Species of the genus Meladema (Dytiscidae, Colymbetinae) are some of the largest macroinvertebrates in the western Palearctic region, being top predators in fishless streams. Two of the three described species, Meladema imbricata (Wollaston, 1871) and Meladema lanio (Fabricius, 1775) are Macaronesian endemics from the Canary Islands and Madeira, respectively, while the third, Meladema coriacea Laporte, 1835, is widely distributed from Morocco and the Iberian Peninsula to Turkey, including the Canary Islands. Previous phylogenetic analysis using only mitochondrial markers revealed the existence of two cryptic lineages within M. coriacea, one restricted to Corsica and the other including the rest of sampled populations. We reconstruct here the evolutionary history of the species of Meladema using a more comprehensive sampling covering its whole geographical range, adding nuclear markers and Bayesian molecular dating. Using environmental niche modelling, we test for possible differences in climatic preferences among lineages and reconstruct their ancestral climatic niche. Our results strongly supported the existence of four monophyletic lineages represented by the three recognized species plus a fourth cryptic lineage with populations of M. coriacea from the Tyrrhenian islands (Corsica, Sardinia and Montecristo). This pattern is not likely to be the result of mitochondrial artefacts due to Wolbachia infection, as all 11 tested individuals were negative for this parasite. Dating analysis placed the origin of Meladema in the Middle Miocene although diversification among extant Meladema lineages started in the early Pleistocene and took place in a relatively short time period. Phylogeographic analysis inferred a continental origin of Meladema, with an independent colonization of the Macaronesian and Mediterranean islands. From the south-western Mediterranean region, the continental M. coriacea expanded its range up to Bulgaria and Turkey in the northern basin and to Tunisia in the southern. Results of niche modelling showed that seasonality is the critical factor in shaping the current distribution of Meladema. Island lineages (M. imbricata, M. lanio and the Tyrrhenian lineage of M. coriacea) occur in sites with low seasonality, within the range of the reconstructed ancestral climatic niche of the genus. On the contrary, continental M. coriacea expanded its range to localities outside the ancestral climatic range of the genus, with a higher seasonality and aridity.

Corresponding author: Ignacio Ribera, Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Passeig Maritim de la Barceloneta 37, 08003 Barcelona, Spain. E-mail: ignacio.ribera@ibe.upf-csic.es

Vít Sykora, Faculty of Science, Charles University in Prague, Albertov 6, 128 43 Praha 2, Prague, Czech Republic. E-mail: vit.sykora@natur.cuni.cz

David García-Vázquez, Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Passeig Maritim de la Barceloneta 37, 08003 Barcelona, Spain. E-mail: david.garcia@ibe.upf-csic.es

David Sánchez-Fernández, Instituto de Ciencias Ambientales, Universidad de Castilla-La Mancha, Campus Tecnológico de la Fábrica de Armas, Toledo, 45071 Spain. E-mail: david.sfernandez@uclm.es

Ignacio Ribera, Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Passeig Maritim de la Barceloneta 37, 08003 Barcelona, Spain. E-mail: ignacio.ribera@ibe.upf-csic.es
Introduction
The diving beetle genus *Meladema* (Dytiscidae, Colymbetinae) is a common inhabitant of Mediterranean streams, in which it is often one of the largest macroinvertebrates. With a body size ranging from 20 to 23 mm, in small, fishless streams it may also be a top predator, both in the adult and larval stages. The genus currently includes three species: *Meladema coriacea* Laporte, 1835, *Meladema imbricata* (Wollaston, 1871) and *Meladema lanio* (Fabricius, 1775) (Ribera et al. 2003; Alarie & Hughes 2006). *Meladema* is confined to the western Palaearctic region, with one species (*M. coriacea*) distributed from Morocco and the Iberian Peninsula to Turkey and the other two endemic to Macaronesia. The widespread *M. coriacea* is common in the western Mediterranean, including the Iberian Peninsula, Southern France, North Africa and all the western Mediterranean islands (Nilsson & Hájek 2015). It seems to become scarcer towards the east, with few records from the Balkan Peninsula (Zaitsev 1972; Nilsson & Hájek 2015) and a single specimen reported from Turkey in the literature (Gueorguiev 1981; Darilmaz & Kiyak 2009). It is also present in the Canary Islands, likely resulting from a recent colonization (Ribera et al. 2003). There are references of the presence of *M. coriacea* in northern France and in Belgium (Gschwendtner 1936), but even when correct, these were undoubtedly occasional records. In contrast, *M. imbricata* and *M. lanio* are endemics to the Canary Islands and Madeira respectively, in where they are of high conservation concern due to their confinement to well preserved streams (Machado 1987; Balke et al. 1990; Ribera et al. 2003). All three species occupy running waters, mainly deeper pools in streams at intermediate altitudes (Ribera et al. 2003).

In a previous study using only mitochondrial DNA sequences, Ribera et al. (2003) found that *Meladema* was in fact constituted by four genetically distinct lineages with unresolved relationships between them. The two Macaronesian endemics *M. imbricata* and *M. lanio* were found to be respectively monophyletic, but the analyses suggested that *M. coriacea* was not monophyletic, as it was split into a lineage restricted to Corsica and another including the rest of sampled populations (Morocco, the Iberian peninsula and southern France). Despite uncertainties in the calibration, diversification within the genus was estimated to be of early Pliocene origin. None of the lineages seem therefore to be a Tertiary relict, as hypothesized for other Macaronesian endemics (Ribera et al. 2003).

In this study, we present an update of the phylogeny and biogeography of the species of *Meladema*, with a more comprehensive sampling encompassing their whole geographical range, nuclear markers and a Bayesian molecular clock dating. We also include environmental niche modelling to test for differences in climatic preferences among individual lineages, trying to understand the factors determining the current and past distribution of the species of this emblematic genus of freshwater macroinvertebrate.

Material and methods
Sampling, DNA extraction and sequencing
We sampled all three known species of the genus *Meladema* (*M. coriacea*, *M. imbricata* and *M. lanio*) from 38 localities covering the whole range of all species (Fig. 1; see Table S1 for a list of specimens included in the study). Specimens were collected and preserved in absolute ethanol directly in the field. We extracted the DNA non-destructively with commercial kits (mostly ‘DNeasy Tissue Kit’; Qiagen GmbH, Hilden, Germany) following the manufacturer’s instructions. Specimens and DNA extractions are kept in the collections of the Institut de Biologia Evolutiva, Barcelona (IBE) and the Museo Nacional de Ciencias Naturales, Madrid (MNCN). We used the data of Ribera et al. (2003), to which we added 17 newly sequenced specimens, especially from areas not previously studied in the Central and Eastern Mediterranean, and nuclear data. Six gene fragments from five different genes (three mitochondrial and two nuclear) were obtained in four different amplification reactions: (i) 3’ end of Cytochrome Oxidase Subunit 1 (*COI-3’*); (ii) 5’ end of 16S tRNA plus tRNA transfer of Leucine plus 3’ end of NADH subunit 1 (*16S + tRNA-Leu + nad1’*); (iii) an internal fragment of the nuclear gene Histone 3 (*H3’*); (iv) a fragment of the nuclear gene Wingless (*Wg*) (see Table S2 for the primers used and general cycling conditions). The new 125 new sequences have been deposited in the EMBL database with accession numbers LT602717-LT602841 (Table S1). Sequences of outgroup taxa were downloaded from GenBank: a representation of other genera of subfamily Colymbetinae, including the closest known relatives of *Meladema* (Morinière et al. 2015, 2016), and one genus of Agabinae (*Agabus*) to root the tree, which is likely the sister subfamily of Colymbetinae (Ribera et al. 2008; Miller & Bergsten 2014). DNA sequences were assembled and edited using GENEIOUS v1.6 (Biomatters, http://www.geneious.com). Alignments of variable length sequences were obtained with the MAFFT plugin v7.017 (Katoh et al. 2002) in GENEIOUS v1.6.

The likely presence of cryptic mitochondrial lineages within *Meladema* (Ribera et al. 2003) raised the possibility that specimens could be infected with *Wolbachia*, a maternally transmitted parasite that can alter the patterns of mtDNA variability (Jiggins 2003). We thus tested for the presence of *Wolbachia* in 11 selected specimens of the two
lineages of *Meladema coriacea* (Table S1). We used specific primers for the *wsp* gene, trying to amplify a 632-bp fragment. A 3 µl sample of the PCR reaction mixture was electrophoresed with a 100-bp DNA ladder on 1% agarose gel to determine the presence and size of the amplified DNA bands that were visualized by Sybr-Safe staining. The same methods have yielded positive results, thus detecting the presence of *Wolbachia* infested individuals, in another genus of the same family (*Deronectes*, García-Vázquez & Ribera 2016).

**Data sets for molecular analyses**

We used three different data sets in our analyses: (i) phylogenetic and divergence time analysis with the complete, combined sequence and with only nuclear data; (ii) coalescence analysis with only mitochondrial data; (iii) phylogeography and climatic niche reconstruction with mitochondrial and nuclear data.

**Phylogenetic and divergence time analysis**

For the phylogenetic and divergence time analysis, we used only specimens with different haplotypes of the *COI-3* gene, plus the corresponding sequences of the rest of the genes. The final length of the data matrix was 2439 nucleotides for 52 specimens. Two specimens of *M. coriacea* known to have mtDNA haplotypes of *M. imbricata* and three specimens of *M. imbricata* from Tenerife were also excluded from the analyses. These specimens were identified in preliminary phylogenetic analyses performed with sequences of all available specimens of the genus *Meladema* (see Results below and Ribera et al. 2003).

The data set was divided into five partitions corresponding to each gene, and we used Partition FINDER v1.1.1 (Lanfear et al. 2012) to estimate the evolutionary model that best fitted the data for each partition separately, using Bayesian Information Criterion (BIC) as selection criteria. We analysed the data with BEAST 1.8 (Drummond et al. 2012) linking substitution models in two partitions corresponding to the mitochondrial (COI-3’, 16S + tRNA-Leu (16S onwards) and *nad1*) and nuclear (*H3* + *Wg*) markers, respectively. Molecular clock models were linked in three partitions corresponding to the mitochondrial protein coding genes (*COI-3’ + nad1*), 16S and nuclear genes (*H3* + *Wg*), and we ran two analyses to test the clock model (strict or lognormal relaxed) that best fitted the data. We applied an *a priori* rate of 0.0145 substitutions/site/MY (standard deviation 0.002) for the protein coding genes and 0.0016 substitutions/site/MY (standard deviation 0.0002) for 16S. These are substitution rates estimated for a related group, family Carabidae, based on a combination of fossils and biogeographic events (Andújar et al. 2012). Clock rates of *H3* and *Wg* were left with uniform priors due to the absence of any suitable estimations of the evolutionary rate for these nuclear genes. We used a Yule speciation prior for all of the partitions and ran the analysis for 100 million generations, logging parameters every 5000 generations. Unless otherwise stated, we used a conservative burn-in of 10% for combined data sets of both mitochondrial and nuclear DNA, after checking convergence of all parameters in TRACER v1.6 (Rambaut et al. 2014).

We performed four runs with different topological constraints: (i) *M. imbricata + M. lanio* as monophyletic; (ii) *M. coriacea* as monophyletic; (iii) *M. lanio, M. imbricata* and *M. coriacea* (only specimens from Corsica, Sardinia and Montecristo) as monophyletic; (iv) without any constraints. We then compared tree likelihoods in TRACER v1.6. A consensus tree was obtained via TREE ANNOTATOR v1.8 (Drummond et al. 2012) and visualized using FIGTREE v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

We also analysed the nuclear genes only to test for potential incongruence between the nuclear and mitochondrial genomes. The data set for the nuclear sequence contained 46 specimens and the aligned sequence was 805 nucleotides long, and was analysed with a fast maximum-likelihood algorithm implemented in RAxML (Stamatakis et al. 2008) with a single partition and a GTR model. Node support was assessed with 500 fast bootstrap replicas.

**Coalescence analysis**

We used all available mitochondrial sequences of the continental *M. coriacea* for analyses of coalescence and Bayesian
Skyline Plots (Drummond et al. 2005) to estimate the population history of this species. A similar methodology has been successfully applied to relate demographic models and geographical expansion of other water beetles species complexes (Hidalgo-Galiana et al. 2014). The aligned nucleotide matrix of 35 specimens was 1632 nucleotides long. Both coalescence and Bayesian Skyline Plot analyses were carried out in BEAST 1.8. The data set was divided into three partitions corresponding to each mitochondrial gene (COI-3', 16S and nad1), with a GTR substitution model. We estimated the best clock model and set an a priori rate of 0.0134 substitutions/site/MY with a standard deviation of 0.001 for the combined sequence, estimated for the family Carabidae (Andújar et al. 2012). Preliminary analyses had shown that if the sequence was split into different partitions, there were severe convergence problems, so we used a single concatenated partition. To identify the best demographic model, we ran four analyses, including constant size, exponential growth, logistic growth and expansion growth. Models were run for 100 million generations. The estimated optimal coalescent model for 100 million generations, using a principal component analysis (PCA) to obtain uncorrelated environmental factors (Varimax rotation).

Values of the first two PCA factors were used to plot a bidimensional climatic space of the study area (e.g. Sánchez-Fernández et al. 2013) and to locate sampling localities of each lineage of Meladema, including the Italian mainland localities obtained from CKMAP (Table S3). To test for significant differences in climatic conditions among Meladema species, we also conducted one-way ANOVA analyses of the first three PCA factors with post hoc tests using the Bonferroni correction. All statistical analyses were conducted using SPSS 15.0.1 and STATISTICA v8.0 (www.statsoft.com, 2007).

To compare the climatic niche between lineages, we generated ecological niche models using MAXENT v3.3.3 (Phillips et al. 2006). To avoid autocorrelation in the variables, we also used the three-first factors of the climatic PCA as predictors. We included all known localities for all lineages except for continental M. coriacea, for which we included only localities with sequenced specimens to avoid uncertainties. The localities with molecular data are widespread in the climatic space, providing a good geographic and climatic subsample of all known records (Fig. 1; Tables S1, S3), minimizing the risk of bias due to unbalanced climatic and spatial sampling effort. MAXENT is a machine-learning method that estimates organisms’ distribution by finding the probability distribution of maximum entropy (i.e. the most uniform), given the constraint that the expected value of each environmental predictor under this estimated distribution matches the empirical average of sample locations (Phillips et al. 2006). MAXENT simulations of realized distributions produce continuous suitability scores for each cell (from 0 to 1), and we calculated Schoener’s D metric (Schoener 1968) using ENMTools (Warren et al. 2010) to assess niche similarity among populations, where $D (p_X, p_Y) = 1 - \frac{1}{2} S (p_X, p_Y)$. This metric assumes probability distributions defined over geographic space, in which $p_X$ (or $p_Y$) denotes the probability assigned by the modelling method to species $X$ (or $Y$) in the $i$ cell ranging from 0 (niches do not overlap) to 1.
(niches completely overlap). The null hypothesis of the background similarity test states that observed niche overlap between taxa is explained by regional similarities in available background environments. This hypothesis involves a two-tailed test, so it is rejected if the observed $D$ between two taxa falls outside the 95% confidence limits of the null distribution. Niche conservatism is supported when the observed value of $D$ is larger than the upper 95% confidence limit of the null distribution, suggesting that niches are more similar than expected based on their background environments (i.e. species are occupying niches that are as similar as possible given what is available). Niche divergence is supported when the observed value of $D$ is smaller than the lower 95% confidence limit of the null distribution, suggesting that niches are more divergent than expected based on background divergence. Because niche differences may be simply result of the spatial autocorrelation of the used explanatory environmental variables (background environmental divergence, Warren et al. 2008), strong evidence for niche divergence requires two conditions: (i) that niche characteristics differ between the two considered populations and (ii) that these differences are greater than the background environmental divergence (McCormack et al. 2010). Niche conservatism, on the other hand, would be supported if niche differences were smaller than the obtained background environmental divergence. Thus, comparison of environmental characteristics from these two classes of data should allow discrimination between differences as a result of simple spatial autocorrelation caused by geographic distance and strong niche divergence that occurs because two species occupy different environmental conditions. To test the null hypothesis that niches are similarly divergent in comparison with background environments, we used the 'background similarity test’ procedure implemented in ENMTools, estimating 100 niche overlap values generated by comparing model suitability values of one population to those generated from random cells drawn from the geographic range of the other population (Warren et al. 2010). The background area of each species should be adjusted to the habitat available and should be biologically realistic (Warren et al. 2010). As three of the lineages within the genus are insular (see Results), the background area was restricted to the islands in which the species appear. For continental $M$. coriacea, we defined the background as the whole study area.

We also performed an ancestral character state reconstruction analysis in BEAST v1.8 to estimate ancestral climatic preferences of individual Meladema lineages. All substitution models and settings were the same as in coalescence analyses above. We created another partition representing the first three best scoring PCA factors used in previous PCA analysis as a continuous trait, with a homogenous Brownian model of evolution. We ran the ancestral character state reconstruction analysis for 100 million generations with parameters logging every 5000 generations. For each of the main nodes, we identified the geographic area with climatic conditions encompassed by the 95% interval of the reconstructed value for all three PCA factors.

**Results**

**Data sets for molecular analyses**

There were no length differences in protein coding genes. The only length differences in alignment were three insertions in ribosomal genes in outgroup taxa and one single nucleotide deletion in several specimens of $M$. coriacea.

The GTR + I + G (Tavaré 1986) was identified as the best substitution model for the COI-3', 16S and nad1 partitions, and the TRN + I + G (Tamura & Nei 1993) for the $H3$ partition. For the $Wg$ partition, the K80 + I (Kimura 1980) was identified as the most suitable, but as it is not implemented in BEAST 1.8, we used the HKY model instead (Hasegawa et al. 1985), with equal base frequencies (Drummond et al. 2012). The same substitution models were used for each gene in all data sets.

**Phylogenetic and divergence time analysis**

The strict clock was significantly better than the lognormal clock when using the combined data (Table 1), and in consequence we used the strict clock in all analyses. The comparison of likelihood scores among trees with topological constraints showed no significant differences between likelihood scores of the tree without any constraints and trees with constrained clades (with <5 units of AICM between the most divergent values, Table 2), so we used the topology without any prior constraints.

The analysis of phylogenetic relationships including only specimens with unique haplotypes supported a sister relationship of Corsican, Sardinian and Montecristo specimens of $M$. coriacea ($M$. coriacea CSM onwards) and the rest of the genus (Bayesian posterior probability, pp = 1; Fig. 2). The rest of specimens of $M$. coriacea (continental $M$. coriacea onwards, although including specimens from the Balearic and Canary Islands and Malta) were placed as

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Clock model</th>
<th>Likelihood score</th>
<th>AICM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phylogeny and dating</td>
<td>Lognormal clock</td>
<td>7875.6</td>
<td>15994.9</td>
</tr>
<tr>
<td>Phylogeny and dating</td>
<td>Strict clock</td>
<td>7889.8</td>
<td>15962.3</td>
</tr>
<tr>
<td>Phylogeny (nuclear genes)</td>
<td>Lognormal clock</td>
<td>2218.6</td>
<td>4494.6</td>
</tr>
<tr>
<td>Phylogeny (nuclear genes)</td>
<td>Strict clock</td>
<td>2240.5</td>
<td>4560.6</td>
</tr>
</tbody>
</table>

AICM, Akaike information criterion. Lower AICM values indicate a better model fit.
sister of a clade including *M. imbricata* and *M. lanio* (Fig. 2), but with a low support (pp < 0.50). On the contrary, the monophyly of *M. imbricata* plus *M. lanio* was strongly supported (pp = 0.96), as well as monophyly of, respectively, *M. lanio* (pp = 1), *M. imbricata* (pp = 1), *M. coriacea* CSM (pp = 1) and the continental *M. coriacea* (pp = 0.98) (Fig. 2).

The dating analysis using an *a priori* rate obtained for the family Carabidae showed a very long branch between the ingroup and outgroup taxa, with an estimation for the origin of the genus *Meladema* at ~14.4 Ma (million years ago) (95% CI of 9.99–20.02 Ma), in the Miocene (Fig. 2). The crown age of the genus *Meladema* was estimated at ~1.49 Ma (95% CI of 1.01–2.05 Ma), with the youngest speciation between *M. imbricata* and *M. lanio* dated to ~1.17 Ma (95% CI of 0.77–1.69 Ma), in a relatively short time period in the Pleistocene (Fig. 2).

**Table 2** Likelihood and AICM scores for the BEAST analyses of the combined data with different topological constraints under a strict clock, calculated in TRACER.

<table>
<thead>
<tr>
<th>Topological constraint</th>
<th>Likelihood score</th>
<th>AICM</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Meladema lanio</em> + <em>Meladema imbricata</em></td>
<td>7890.2</td>
<td>15964.0</td>
</tr>
<tr>
<td><em>Meladema coriacea</em></td>
<td>7889.3</td>
<td>15959.5</td>
</tr>
<tr>
<td><em>M. lanio</em> + <em>M. imbricata</em> + <em>M. coriacea</em> CSM</td>
<td>7889.9</td>
<td>15959.4</td>
</tr>
<tr>
<td>No constraints</td>
<td>7889.8</td>
<td>15962.3</td>
</tr>
</tbody>
</table>

AICM, Akaike information criterion. Lower AICM values indicate a better model fit.

---

**Fig. 2** Reconstructed phylogeny of *Meladema* obtained with BEAST with the combined nuclear and mitochondrial sequence. Numbers at nodes, estimated age (Ma); coloured points, node support. See Table S1 for details of the specimens.
For nuclear partitions, the lognormal relaxed clock was considered as a more suitable clock model (Table 1). In the H3 nuclear protein coding gene sequence, there were only two variable positions, one synonym change synapomorphic for M. imbricata plus M. lanio and another with ambiguous callings in most individuals, without a clear pattern. In the Wingless nuclear protein coding gene sequence, there were three variable nucleotide positions with changes shared by: (i) M. imbricata + M. lanio; (ii) M. coriacea CSM and (iii) M. coriacea CSM + M. imbricata + M. lanio. As happened in H3, there was more variability among outgroup taxa. The monophyly of M. coriacea CSM and the sister relationship of M. imbricata and M. lanio were thus corroborated by nuclear DNA (Fig. S1).

**Coalescence analysis**

The best demographic model for the continental M. coriacea was an expansion growth, although with an AICM almost identical to the constant size model (Table 3). The analysis using the expansion growth model dated the age of the most recent common ancestor (MRCA) of the continental M. coriacea lineage to ~145 thousand years before present (Ka) (95% confidence interval 71–246 Ka), in the Late Pleistocene (Table 3). All analyses using different demographic models estimated a Late Pleistocene age of the MRCA of continental M. coriacea (Table 3).

The Bayesian Skyline Plot showed a continuous increase in the population size of continental M. coriacea during the last 56 Ka, with a slight decrease towards the present (Fig. S2).

**Phylogeographic analysis**

The phylogeographic analysis inferred a continental origin of Meladema, although the lack of well-defined outgroups did not allow a precise reconstruction. Macaronesian and Mediterranean islands were colonized independently from the continent. There were two main clades within the M. coriacea CSM lineage, one with exclusively specimens from Corsica and a second with specimens from Corsica, Sardinia and Montecristo. The basal splitting events within continental M. coriacea took place in north-western Africa, with subsequent expansions through the western Mediterranean and possibly several secondary colonizations of the Canary Islands. The eastern most sampled specimens (Malta and Turkey) were not derived from the northern lineages (Iberian peninsula and southern France), but originated directly from within the north African stock (see Data S1 for a phylogeographic reconstruction visualized in Google Earth).

**The presence of Wolbachia**

None of the 11 tested specimens was positive for the presence of Wolbachia.

**Environmental niche modelling**

The main two factors of the climatic PCA jointly accounted for a high percentage of the total variance (78.7%; the first three factors 85.8%). The first factor was positively correlated with annual mean temperature and negatively with annual precipitation and was interpreted as representing an ‘aridity’ gradient with higher values for hot and dry sites. The second axis was negatively correlated with temperature seasonality and was interpreted as representing a ‘seasonality’ gradient. The third factor (with <10% of variance explained) showed a positive correlation with precipitation of the coldest quarter and negative with mean temperature of the wettest quarter, and its interpretation was less clear. The environmental space of continental M. coriacea almost covered that of all other lineages, with the only exception of M. lanio (Fig. 3). Among the island lineages, M. lanio and M. imbricata were closer to each other, although M. lanio occupied sites with lower seasonality, while M. coriacea CSM occupied relatively more seasonal sites (Fig. 3). The environmental space of the peninsular Italian M. coriacea fitted into the space occupied by continental M. coriacea and M. coriacea CSM, with a higher seasonality than in M. imbricata and M. lanio (Fig. 3).

Results of the ANOVA analyses showed differences among the four main lineages for the first three main PCA factors (Fig. 4). Post hoc analyses with Bonferroni corrections showed that continental M. coriacea occurs in significantly more arid localities than M. coriacea CSM (first factor), and M. lanio in significantly less seasonal localities than continental M. coriacea and M. coriacea CSM (second factor; Fig. 4).

Results of the MAXENT analysis estimated by Schoener’s D showed a low niche overlap among the four main lineages (Fig. S3), with the highest between continental M. coriacea and M. coriacea CSM (0.48) and M. imbricata and M. lanio (0.41). The lowest overlaps were between M. lanio

---

Table 3 Ages (in thousands of years before present) of the MRCA of continental Meladema coriacea estimated by the coalescence analysis, with the 95% confidence intervals of the MRCA ages and likelihood and AICM scores of individual demographic models

<table>
<thead>
<tr>
<th>Demographic model</th>
<th>MRCA age (Ka)</th>
<th>95% CI (Ka)</th>
<th>Likelihood score</th>
<th>AICM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant size</td>
<td>0.170</td>
<td>(0.084–0.281)</td>
<td>2327.1</td>
<td>4692.2</td>
</tr>
<tr>
<td>Logistic growth</td>
<td>0.117</td>
<td>(0.061–0.194)</td>
<td>2328.7</td>
<td>4698.7</td>
</tr>
<tr>
<td>Expansion growth</td>
<td>0.145</td>
<td>(0.071–0.246)</td>
<td>2327.2</td>
<td>4692.1</td>
</tr>
<tr>
<td>Exponential growth</td>
<td>0.123</td>
<td>(0.061–0.209)</td>
<td>2327.7</td>
<td>4697.5</td>
</tr>
</tbody>
</table>

AICM, Akaike information criterion; MRCA most recent common ancestor. Lower AICM values indicate a better model fit.

© 2017 Royal Swedish Academy of Sciences
and both lineages of *M. coriacea*, continental (0.08) and CSM (0.09) (Fig. 5; Table S4). In the background similarity tests, the null hypothesis of no niche difference was rejected in both directions in four of the six pairwise comparisons. In the other two, the comparison of *M. imbricata* against *M. coriacea* CSM and of *M. lanio* against *M. coriacea* CSM only one direction showed niche divergence (Fig. 5).

Reconstruction of ancestral climatic preferences

We obtained the 95% Highest Posterior Density (HPD) intervals in the BEAST analysis for the first two PCA factors (Table S5). The analysis showed that all studied specimens of *M. coriacea* CSM, *M. imbricata* and *M. lanio* occurred in climatic conditions reconstructed as being ancestral for the whole genus *Meladema* (Fig. 6). On the contrary, some current localities of continental *M. coriacea* (mostly in Morocco, Spain and Turkey) did not fit into the 95% HPD interval for at least one reconstructed PCA factor, suggesting occurrence in non-ancestral conditions for the genus (Fig. 6; Table S5). When considering the ancestral conditions of each lineage individually, for all insular lineages, the climatic conditions of the sampled localities fall within the 95% HPD interval of the ancestral conditions of the same lineage only. This is in contrast to some specimens of continental *M. coriacea*, which fall within the 95% HPD interval of other lineages (Fig. 6; Table S5).

Discussion

Phylogeny and age of *Meladema*

Previous work on the genus *Meladema* had revealed the existence of a fourth, cryptic lineage in addition to the three recognized species (Ribera et al. 2003). This lineage was apparently morphologically indistinguishable from the widespread *M. coriacea*, but found only in Corsica. This study had, however, two important shortcomings: it was based on mitochondrial data only, and the geographical sampling was very incomplete. Our results using both mitochondrial and nuclear DNA and a more comprehensive sampling corroborated the existence of this cryptic lineage within the genus. Although we found very small amount of variability in the nuclear genes used in our analyses, the variable nucleotide positions within them supported the existence of a distinct lineage of *M. coriacea* from Corsica, Sardinia and the island of Montecristo, in the Tuscan archipelago (*M. coriacea* CSM). The genus and each of four previously suggested lineages were strongly supported as monophyletic, although the internal relationships within the genus were not resolved reliably except for the sister relationship between the Macaronesian species *M. lanio* and *M. imbricata*.

The origin of *Meladema* was dated back to the Middle Miocene. However, the diversification among extant *Meladema* lineages started much later, in the Early Pleistocene. According to our estimate, these events took place in a relatively short time period, suggesting the possibility of a nearly simultaneous speciation. The poor support for the relationships among the lineages with an exception of the sister relationship of *M. lanio* and *M. imbricata* would also imply that the speciation took place in a short sequence of events. These age estimates are very close to those suggested by the analysis of Ribera et al. (2003), which used a standard rate of variation for the mitochondrial genes of
It is likely that in the case of One explanation could be a difference in population sizes. This could have several possible reasons. Another explanation could be an occasional gene flow between the CSM lineage and continental M. coriacea preventing morphological differentiation, while not being enough to leave a clear signal in mtDNA. However, we do not have any clear evidence for either of these two scenarios (which are not mutually exclusive). A third possible explanation for the absence of morphological differences between M. coriacea CSM and continental M. coriacea could have been the presence of Wolbachia, which can cause cytoplasmic incompatibility and thus reproductive isolation (Werren et al. 2008). Recent works gave an estimate of 40% of Wolbachia prevalence among arthropods (Zag & Hammerstein 2012) and 31% in a group of families of aquatic beetles, among them Dytiscidae (Sontowski et al. 2015), a frequency agreeing with recent reports of the same family (García-Vázquez & Ribera 2016). However, none of the specimens of Meladema tested was positive for the presence of these endosymbiotic bacteria.

Phylogeography and climatic preferences of Meladema

Our phylogeographic analysis suggests the origin of Meladema in the western Mediterranean, most probably in southern Iberia or in north-western Africa, during the Middle Miocene. This implies that for most of their evolutionary history, the ancestors of extant Meladema lineages evolved under the tropical climate present in southern Europe and North Africa during that time (Ruddiman et al. 1989; Griffin 2002; Jiménez-Moreno et al. 2010; Köhler et al. 2010). The climatic change from wet tropical to Mediterranean, with a stronger seasonality and dry summers, occurred at ~3.2 Ma, in the Pliocene (Suc 1984). This could have confined the ancestral Meladema to regions with a climate more similar to that in Miocene. Our reconstruction of ancestral climatic preferences implies that M. lanio, M. imbricata and M. coriacea CSM are probably relicts inhabiting sites with conditions closer to the ancestral for the whole genus. In contrast, continental M. coriacea was the only lineage able to colonize also sites with climatic conditions non-ancestral for Meladema. Environmental niche modelling analyses showed that seasonality is the most important factor in current distribution of Meladema. All island lineages (M. lanio, M. imbricata and M. coriacea CSM) inhabit on average less seasonal sites than continental M. coriacea, which is the lineage occupying the highest diversity of climatic conditions. Among the island lineages, endemic Macaronesian species inhabit the least
seasonal sites, while *Meladema* CSM generally occupies moderately seasonal localities. The estimated time of the colonization of Macaronesia by the ancestors of *M. lanio* and *M. imbricata* clearly excludes the possibility of them being tertiary relicts of laurisilva forests on Madeira or the Canary Islands, as already suggested by Ribera *et al.* (2003).

There are several possible scenarios for the colonization of these archipelagos. One is that both were independently colonized.
colonized from the continent; the Canary Islands (M. imbricata) likely from north-western Africa, and Madeira (M. lanio) from south-western Iberia. During the Pleistocene, there were several islands between Madeira and continental Europe, even during the Last Glacial Maximum (Fernández-Palacios et al. 2011). The two colonizations could have originated from a common continental stock or they may have previously speciated, reaching their recovered sister relationship through a process of coalescence with the extinction of some of the continental lineages. A second possibility is that there was a single colonization process from the continent, either to Madeira or the Canary Islands, with subsequent dispersal between islands. However, as Meladema is absent on Salvages and no other islands were probably present between Madeira and Canary Islands in the past (Fernández-Palacios et al. 2011), the dispersion between both archipelagos seems rather unlikely.

The estimated age of the M. coriacea CSM lineage (Pleistocene) indicates that the isolation occurred much later than the end of the Messinian Salinity Crisis, which lasted from 5.93 to 5.33 Ma and was the last time when islands in Tyrrenian Sea were connected to the continental Europe (Krijgsman et al. 1999; Manzi et al. 2013; Pérez-Asensio et al. 2013). This suggests that ancestral Meladema were able to colonize islands across the sea, probably either from France or from Italy. Although we sampled only Corsica, Sardinia and Montecristo, it is likely that also other islands such as Pianosa, Capraia or Elba are inhabited by M. coriacea CSM (Rocchi et al. 2014). On the contrary, Malta (and likely Sicily, from where no specimen could be obtained) were likely colonized through other routes.

After the speciation, continental M. coriacea started to spread to north-western Africa, Iberia, southern France, Turkey and several Mediterranean islands (Malta, Balearic Islands), including one or more secondary recent colonizations of the Canary Islands. The phylogenetic assignment of continental Italian, Sicilian and Balkan M. coriacea remains uncertain, as no specimens from these regions could be sequenced. In the case of Greece, M. coriacea was reported for the south, in the Attica region and the Cyclades (Oertzen 1886). As the Turkish specimen from the Izmir province included in our analysis is nested deep within the continental M. coriacea clade, it seems that also Bulgaria, Greece and Cyclades are likely inhabited by the same lineage. The phylogenetic assignment of M. coriacea from mainland Italy is more problematic. According to the ‘Checklist of Italian Fauna’ database (CKMAP2000; http://www.faunaitalia.it/ckmap/), M. coriacea is widely distributed in the Italian peninsula and Sicily, with the northernmost records from Liguria. The fact that M. coriacea CSM
inhabits islands in the Tuscany archipelago very close to the Italian mainland (Montecristo) cannot be taken as evidence that Italian Meladema also belong to the CSM lineage. These islands are known to be inhabited by some Corso-Sardinian endemics not found in mainland Italy (e.g. Agabus aubei Perris, 1869, Agabus rufulus Fairmaire, 1859 or Rhithrohydus sexguttatus Aubé, 1838 among the Dytiscidae, Rocchi et al. 2014). Moreover, Meladema has never been found in North Africa east of Tunisia (Touaylia et al. 2010), or in the Near East but Turkey. This suggests that M. coriacea reached western Turkey and Greece probably from southern Italy or the Balkans, which would thus be inhabited by the same lineage (continental M. coriacea). Results of niche similarity test do not provide clear evidences from a climatic niche perspective. The single specimen from Malta belongs to the continental M. coriacea lineage, suggesting that the Sicilian populations may also belong to the same lineage. The aquatic Coleoptera fauna of Malta is known to be strongly related to that of Sicily and mainland Italy (e.g. among the Hydraenidae, with a high degree of local endemism, all species known from Malta occur also in Sicily, Mifsud et al. 2004), which can also be taken as evidence of the absence of M. coriacea CSM outside the northern Tyrrhenian islands.

We can only speculate what is the reason of the wide distribution of continental M. coriacea through the Mediterranean, while all other lineages are exclusively restricted to different archipelagos. Similarly to all other Meladema lineages, continental M. coriacea could have initially been restricted to sites with climate resembling the wet tropical climate of the Miocene, but developed capabilities necessary for survival during dry periods after the climatic change from a wet tropical to Mediterranean climate 3.2 Ma (Suc 1984). To test this hypothesis, it would be necessary to obtain experimental data on the physiological tolerance of the different species. However, it could also be that continental Meladema expanded its range, at least initially, just because it was in the continent, that is, without geographical barriers to impede its displacement, similarly to what is hypothesized to have happened in other groups of aquatic beetles during the interglacials (García-Vázquez & Ribera 2016). Only at a later stage, it may have acquired the dispersal or physiological capabilities necessary to cross the Straits of Gibraltar or to re-colonize the Canary Islands.

Conclusions
Our results strongly support the existence of a cryptic lineage of Meladema on the northern Tyrrenian islands. Thus, M. coriacea as currently understood is not a monophyletic species. In contrast, the Macaronesian species M. larnio and M. imbricata were found, respectively, monophyletic and sisters. The speciation among extant lineages likely took place in a very short time period in the Early Pleistocene. This could have been the result of the change from a tropical climate in the Miocene to the typical Mediterranean seasonality in the Pliocene. Climate niche modelling showed that seasonality is the most important factor in the distribution of Meladema. All three island lineages inhabit less seasonal sites, likely relicts of climatic conditions ancestral to Meladema. On the contrary, continental M. coriacea was able to colonize more seasonal sites and expand its range across the Mediterranean.

Acknowledgements
We especially thank S. Bouzid, A. Castro, A. Cieslak, M.G. París, A. Rudoy and R. Vila for providing specimens for study; A. Izquierdo and H. Lehmann (MNCN) and R. Alonso (IBE) for laboratory work; D.T. Bilton for discussions on our work with Meladema and two anonymous referees for comments. DG-V had a FPI PhD grant from the Spanish Government. DSF was supported by a ‘Juan de la Cierva’ postdoctoral contract from the Spanish Ministry of Economy and Competitiveness and another postdoctoral contract funded by Universidad de Castilla-La Mancha and the European Social Fund (ESF). This work has been partly funded by projects CGL2010-15755 and CGL2013-48950-C2-1-P to IR.

References
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.


Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. List of specimens included in the different analyses with specimen code, voucher, locality, geographical coordinates and Genbank accession numbers. In bold, sequences newly obtained for this study. Specimens with * were tested for the presence of Wolbachia infection

Table S2. DNA extraction and sequencing. (A) Primers used for amplification and sequencing. In brackets, length of the amplified fragment (ingroup only). (B) Standard PCR conditions for the amplification of the studied fragments

Table S3. (A) Bioclimatic variables of each locality obtained in WORLDCLIM (http://www.worldclim.org) used in the PCA (Principal Component Analysis). (B) Geographical coordinates and values of the first three axes of the climatic PCA of the four lineages of Meladema

Table S4. Pairwise values of observed niche overlap (Schoener’s D-value) between the four lineages of Meladema

Table S5. Specimens used in the reconstruction of the ancestral niche from the scores of the first two principal component analysis (PCA) axes. Overlap “YES” indicates that the two PCA scores of the climatic conditions of the locality in which the specimen was found fall within the 95% Highest Posterior Density (HPD) intervals of the reconstructed conditions of the most recent common ancestor of the respective lineage

Figure S1. Ultrametric phylogenetic tree of Meladema obtained with RAxML using only the nuclear sequence (H3 + Wg) (see Methods for details). Numbers in nodes, bootstrap support values

Figure S2. Bayesian Skyline Plot of continental Meladema coriacea showing reconstructed changes in its effective population size during the last 60 Ka. Thin blue lines, 95% confidence intervals; horizontal axis, time (1000 years ago); vertical axis, effective population size (Neff)

Figure S3. Environmental niche models (ENMs) using MAXENT for: (A) continental Meladema coriacea; (B) M. coriacea CSM; (C) Meladema imbricata and (D) Meladema lanio. Higher MAXENT probabilities (red colours) represent areas more suitable for the species according to the MAXENT models, lower values (blue) represent less suitable areas. These continuous measures of habitat suitability were used in ENM-based tests of niche similarity (see Fig 5 and Table S4)

Data S1. kml file obtained with BEAST and SPREAD to visualize in Google Earth the phylogeographic reconstruction through time of the genus Meladema (see main text for details).