 1995). In our implementation, BFs

were interpreted as requiring at least a ten-unit increase in marginal likelihood per additional free parameter before accepting a more complex model (Pagel & Meade, 2004; Miller, Bergsten & Whiting, 2009). We assumed one more parameter in ULN analyses compared with the SC assumption (Drummond *et al*., 2006). In these analyses, every nucleotide substitution was modelled with a prior uniform probability function ranging from 0 to 10, the rate of molecular evolution between 0 and 1, and the Yule prior parameter (yule.birthRate) between 0 and 20. These constraints were selected empirically against default priors for their enhancing stability and convergence of different runs. Other priors and settings were used as default options. Trace plots and ESSs of likelihoods were visualized using TRACER 1.5 to confirm that the stationary phase was reached and to assess the convergence of independent runs. Samples from two independent runs were pooled using LOG-COMBINER 1.6.1 (Drummond & Rambaut, 2007) after removing the initial 10% of results as a burn-in. Consensus trees were estimated in TREEANNOTA-TOR 1.6.1 (Drummond & Rambaut, 2007).

Two nodes were employed for tree time calibration, and were defined as gamma-constrained ages used as prior age information: (1) node G1 (gamma prior age: shape = 63.787, scale = 0.118, and offset = 0), representing the Miocene split between *Carabus (Macrothorax) rugosus* Fabricius, 1792 and *Carabus (Macrothorax) morbillosus* Fabricius, 1792; and (2) node G2 (gamma prior age: shape = 66.361, scale = 0.144, and offset = 0), representing the Miocene split between the subgenera *Eurycarabus* Géhin, 1876 and *Nesaeocarabus* Bedel, 1895. These nodes, both present in the phylogenies and affecting lineages outside, but closely related to *Mesocarabus*, were selected to be old enough to avoid time-dependence effects (Ho *et al*., 2011), but not so deep as to be excessively affected by the saturation of molecular change. The ages for these two calibration nodes (node G1, mean 7.5 Mya, 95% HPD 6.0–9.1 Mya; node G2, mean 9.5 Mya, 95% HPD 5.5–11.6 Mya) were obtained from dating analysis based on *nd5* data calibrated using eight calibration points across the phylogeny of *Carabus* (Andújar *et al*., 2012). We used TreeStat 1.6.1 (Rambaut & Drummond, 2010) to retrieve ages for the nodes of interest from the corresponding MCMC sample in BEAST, and used the 'fitdistr' option of the R package MASS (Venables & Ripley, 2002) to obtain a gamma function adjusting the distribution of sampled ages.

DISPERSAL–VICARIANCE ANALYSES

Dispersal–vicariance analyses were conducted in DIVA (Ronquist, 1997a, 2001) using the statistical graphic interface of S-DIVA 1.5 (Yu, Harris & He, 2010), based on 10 000 random trees from the Bayesian posterior probability tree distribution of BEAST analyses for the *ALL-B* data set (71 taxa), after discarding the initial 10% of trees. The strict consensus tree from this sample was used to visualize the results. For simplicity, and given the intrinsic limitations of the software to deal with complex biogeographic scenarios, we only considered three areas, and ancestral area reconstruction was consequently limited to 'maxareas = 3': A, Iberian Peninsula; B, Eurasia; C, North Africa and the Canary Islands. This event-based parsimony method minimizes dispersal and extinction events, favouring vicariance by the use of a cost matrix to estimating the most parsimonious ancestral ranges in a phylogeny (Ronquist, 1997b). This method is considered as acceptable to reconstruct reticulate biogeographical scenarios, as there is no hierarchical pattern of constraints for area relationships (Sanmartín, 2003). The hypothetical ancestral area distributions and the age estimates of the calibrated molecular phylogeny were interpreted and compared with data on the geological history of the western Mediterranean region.

RESULTS

PHYLOGENETIC INFERENCE

Sequences for mitochondrial genes showed no length variation in the case of protein-coding genes, and variation was low in the case of *rrnL*, with only 1.6% of gapped position in the aligned matrix. Overall, mitochondrial genes were characterized by moderate G + C composition, ranging from 29.0 to 33.9%, and transition/transversion ratios ranging from 1.2 to 1.6. The optimal substitution models for these genes always included invariants and gamma parameter modelling rate variation. Nuclear genes had roughly similar values for transition/transversion ratios (1–2.2), with the exception of *LSU-a* (3.3), but had higher G + C compositions (43.7–57.6%) than mitochondrial genes. Nuclear protein-coding genes showed no length variation and required both invariants and gamma parameters in their optimal model of evolution. *HUWE1* showed length variation and had a stop codon close to its 3′ end, and was therefore processed as a non-coding gene; the substitution model for this marker only required the gamma parameter. The alignment of most variable nuclear ribosomal genes, *LSU-b* and *ITS2*, required a relatively high proportion of gapped positions, 35.3 and 57.7%, respectively. The optimal substitution model and additional information about the loci studied, and their final alignments, are provided in Table 2.

The node height of the MRCA of *Carabus*, as obtained in ML analyses, was within the 95% HPD branch length interval of BEAST analyses for both

Table 3. Tree length from tips to the node (expressed in nucleotide substitutions) for the most recent common ancestor of *Carabus*

Data set	ML(RAxML)	BA (MrBayes)*	BA (BEAST)*
MIT	0.447	0.410 (0.331–0.731)	0.475 (0.407–0.547)
NUC	0.3110	1.047 (0.818–1.292)	0.216 (0.177–0.261)
ALL-A	0.417	0.695 (0.554–0.823)	0.552 (0.471–0.652)
ALL-B	0.432	0.641 (0.485–0.830)	0.448 (0.394–0.557)

*Median and 95% highest posterior density intervals values of the Bayesian posterior probability trees.

MIT and *ALL-B* data sets, and was only slightly lower for the *ALL-A* data set and slightly higher for the *NUC* data set. Conversely, the 95% HPD length interval from the results of MrBayes only included the height of the optimal ML tree in the case of the *MIT* data set, and the *NUC* data set resulted in unrealistic long branch lengths, but even in the latter case, the tree topologies were very similar to those obtained under ML and BEAST analyses (Table 3).

Overall, Bayesian trees obtained for individual gene fragments showed low support for several nodes in the basal part of the tree, despite all of them supporting the monophyly of the genus *Carabus* with respect to its sister *Calosoma* (posterior probability, PP ≥ 0.95; Figs S1–S13). Both *MIT* and *NUC* combined data sets produced phylogenies showing high support for most nodes, independently of the phylogenetic method employed (PP ≥ 0.99; bootstrap support, BS ≥ 95%). Nevertheless, some nodes representing initial splits on the evolution of the genus *Carabus* appeared with low support (Fig. 3). SH and PHT tests revealed significant incongruence between *MIT* and *NUC* data sets ($P = 0.000$ and $P = 0.001$, respectively), despite PHT not detecting any conflict when the *ALL-B* data set was partitioned by gene and analysed ($P = 0.467$). MIT and NUC combinations partitioned by gene also produced no significant incongruence ($P = 1.00$ in both cases). Pairwise comparisons of individual loci revealed inconsistencies between several pair combinations of mitochondrial and nuclear genes, and overall consistency between genes of the same genome, with some exceptions for nuclear ribosomal genes (Table S3). Consequently, we studied *MIT* and *NUC* combined data sets separately, but we also investigated the combination of all genes to better resolve the consistent parts of the *Carabus* phylogeny. Particular inconsistencies were taken into account for the interpretation of the evolutionary history of the group of interest. For instance, some apparent inconsistencies between mitochondrial and nuclear DNA-based trees affected highly supported internal nodes, including: the position of *Orinocarabus*, sister to the (*Nesaeocarabus* + *Eurycarabus*) clade in the nuDNA tree, and to (*Oreocarabus* + *Mesocarabus*) in the mtDNA tree; the

position of *Oreocarabus amplipennis* Vacher de Lapouge, 1924 within *Oreocarabus*; or the position of *C. problematicus* within *Mesocarabus* (nodes labelled with a star in Fig. 3). Particular individuals within the *Mesocarabus* clade also showed alternative positions depending on the source of phylogenetic data (Fig. 3).

Both strategies of global data combination, *ALL-A* and *ALL-B*, produced similar topologies regardless of the phylogenetic method employed, and with high support for most nodes, including basal splits in the evolution of *Carabus*, which were highly supported, at least in Bayesian analyses (PP ≥ 0.95; Fig. 4). *Mesocarabus* was included within Metacarabi, appearing as a monophyletic and strongly supported group (node B), sister to Iberian *Oreocarabus*, with high support (node A). The latter subgenus (*Oreocarabus sensu* Casale & Kryzhanovskij, 2003) was recovered as polyphyletic, as *Carabus hortensis* Linnaeus, 1758 and *Carabus glabratus* Paykull, 1790 did not appear closely related to Iberian *Oreocarabus*. Six of the 13 individual gene fragments also recovered Iberian *Oreocarabus* species as the sister taxon to *Mesocarabus* (node A, PP ≥ 0.95), whereas the monophyly of *Mesocarabus* (node B) was also supported by six individual data sets (Table 4). All analyses on combined data sets (*MIT*, *NUC*, *ALL-A*, and *ALL-B*) recovered, with high support, the sister relationships of *Mesocarabus* and Iberian *Oreocarabus* (node A), the monophyly of *Mesocarabus* and the sister relationship of European *Mesocarabus* with *Carabus* (*Mesocarabus*) *riffensis* (node B), the monophyly of European *Mesocarabus* (node C), and the monophyly of Iberian *Oreocarabus* (Node D: Figs 2–4; Table 4). *Orinocarabus* was found to be the sister group to *Mesocarabus* + Iberian *Oreocarabus*, but only for the *MIT* and the *ALL* data sets (Figs 2 and 3); it was clustered with the remaining European and North African Metacarabi for the *NUC* data set (Fig. 3). The Metacarabi (*sensu* Deuve, 2004) did not constitute a monophyletic group, as they included some taxa traditionally considered as Digitulati (subgenera *Nesaeocarabus* and *Eurycarabus*), and *Carabus* (*Cavazzutiocarabus*) *latreilleanus* Csiki, 1927 appeared in another clade (Fig. 4). We also
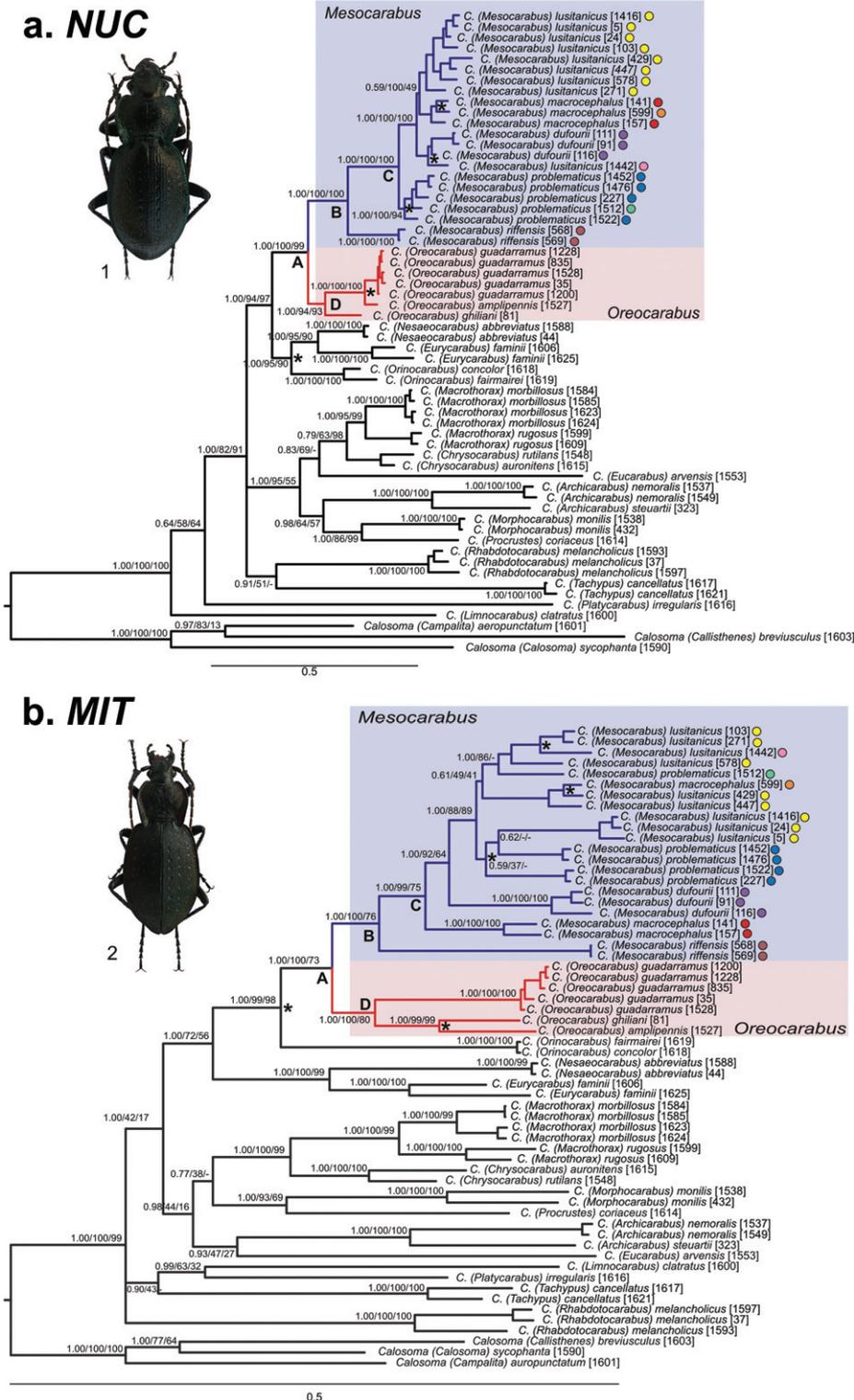
Figure 3. Bayesian 50% majority rule consensus trees from (a) nuclear (*NUC*) and (b) mitochondrial (*MIT*) data sets. The numbers beside nodes represent posterior probabilities and bootstrap values for maximum-likelihood and maximum-parsimony analyses, respectively. Labels A–D indicate cladogenetic events referred to in the text for *Mesocarabus* (in blue) and Iberian *Oreocarabus* (in red). Asterisks indicate incongruent nodes between *MIT* and *NUC* data sets. The species colour codes are as described in Figure 1. Voucher numbers are indicated in brackets. Specimens illustrated: 1, *Carabus (Mesocarabus) lusitanicus* from Salamanca, Spain; 2, *Carabus (Oreocarabus) ghiliani* from Segovia, Spain.
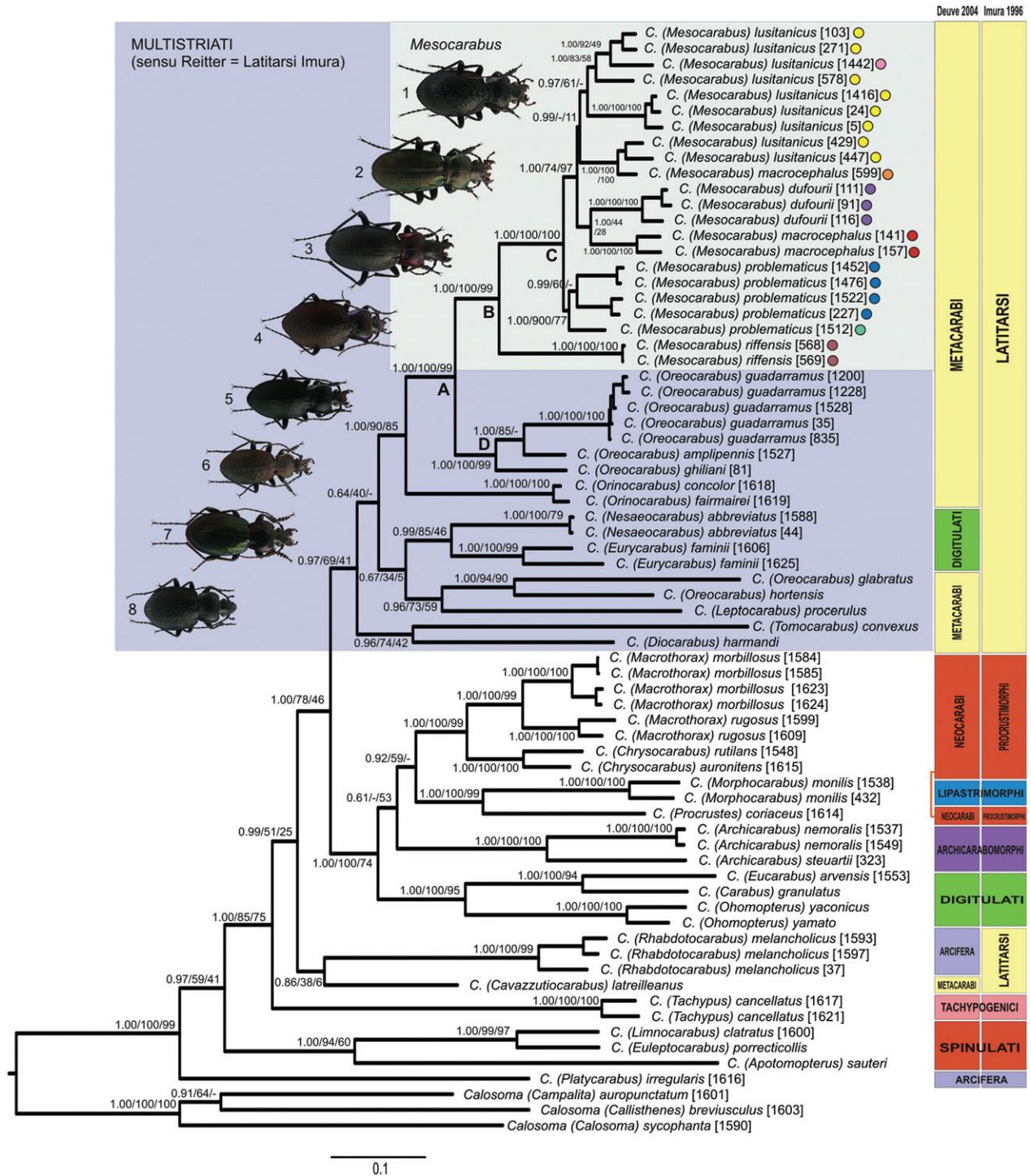
Figure 4. Bayesian 50% majority rule consensus tree for the total evidence data set (*ALL-B*). Numbers besides nodes represent posterior probabilities and bootstrap values for maximum-likelihood and maximum-parsimony analyses, respectively. Labels A–D indicate the cladogenetic events for *Mesocarabus* and Iberian *Oreocarabus* referred to in the text. The species colour codes are as described in Figure 1. Voucher numbers are indicated in brackets. Vertical bars represent the main lineages, as proposed by Imura (1996) and Deuve (2004). Specimens illustrated: 1, *Carabus (Mesocarabus) lusitanicus* from Tarragona, Spain; 2, *Carabus (Mesocarabus) macrocephalus* from León, Spain; 3, *Carabus (Mesocarabus) riffensis* from El Biutz, Morocco; 4, *Carabus (Oreocarabus) guadarramus* from Madrid, Spain; 5, *Carabus (Oreocarabus) amplipennis* from León, Spain; 6, *Carabus (Orinocarabus) concolor* from Bex, Switzerland; 7, *Carabus (Nesaeocarabus) abbreviatus* from Tenerife, Spain; 8, *Carabus (Eurycarabus) faminii* from Rif Massif, Morocco.

Table 4. Node support [Bayesian posterior probability (PP) for individual data sets; PP, maximum likelihood and parsimony bootstrap for combined data] for relevant splits in the evolution of *Mesocarabus* and Iberian *Oreocarabus*. Nodes as shown in Figure 3

Data set	NODE A	NODE B	NODE C	NODE D
nd5	1	0.99	1	0.91
cox1-a	0.68	0.78	–	0.97
cox1-b	–	1	1	0.56
cob	1	1	0.56	1
rrnl	–	–	0.59	0.58
LSU-a	–	–	–	–
LSU-b	1	1	1	0.67
ITS2	0.91	1	1	0.73
SSU	–	–	–	–
TP	1	–	1	1
HUWE1	0.99	0.97	0.98	0.96
PEPCK	1	x	–	x
WG	–	–	–	–
MIT	1.00/100/73	1.00/100/76	1.00/99/75	1.00/100/80
NUC	1.00/100/99	1.00/100/100	1.00/100/100	1.00/100/100
ALL-A	1.00/100/99	1.00/100/99	1.00/100/100	1.00/100/99
ALL-B	1.00/100/99	1.00/100/99	1.00/100/100	1.00/100/99

found an unexpected relationship between *Carabus (Morphocarabus) monilis* Fabricius, 1792 and *Carabus (Procrustes) coriaceus* Linnaeus, 1758 in both mitochondrial and nuclear phylogenies that deserves further study.

CALIBRATION AND BIOGEOGRAPHIC ANALYSES

BEAST phylogenetic calibration analyses on the *ALL-B* data set resulted in similar topologies and posterior probabilities than BI analyses conducted in MrBayes (Fig. 5). These analyses dated the time for the MRCA of *Carabus* at 27.4 Mya (95% HPD interval 19.6–36.4 Mya), at the end of the Oligocene epoch (Fig. 5). All cladogenetic events leading to the main extant lineages and subdivisions proposed by Deuve (2004) occurred between 25 and 15 Mya according to our estimations. The split between *Mesocarabus* and Iberian *Oreocarabus* (node A) was dated at 11.8 Mya (15.3–8.7 Mya), and that between Moroccan *C. riffensis* and European *Mesocarabus* (node B) was dated at 9.5 Mya (12.5–7.0 Mya), during the Miocene. The diversification of European *Mesocarabus* into several lineages within the Iberian Peninsula and continental Europe was dated between 6.4 and 4.5 Mya (8.5–3.2), during the Messinian and early Pliocene epoch (Fig. 5; Table 5).

Ancestral area reconstructions inferred the origin of the MRCA of the Metacarabi lineage out of the Iberian Peninsula during the Miocene, at around 13.6–23.2 Mya (Fig. 6). The ancestor of *Mesocarabus* was inferred inhabiting areas AC (55%), BC (18%), or ABC (27%), whereas the ancestor of both *Mesocarabus* and Iberian *Oreocarabus* showed a higher probability to have occurred in the Iberian Peninsula (A = 64%), or in this area combined with others (AB = 18%; ABC = 18%; Fig. 6).

DISCUSSION

THE EVOLUTIONARY HISTORY OF *MESOCARABUS*

Mesocarabus has been retrieved as part of a paraphyletic Metacarabi division, *sensu* Deuve (2004). Most of our findings relating to the high-level taxonomy of *Carabus* are consistent with those found in Deuve *et al*. (2012). It seems that the Metacarabi, defined on the basis of morphological characters (e.g. the endophallus of the median lobe of male genitalia), are not fully congruent with those derived from molecular data. Thus, *Carabus (Rhabdotocarabus) melancholicus* Fabricius, 1798, *Carabus (Tachypus) cancellatus* Illiger, 1798, and *Carabus (Cavazzutiocarabus) latreilleanus* should not be included within the Metacarabi (Fig. 4) (Sota & Ishikawa, 2004). *Carabus latreilleanus* was considered as part of Metacarabi based on characteristics of the endophallus and *nd5* phylogenies (Imura, 2002; Deuve, 2004), but we find it here as an early split (albeit with low support) to the subgenus *Rhabdotocarabus* Seidlitz, 1887. In turn, *Nesaeocarabus* and *Eurycarabus*, currently included within the Digitulati (Deuve, 2004), and not sampled by Sota & Ishikawa (2004), should be considered part of the Metacarabi, as previously suggested by Arndt *et al*. (2003) based on their larval
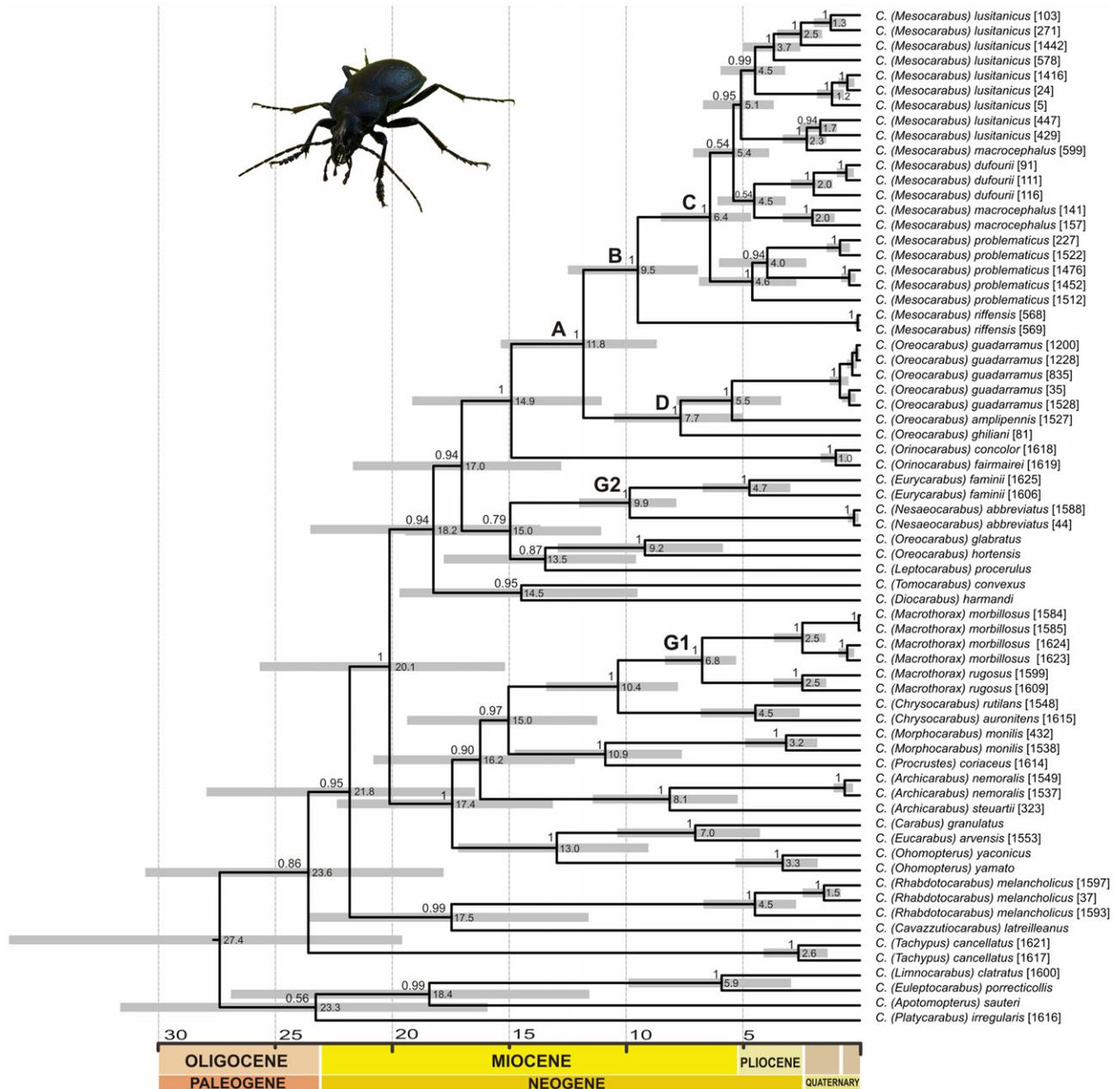
Figure 5. Ultrametric time-calibrated tree for combined DNA markers (*ALL-B* data set) in *Carabus*. Numbers above nodes represent posterior probabilities. Grey bars on nodes represent the 95% confidence intervals for node ages (Myr), with mean ages indicated inside the bars. Labels A–D indicate the cladogenetic events for *Mesocarabus* and Iberian *Oreocarabus* referred to in the main text; labels G1 and G2 indicate nodes used as calibration priors. Specimen illustrated: *Carabus (Mesocarabus) lusitanicus* from Albacete, Spain.

morphology. Our results also show that *Oreocarabus sensu* Casale & Kryzhanovskij (2003) is a polyphyletic taxon, including at least two lineages (one Iberian endemic and the other trans-Pyrenean). The validity of *Oreocarabus* is thus questioned, in agreement with the results of Deuve *et al.* (2012), and the systematic arrangement of the species involved must be settled in accord with the new molecular evidence.

Mesocarabus is unambiguously retrieved as sister taxon to Iberian *Oreocarabus* and more distantly related to *Orinocarabus*. The integration of phylogenetic calibration and ancestral area reconstruction analyses suggests that the evolutionary history for western European *Mesocarabus*, *Oreocarabus*, and *Orinocarabus* subgenera is linked to the complex geological history and climatic changes that occurred in

Table 5. Ages (Myr) and 95% highest posterior density interval obtained in BEAST with the different combined data sets for relevant nodes in the evolution of *Mesocarabus* and Iberian *Oreocarabus*. Nodes as shown in Figure 3

Data set	ROOT	NODE A	NODE B	NODE C	NODE D
MIT	21.4 (16.3–26.6)	11.78 (9.1–14.7)	10.0 (7.7–12.6)	7.9 (6.1–10.0)	8.4 (6.7–11.5)
NUC	26.6 (16.0–41.4)	13.5 (8.4–20.5)	10.1 (5.8–15.3)	6.4 (3.9–10.1)	8.3 (3.7–14.2)
ALL-A	25.5 (18.1–35.0)	12.3 (9.0–16.2)	9.7 (6.9–12.9)	6.2 (4.5–8.1)	7.7 (5.1–10.7)
ALL-B	27.4 (19.6–36.4)	11.8 (8.7–15.3)	9.5 (7.0–12.5)	6.4 (4.7–8.5)	7.7 (5.0–10.5)

the Western Palaearctic region during the Cenozoic Era (Andeweg, 2002; Krijgsman, 2002; Rosenbaum, Lister & Duboz, 2002; Meulenkamp & Sissingh, 2003) (Fig. 6). The colonization of the Iberian Peninsula by the Eurasian Metacarabi ancestor is dated between 17 and 14.9 Mya (95% HPD 21.6–11.1 Mya; Fig. 6), and was probably accompanied by the split into an Iberian clade (originating the ancestor of both *Mesocarabus* and Iberian *Oreocarabus*) and one European clade (ancestor of *Orinocarabus*). This hypothesis is congruent with the relationships inferred from the data set combining all genes (Fig. 4) and on mitochondrial data alone (Fig. 3b). However, the nuclear phylogeny is compatible with *Orinocarabus* sharing a common ancestor with other representatives of the Metacarabi lineage. Therefore, these incongruent phylogenetic scenarios suggest that the split of the Iberian clade was the result of an early colonization of the Iberian Peninsula, about 17 Mya (95% HPD 23.5–13.7 Mya), probably followed by an episode of mitochondrial capture from the Iberian taxa into the *Orinocarabus* lineage, which could be dated around 14.9 Myr (95% HPD 19.1–11.1 Mya).

Dates obtained for the splits between *Mesocarabus* and Iberian *Oreocarabus* (15.3–8.7 Mya), and between Moroccan *C. riffensis* and European *Mesocarabus* (12.5–7.0 Mya), are suggestive of an early colonization of the Betic-Riffian plate by ancestral *Mesocarabus* in the Upper Miocene, when land bridges started to be available between these land masses (Fig. 6). Present evidence does not allow rejecting whether the split between *Oreocarabus* and *Mesocarabus* occurred in the Iberian Peninsula before the colonization of the Betic-Riffian plate, or was a result of a vicariant event. The first hypothesis requires only one dispersal event of *Mesocarabus* from Iberia to the Betic-Riffian plate at around 9.5 Mya (95% HPD 12.5–7.0 Mya), whereas the second implies an earlier colonization of the Betic-Riffian plate about 11.8 Mya (95% HPD 15.3–8.7 Mya) by the ancestor of *Mesocarabus* and Iberian *Oreocarabus* (where *Mesocarabus* evolved), and a return of *Mesocarabus* to colonize the Iberian Peninsula, where they diversified. Both alternative scenarios are geologically possible thanks to connections between the

Iberian and the Betic-Riffian plates during the late Miocene (Martin, Braga & Betzler, 2001; Andeweg, 2002; García-Castellanos et al., 2009), where different episodes of dispersal and vicariance could have occurred at different times.

The diversification of European *Mesocarabus* into several lineages within the Iberian Peninsula and continental Europe is dated between 6.4 and 4.5 Mya (95% HPD 8.5–3.2 Mya), during the Messinian and early Pliocene epoch. The reconstruction of the evolutionary history of this lineage is hindered by the complex geological changes of the Iberian Peninsula, and the occurrence of major barriers, such as large continental basins and transversal mountain chains. Moreover, secondary contact and hybridization between entities that account for different degrees of differentiation probably obscure the evolutionary history of the group. This reconstruction will be explored in depth in a future work. Yet, some interesting insight about the geographic drivers for the evolution of this group can be derived from our data. The current distribution of most *Mesocarabus* species in the Iberian Peninsula and their sister *C. problematicus* in most of non-peninsular Europe may be suggestive of the classical pattern with dispersal and (phylo)genetic differentiation from southern refugia accompanying Quaternary climatic changes (Hewitt, 1999, 2004). However, the split between the European *C. problematicus* and the Iberian *Mesocarabus* occurred at 6.4 Mya (95% HPD 8.5–4.7 Mya), which strongly suggests that the origin of *C. problematicus* was not associated with the climatic oscillations of the Pleistocene, but rather pre-dated them. Instead, the latter species could have had an allopatric origin in western Europe, with the Pyrenees acting as a major isolation barrier.

INCONGRUENCE BETWEEN MITOCHONDRIAL AND NUCLEAR PHYLOGENIES

Incongruence between mtDNA and nuclear DNA (nuDNA) phylogenies are common (e.g. Shaw, 2002; Gómez-Zurita & Vogler, 2003; Leaché & McGuire, 2006; Ting et al., 2008; Spinks & Shaffer, 2009), and this has also been described in different studies of
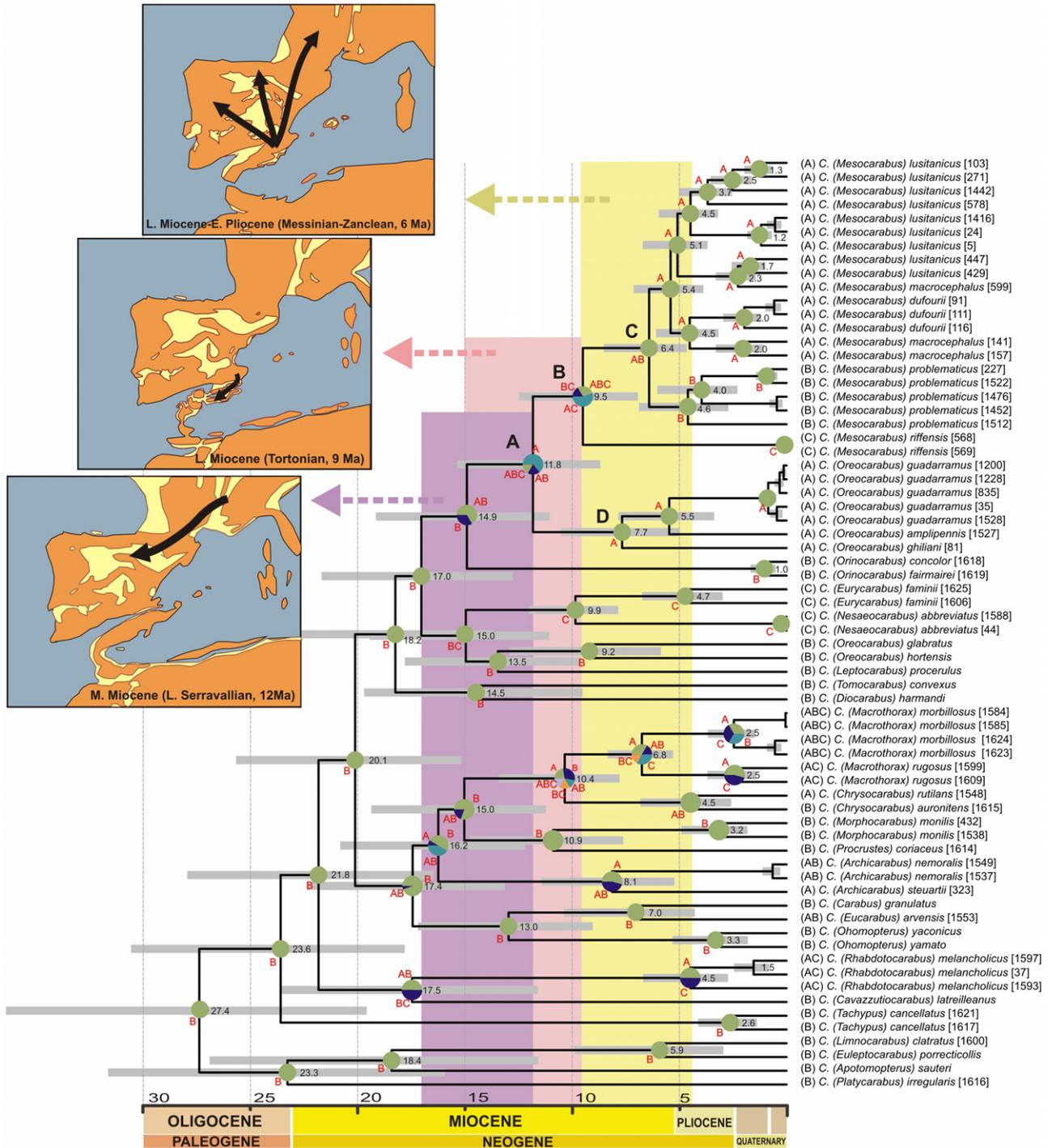
Figure 6. Ultrametric time-calibrated tree for combined DNA markers (*ALL-B* data set) in *Carabus* showing ancestral area inferences (A, Iberian Peninsula; B, Eurasia; C, North Africa and Canary Islands). Pie charts represent the probability for each area reconstruction. The grey bars on nodes represent the 95% confidence intervals for node ages (Myr), with mean ages indicated inside bars. The palaeogeographic reconstructions are taken from Andeweg (2002).

Carabus, e.g. with the subgenus *Ohomopterus* Reitter, 1896 in Japan (Sota & Vogler, 2001, 2003; Sota *et al.*, 2001; Nagata, Kubota & Sota, 2007; Nagata *et al.*, 2007) or with *Chrysocarabus* Thomson, 1875 in Europe

(Prüser, 1996; Streiff *et al.*, 2005; Düring, Bruckner & Mossakowski, 2006). In our phylogenetic study centred in *Mesocarabus*, we have observed inconsistent results between mtDNA and nuDNA phylogenies spanning

several taxonomic levels within *Carabus* (Fig. 3). There is incongruence in the relationship of the subgenus *Orinocarabus* within Metacarabi at the subgeneric level, but also for the specific relationships of *C. (Oreocarabus) amplipennis* and *C. (Mesocarabus) problematicus* within *Oreocarabus* and *Mesocarabus*, respectively. But incongruence also affects particular *Mesocarabus* specimens, such as *Carabus (Mesocarabus) macrocephalus* (voucher ref. 599) and *Carabus (Mesocarabus) lusitanicus* (voucher ref. 1442) (see Fig. 3). A number of evolutionary processes may cause incongruence between independent molecular markers, including low mutation rate, natural selection, ancestral polymorphism, and introgressive hybridization (Funk & Omland, 2003; Ballard & Whitlock, 2004), and it is difficult to distinguish among them.

However, some well-known characteristics of this genus point to hybridization, leading to introgression as the most plausible explanation for the patterns observed. Introgression has been proven for three species of the subgenus *Chrysocarabus* inhabiting the southern slopes of the Pyrenees (Düring *et al*., 2006), and natural hybridization between *Carabus (Chrysocarabus) lineatus* Dejean, 1826 and *Carabus (Chrysocarabus) splendens* Olivier, 1790 at both sides of the western Pyrenees was also documented on the basis of allozymes and morphological traits (Mossakowski, Roschen & Vaje, 1986). Furthermore, there is a large background of great success in artificial interspecific crosses for *Carabus* species (Deuve, 2004), referred to species within *Mesocarabus* (Puisségur, 1987), *Chrysocarabus* (Puisségur, 1987; Godeau, Malausa & Drescher, 1991), and *Macrothorax* (Godeau *et al*., 1991; Malausa *et al*., 1991). Crosses between species of different subgenera have also been successfully obtained (Imura, 1989; Deuve, 1994, 2004). In support for the idea of introgression affecting *Mesocarabus* and related taxa, the observed cases of incongruence between nuDNA and mtDNA can always be mapped to situations where the taxa involved in the hybridization process meet (Fig. 1), thereby satisfying the prerequisite of spatial coexistence that is required for hybridization (Gómez-Zurita & Vogler, 2003).

Carabus lusitanicus baguenai Breuning, 1926 has been traditionally considered a subspecies of *C. lusitanicus* based on external morphology (Serrano, 2003), an assignment supported by mitochondrial data. However, using nuclear genes our results based on a specimen from Sierra Alhamilla (voucher number 1442) show that it is clearly related to parapatric *C. dufourii*, as already suggested by the characteristics of the everted endophallus (Anichtchenko, 2004). These results suggest that historic hybridization between *C. dufourii* and *C. lusitanicus* left a signature of morphological and molecular character admixture in geographically intermediate populations between both species. These molecular data may be interpreted as supporting the proposal of Anichtchenko (2004) of a *C. dufourii baguenai* subspecies. If this combination of characters is fixed across the distribution range of this taxon, and considering its relatively old age for both mitochondrial (1.59–3.01 Mya) and nuclear (0.57–3.30 Mya) time-calibrated phylogenies, it would be possible to consider *C. baguenai* as a valid species of hybrid origin.

The case of the specimen of *Carabus macrocephalus barcelecoanus* Lapouge, 1925 from Puerto de Lunada (voucher number 599) is somehow different, as both nuclear phylogeny and aedeagal and external morphology agree with those of typical *C. macrocephalus*, whereas only mitochondrial DNA corresponds to *C. lusitanicus*. Thus, *C. macrocephalus barcelecoanus* should be retained within *C. macrocephalus*, and possibly affected by a past episode of mtDNA capture. Similarly, all available evidence indicates that the specimen of *C. problematicus* from Ochagavía (voucher number 1512) has an introgressed mitochondrial DNA from *C. lusitanicus*. In summary, hybridization followed by introgression between *Mesocarabus* lineages are not uncommon events, and they seem to have given rise to intermediate populations, the status of which as independent lineages poses new evolutionary problems that we are currently investigating.

REFERENCES

Andeweg B. 2002. *Cenozoic tectonic evolution of the Iberian Peninsula: causes and effects of changing stress fields*. Amsterdam: Netherland Research School of Sedimentary Geology (NSG).

Andújar C, Serrano J, Gómez-Zurita J. 2012. Winding up the molecular clock in the genus Carabus (Coleoptera: Carabidae): assessment of methodological decisions on rate and node age estimation. *BMC Evolutionary Biology* **12:** 40.

Anichtchenko AV. 2004. Notas taxonómicas sobre el subgénero Mesocarabus Thomson, 1875 (Coleoptera, Carabidae) de la Península Ibérica. Primera nota. *Boletín De La Sociedad Española De Entomología* **28:** 89–103.

Arndt E, Brücker M, Marciniak K, Mossakowski D, Prüser F. 2003. Phylogeny. In: Turin H, Penev L, Casale A, eds. *The genus Carabus L. In Europe, a synthesis*. Sofia & Leiden: Pensoft & European Invertebrate Survey, 307–325.

Ballard JWO, Whitlock MC. 2004. The incomplete natural history of mitochondria. *Molecular Ecology* **13:** 729–744.

Bengtsson S. 1927. Die Larven der nordischen Arten von Carabus L. Eine morphologische studie. Kungliga Fysiografiska sällskapet i Lund. *Förhardlingar. (N.F.)* **39:** 1–89.

Breuning S. 1932–1937. Monographie der Gattung Carabus L. Troppau.

Brown JM, Hedtke SM, Lemmon AR, Lemmon EM. 2010. When trees grow too long: investigating the causes of highly inaccurate bayesian branch-length estimates. *Systematic Biology* **59:** 145–161.

Casale A, Kryzhanovskij OL. 2003. Key to the adults. In: Turin H, Penev L, Casale A, eds. *The genus Carabus L. In Europe, a synthesis*. Sofia & Leiden: Pensoft & European Invertebrate Survey, 73–123.

Deuve T. 1994. *Une classification du genre Carabus*. Paris: Muséum National d'Histoire Naturelle.

Deuve T. 2004. *Illustrated catalogue of the genus Carabus of the World (Coleoptera: Carabidae)*. Sofia: Pensoft.

Deuve T, Cruaud A, Genson G, Rasplus JY. 2012. Molecular systematics and evolutionary history of the genus Carabus (Col. Carabidae). *Molecular Phylogenetics and Evolution* **65:** 259–275.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *Plos Biology* **4:** 699–710.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 214.

Düring A, Bruckner M, Mossakowski D. 2006. Different behaviour of mitochondrial and nuclear markers: introgression and the evolutionary history of Chrysocarabus (Coleoptera: Carabidae). *Entomologica Fennica* **17:** 200–206.

Farris JS, Kállersjö M, Kluge AG, Bult C. 1995. Testing significance of incongruence. *Cladistics* **10:** 315–319.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39:** 783–791.

Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* **34:** 397–423.

García-Castellanos D, Estrada F, Jiménez-Munt I, Gorini C, Fernández M, Verges J, De Vicente R. 2009. Catastrophic flood of the Mediterranean after the Messinian salinity crisis. *Nature* **462:** 778–781.

Godeau B, Malausa JC, Drescher J. 1991. L'hybridation expérimentale entre le genre Macrothorax Desmarest et l'ensemble des espèces du genre Chrysocarabus Thomson. *Bulletin De La Société Entomologique De France* **96:** 144.

Goloboff P, Farris JS, Nixon K. 2003. T.N.T.-Tree analysis using new technology. Vers. 1.1. (December 2007). Available at: http://www.zmuc.dk/public/phylogeny

Gómez-Zurita J, Vogler AP. 2003. Incongruent nuclear and mitochondrial phylogeographic patterns in the Timarcha goettingensis species complex (Coleoptera, Chrysomelidae). *Journal of Evolutionary Biology* **16:** 833–843.

Hewitt GM. 1999. Post–glacial re–colonization of European biota. *Biological Journal of the Linnean Society* **68:** 87–112.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B* **359:** 183–195.

Ho SYW, Lanfear R, Bromham L, Phillips MJ, Soubrier J, Rodrigo AG, Cooper A. 2011. Time-dependent rates of molecular evolution. *Molecular Ecology* **20:** 3087–3101.

Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17:** 754–755.

Imura Y. 1989. Natural hybrids of the Damaster species (Coleoptera, Carabidae) in Hokkaido, Northern Japan. *Japanese Journal of Entomology* **57:** 67–71.

Imura Y. 1996. A revised classification of the major divisions and subdivisions of Carabus (s. lat.) (Coleoptera, Carabidae). *Elytra, Tokyo* **24:** 5–12.

Imura Y. 2002. Classification of the subtribe Carabina (Coleoptera, Carabidae) based on molecular phylogeny. *Elytra, Tokyo* **30:** 1–28.

Ishikawa R. 1978. A revision of the higher taxa of the subtribe Carabina (Coleoptera, Carabidae). *Bulletin of the National Science Museum Series A: Zoology* **4:** 15–68.

Kass RE, Raftery AE. 1995. Bayes factors. *Journal of the American Statistical Association* **90:** 773–795.

Katoh K, Asimenos G, Toh H. 2009. Multiple alignment of DNA Sequences with MAFFT. In: Posada D, eds. Bioinformatics for DNA sequence analysis. *Methods in Molecular Biology* **537:** 39–64.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30:** 3059–3066.

Katoh K, Toh H. 2008. Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC Bioinformatics* **9:** 212.

Krijgsman W. 2002. The Mediterranean: mare nostrum of Earth sciences. *Earth and Planetary Science Letters* **205:** 1–12.

Leaché AD, McGuire JA. 2006. Phylogenetic relationships of horned lizards (Phrynosoma) based on nuclear and mitochondrial data: evidence for a misleading mitochondrial gene tree. *Molecular Phylogenetics and Evolution* **39:** 628–644.

Letsch HO, Kück P, Stocsits RR, Misof B. 2010. The impact of rRNA secondary structure consideration in alignment and tree reconstruction: simulated data and a case study on the phylogeny of hexapods. *Molecular Biology and Evolution* **27:** 2507–2521.

Malausa JC, Sapaly S, Godeau B, Drescher J. 1991.

Premières hybridations expérimentales dans le genre Macrothorax Desmarest. *Bulletin De La Société Entomologique De France* **96:** 23–30.

Marshall DC. 2010. Cryptic failure of partitioned Bayesian phylogenetic analyses: lost in the land of long trees. *Systematic Biology* **59:** 108–117.

Marshall DC, Simon C, Buckley TR. 2006. Accurate branch length estimation in partitioned bayesian analyses requires accommodation of among-partition rate variation and attention to branch length priors. *Systematic Biology* **55:** 993–1003.

Martin JM, Braga JC, Betzler C. 2001. The Messinian Guadalhorce corridor: the last northern, Atlantic-Mediterranean gateway. *Terra Nova* **13:** 418–424.

Meulenkamp JE, Sissingh W. 2003. Tertiary palaeogeography and tectonostratigraphic evolution of the Northern and Southern Peri-Tethys platforms and the intermediate domains of the African–Eurasian convergent plate boundary zone. *Palaeogeography, Palaeoclimatology, Palaeoecology* **196:** 209–228.

Miller KB, Bergsten J, Whiting MF. 2009. Phylogeny and classification of the tribe Hydaticini (Coleoptera: Dytiscidae): partition choice for Bayesian analysis with multiple nuclear and mitochondrial protein-coding genes. *Zoologica Scripta* **38:** 591–615.

Mossakowski D, Roschen A, Vaje S. 1986. Hybridization in Chrysocarabus. In: Den Boer P, Luff ML, Mossakowski D, Weber F, eds. *Carabid beetles, their adaptations and dynamics*. Stuttgart: Gustav Fischer, 281–295.

Nagata N, Kubota K, Sota T. 2007. Phylogeography and introgressive hybridization of the ground beetle Carabus yamato in Japan based on mitochondrial gene sequences. *Zoological Science* **24:** 465–474.

Nagata N, Kubota K, Yahiro K, Sota T. 2007. Mechanical barriers to introgressive hybridization revealed by mitochondrial introgression patterns in Ohomopterus ground beetle assemblages. *Molecular Ecology* **16:** 4822–4836.

Osawa S, Su ZH, Imura Y. 2004. *Molecular phylogeny and evolution of carabid ground beetles*. Tokyo: Springer-Verlag.

Pagel M, Meade A. 2004. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Systematic Biology* **53:** 571–581.

Posada D. 2008. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* **25:** 1253–1256.

Posada D, Buckley TR. 2004. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* **53:** 793–808.

Prüser F. 1996. Variabilität mitochondrialer DNA-Sequenzen und die Phylogenie der Gattung Carabus Linné 1758 (Coleoptera: Carabidae) PhD. Bremen: University of Bremen.

Puisségur C. 1987. Hybridation plurispécifique chez les Chrysocarabus et Hadrocarabus. *Bulletin De La Société Des Sciences Nat* **56:** 1–19.

Rambaut A. 2008. FigTree v.1.1.2. Available at: http://tree.bio.ed.ac.uk/software/figtree/

Rambaut A, Drummond AJ. 2007. Tracer v1.5. Available at: http://beast.bio.ed.ac.uk/Tracer

Rambaut A, Drummond AJ. 2010. TreeStat v1.6.1: tree statistic calculation tool.

Ronquist F. 1997a. Dispersal-Vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology* **46:** 195–203.

Ronquist F. 1997b. Phylogenetic approaches in coevolution and biogeography. *Zoologica Scripta* **26:** 313–322.

Ronquist F. 2001. *DIVA version 1.2. Computer program for macos and Win32*. Uppsala: Evolutionary Biology Centre, Uppsala University.

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

Rosenbaum G, Lister GS, Duboz C. 2002. Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. In: Rosenbaum G, Lister GS, eds. Reconstruction of the evolution of the Alpine-Himalayan Orogen. *Journal of the Virtual Explorer* **8:** 107–130.

Sanmartín I. 2003. Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of Biogeography* **30:** 1883–1897.

Serrano J. 2003. *Catálogo de los Carabidae (Coleoptera) de la Península Ibérica*. Zaragoza: Monografías de la Sociedad Entomológica Aragonesa 9.

Shaw KL. 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. *Proceedings of the National Academy of Sciences* **99:** 16122–16127.

Shimodaira H, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* **16:** 1114.

Sota T, Ishikawa R. 2004. Phylogeny and life-history evolution in Carabus (subtribe Carabina : Coleoptera, Carabidae) based on sequences of two nuclear genes. *Biological Journal of the Linnean Society* **81:** 135–149.

Sota T, Ishikawa R, Ujiie M, Kusumoto F, Vogler AP. 2001. Extensive trans-species mitochondrial polymorphisms in the carabid beetles Carabus subgenus Ohomopterus caused by repeated introgressive hybridization. *Molecular Ecology* **10:** 2833–2847.

Sota T, Vogler AP. 2001. Incongruence of mitochondrial and nuclear gene trees in the carabid beetles Ohomopterus. *Systematic Biology* **50:** 39–59.

Sota T, Vogler AP. 2003. Reconstructing species phylogeny of the carabid beetles Ohomopterus using multiple nuclear DNA sequences: heterogeneous information content and the performance of simultaneous analyses. *Molecular Phylogenetics and Evolution* **26:** 139–154.

Spinks PQ, Shaffer HB. 2009. Conflicting mitochondrial and nuclear phylogenies for the widely disjunct Emys (Testudines: Emydidae) species complex, and what they tell us about biogeography and hybridization. *Systematic Biology* **58:** 1–20.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22:** 2688–2690.

Stamatakis A, Blagojevic F, Nikolopoulos D, Antonopoulos C. 2007. Exploring new search algorithms and hardware for phylogenetics: RAxML meets the IBM cell. *The Journal of VLSI Signal Processing* **48:** 271–286.

Streiff R, Veyrier R, Audiot P, Meusnier S, Brouat C. 2005. Introgression in natural populations of bioindicators: a case study of Carabus splendens and Carabus punctatoauratus. *Molecular Ecology* **14:** 3775–3786.

Su ZH, Imura Y, Osawa S. 2001. Evolutionary discontinuity of the Carabine ground beetles. *Journal of Molecular Evolution* **53:** 517–529.

Su ZH, Imura Y, Zhou HZ, Okamoto M, Osawa S. 2003. Mode of morphological differentiation in the Latitarsi-ground beetles (Coleoptera, Carabidae) of the world inferred from a phylogenetic tree of mitochondrial ND5 gene sequences. *Genes, Genetic Systems* **78:** 53–70.

Swofford DL. 2003. *Paup*. Phylogenetic analysis using parsimony (*and other methods). Version 4*. Sunderland, MA: Sinauer Associates.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetic analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24:** 1596–1599.

Ting N, Tosi AJ, Li Y, Zhang YP, Disotell TR. 2008. Phylogenetic incongruence between nuclear and mitochondrial markers in the Asian colobines and the evolution of the langurs and leaf monkeys. *Molecular Phylogenetics and Evolution* **46:** 466–474.

Turin H, Penev L, Casale A. 2003. *The genus Carabus in Europe. A synthesis*. Sofia & Leiden: Pensoft & European Invertebrate Survey.

Venables WN, Ripley BD. 2002. *Modern applied statistics with S*, Fourth edn. New York: Springer.

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

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Figures S1–13. Phylogenetic trees obtained with MrBayes for each single gene fragment data set: S1, *cox1-a*; S2, *cox1-b*; S3, *cob*; S4, nd5; S5, *rrnL*; S6, *SSU*; S7, *LSU-a*; S8, *LSU-b*; S9, *ITS2*; S10, *HUWE1*; S11, *TP*; S12, *PEPCK*; S13, *WG*.
Table S1. Species, voucher reference, and accession numbers for each specimen and sequence. The collection localities are listed in Table 1.
Table S2. Primers used in the study.
Table S3. Results of a partition homogeneity test for pairwise combinations of genes.