


DNA barcoding of large oak-living cerambycids: diagnostic tool, phylogenetic insights and natural hybridization between *Cerambyx cerdo* and *Cerambyx welensii* (Coleoptera: Cerambycidae)

L.M. Torres-Vila^{1*}  and R. Bonal²

¹Servicio de Sanidad Vegetal, Consejería de Medio Ambiente y Rural PAyT, Junta de Extremadura, Avda. Luis Ramallo s/n, 06800 Mérida, Badajoz, Spain, ²Forest Research Group, INDEHESA, Escuela de Ingeniería Forestal, Universidad de Extremadura, Avda. Virgen del Puerto 2, 10600 Plasencia, Cáceres, Spain

Abstract

Three large saproxylic cerambycids with different pest/legal status co-occur in the Iberian oak woodlands, *Cerambyx welensii* (Cw), *Cerambyx cerdo* (Cc) and *Prinobius myardi* (Pm): Cw is an emerging pest, Cc is a protected but sometimes harmful species and Pm is a secondary/minor pest. A precise taxonomic diagnosis is necessary for research, management or protection purposes, but may be problematic mainly because Cw and Cc larvae are morphologically indistinguishable. To resolve this constraint, we genotyped adults, larvae and eggs collected over a wide geographical range using the mitochondrial barcoding of the cytochrome *c* oxidase subunit I (*COI*). A Neighbour-Joining tree phylogram revealed three distinct clusters corresponding to Cw, Cc and Pm. We further first sequenced for Cw and Cc two mitochondrial (12S rRNA and 16S rRNA) and one nuclear (28S rRNA) gene fragments. For the first two genes, interspecific divergence was lower than in *COI*, and for the 28S (lower mutation rate), the two species shared identical haplotypes. Two approaches for species delimitation (General Mixed Yule Coalescent (GMYC), Barcode Index Number (BIN)) confirmed the species distinctiveness of Cc and Cw. The Bayesian *COI* gene tree showed a remarkable genetic divergence between Cc populations from Iberia and the rest of Europe. Such divergence has relevant taxonomic connotations and stresses the importance of a wide geographical scale sampling for accurate DNA barcoding species identification. Incongruities between morphology/lineage and *COI* barcodes in some individuals revealed natural hybridization between Cw and Cc. Natural hybridization is important from a phylogenetic/evolutionary perspective in these cerambycids, but the prevalence of (and the behavioural/ecological factors involved in) interspecific cross-breeding remain to be investigated.

Keywords: BIN approach, *Cerambyx cerdo*, *Cerambyx welensii*, GMYC species delimitation, larval taxonomy, natural hybridization, *Prinobius myardi*

(Accepted 7 November 2018; First published online 5 December 2018)

*Author for correspondence
Phone: +34 924 00 25 30
E-mail: luismiguel.torres@juntaex.es; luismiguel.torresvila@gmail.com

Introduction

Man-made oak savannah-like woodlands are landscapes widespread in southern Iberia (the so-called *dehesas*) with outstanding socio-economic, ecological and biodiversity values protected under the EU Habitats Directive (CEC, 1992; Montero *et al.*, 1998; Bugalho *et al.*, 2011; Ramírez-Hernández *et al.*, 2014). Among the most important wood-boring insects associated with these oaks, three large longhorn beetles stand out: two Cerambycinae, namely *Cerambyx welen-sii* Küster, 1846 (Cw) and *Cerambyx cerdo* Linnaeus, 1758 (Cc), and one Prioninae, *Prinobius myardi* Mulsant, 1842 (Pm) (Naveiro & Morcuende, 1994; López-Pantoja *et al.*, 2008, 2016; Torres-Vila *et al.*, 2017a, b). The life cycle of these longhorns is typically associated with old, decayed or diseased oak trees, and they are included among the highly diverse assemblage of saproxylic (wood-dwelling) insects. This functional group is important in wood degradation and hollow formation processes in old trees, later used as shelters for a large array of species, so these longhorns greatly contribute to biodiversity in oak forests (Speight, 1989; Grove, 2002; Buse *et al.*, 2008; Micó *et al.*, 2015). However, since Cw and Cc are primary saproxylic beetles that need living and fresh wood to feed on and develop, they also colonize young/healthy living trees and can become harmful or pest species (López-Pantoja *et al.*, 2008, 2011; Carrasco, 2009; Torres-Vila *et al.*, 2012, 2017b; Torres-Vila, 2017).

The pest and legal status of these longhorn species differs markedly depending on the geographical and forest context. Cw is currently considered an emerging pest involved in oak decline in Iberia (Martín *et al.*, 2005; López-Pantoja *et al.*, 2008; Carrasco, 2009; Torres-Vila *et al.*, 2012, 2013); Cc is a protected species in Europe (CE, 1979; CEC, 1992; IUCN, 2010), although it has been also reported as a harmful or pest species (Torres-Vila, 2017); and Pm is a secondary or minor pest (López-Pantoja *et al.*, 2011; Torres-Vila *et al.*, 2017b). The three longhorns exhibit a typical western Palaearctic distribution with respective ranges widely overlapping in southern Europe. Main differences in geographical range are that Cw and Pm are rare or absent in central and northern Europe (attributed to their more thermophilic nature), and that Cw does not occur in North Africa (Bense, 1995; Vives, 2000; González-Peña *et al.*, 2007; Sama, 2013). In southern Iberia, Cw, Cc and Pm are widespread, share a similar ecological niche and often coexist in sympatry in the same stands, even in a single tree or branch section. Wood quality and host preference are major factors shaping larval resource partitioning between these longhorn species, although heterospecific adults may interact with one another in the same oak trunk (Torres-Vila *et al.*, 2017b).

It follows that there is a need to accurately distinguish these cerambycid species in forests harbouring mixed populations, not only for control, management or protection purposes, but also to investigate ecological aspects such as demography, habitat occupancy, larval assemblages and intraguild competition (Torres-Vila *et al.*, 2017a, b). The identification of these three species can be problematic, however, depending on species and development stage. Adults of *Prinobius* and *Cerambyx* are similar in size and brownish colour, but they can be easily recognized by their general look, pronotum shape and coxae prosternal process. Differences between Cw and Cc adults are subtler but reliable, even if identification may be sometimes difficult in intermediate or small adults, which are not

uncommon in the wild (Starzyk & Strojny, 1985; Torres-Vila, 2017; Torres-Vila *et al.*, 2018). Main diagnostic characters are elytral shape and sculpture, apex truncation, sutural spine robustness and hairiness pattern of the hind-tarsus second-segment underside. As consequence, various useful dichotomous keys for the identification of these species in adult stage exist (Picard, 1929; Villiers, 1946, 1978; Bense, 1995; Vives, 2000; Verdugo, 2004; Özdikmen & Turgut, 2009). Some studies also provide valuable information for egg taxonomy in *Cerambyx* species, which is mainly based in chorion sculpture and micropyle morphology (Hernández, 1991; Vitali, 2001; see also Duffy, 1953; Morcuende & Naveiro, 1993).

Regarding larval taxonomy, various morphological descriptions (some of them succinct or very old) for Pm (Švácha & Danilevsky, 1987; Naveiro & Morcuende, 1994), Cc (Ratzeburg, 1839 [as *Cerambyx heros* Fabricius]; Duffy, 1953; Švácha & Danilevsky, 1988) and Cw (Xambeu, 1895; Naveiro & Morcuende, 1994) exist. Larvae of *Cerambyx* and *Prinobius* may be easily differentiated because *Cerambyx* larvae present a conspicuous ferruginous-pigmented band on the pronotum frontal margin and rounded mandible tips, while in *Prinobius* larvae the frontal band is missing and mandibles are pointed (Torres-Vila *et al.*, 2017b). The major taxonomic drawback arises based on the fact that Cw and Cc larvae are practically indistinguishable using morphometric criteria, a constraint further aggravated when larvae are small or coexist in larval assemblages with other related cerambycid species. To our knowledge, there is no integrating diagnostic key allowing an unambiguous differentiation of Cw and Cc larvae, assuming that larvae of both species could actually be differentiated due to subtle morphological differences and large intraspecific variability. In fact, larval diagnosis in the Cerambycidae family below the genus (or even subfamily) level through traditional taxonomy is difficult or even impossible (Kelley *et al.*, 2006). In recent studies, we have used the alternative method of rearing field-derived larvae on artificial diet until adulthood, and then determining species membership (Torres-Vila *et al.*, 2017b, 2018), but this procedure entails constraints inherent to the long larval development (see Study species) such as mortality rates, consumption of time/resources and ultimately slow diagnosis. Larval taxonomic tools are important in ecological studies and are compulsory requirements for routine work in diagnostic laboratories and plant health services, both for an early detection of alien species and the diagnosis of pest species.

Precise and reliable larval diagnosis is a big challenge (Gossner & Hausmann, 2009) that may greatly complement classical taxonomy (Miller, 2007), even if taxonomic expertise appears to be regrettably collapsing (Hebert *et al.*, 2003). The use of certain gene sequences (usually a fragment of the cytochrome *c* oxidase subunit I, [COI]) as molecular species-specific tags for identification purposes, the so-called DNA barcoding, has revolutionized taxonomy (e.g., Hebert & Gregory, 2005; Bergsten *et al.*, 2012; Karahan *et al.*, 2017; Porter & Hajibabaei, 2018). DNA barcoding could provide a solution for the diagnostic problem described above, ideally integrating DNA sequencing not only with traditional taxonomy but also with species ecological background (Dasmahapatra & Mallet, 2006). For insects, DNA barcoding has succeeded as an effective identification tool especially in those species with no morphological keys for larvae/pupae (e.g., Ahrens *et al.*, 2007), and is widely used in agronomy, conservation and pest control (see Bergsten *et al.*, 2012 for a

review). Barcoding is supposed to allow the identification of any species at any life stage and permits a quick identification of large sample sizes to the species level.

The logic behind DNA barcoding relies on the structure of genetic variability at species level. Individuals of the same species are less genetically variable among themselves than with other species (Hebert & Gregory, 2005; Bonal *et al.*, 2011; Ratnasingham & Hebert, 2013; Porter & Hajibabaei, 2018). *COI* intra-specific divergence rarely exceeds 2% whereas interspecific one usually surpasses it (Hebert *et al.*, 2003), hence, this value has been usually the limit established for species delimitation (see review in Ratnasingham & Hebert, 2013). DNA barcoding has some pitfalls that need to be considered to avoid misidentification, including heteroplasmy, the presence of NUMTs and *Wolbachia* infections (e.g., Arthofer *et al.*, 2010). Also, hybridization artefacts among closely related species may occur when individuals externally looking like one species bear a haplotype of another related species, due to either current interspecific mating, past interbreeding or incomplete lineage sorting (e.g., Nicholls *et al.*, 2012; van Velzen *et al.*, 2012). Moreover, intraspecific genetic distance may increase along with the geographical scale, what may hamper the identification of query sequences if the reference DNA barcodes come from individuals sampled over a small range (Bergsten *et al.*, 2012, but see Huemer *et al.*, 2014). These potential problems have triggered the development of alternative methods for putative species (operational taxonomic units, OTUs) delimitation not based on fixed intra- and interspecific genetic divergence thresholds (Hebert *et al.*, 2003). The General Mixed Yule Coalescent (GMYC) model (Pons *et al.*, 2006) or the Barcode Index Number (BIN) approach implemented in the Barcode of Life Data System (BOLD) (Ratnasingham & Hebert, 2013) are among them. Both methods also allow the identification of OTUs when DNA barcodes previously assigned to Linnaean species names are available.

The overall number of DNA barcodes reported for our target longhorn beetles is extremely scarce, especially in southern Europe. For the Iberian Peninsula, no barcode sequence is publicly available so far for Cw, Cc and Pm in GenBank or BOLD (<http://www.barcodinglife.org>). Moreover, the degree of phylogenetic closeness between Cc and Cw remains largely unknown, as well as associated potential problems like incomplete lineage sorting between them or potential hybridization in the wild that might reduce the accurateness of an identification based on DNA barcodes. In this study, we genotyped specimens of the three mentioned cerambycids collected at a wide scale sampling in southwestern Spain using one nuclear and three mitochondrial genes. We specifically investigated: (i) the resolution of each genetic marker for a reliable taxonomic identification of individuals at different life stages, and (ii) the phylogenetic proximity between Cc and Cw to assess species delimitation according to the GMYC model and the BIN approach. This proximity could provoke erroneous molecular species identification if hybridization occurs in the wild, an aspect serendipitously detected in our study with major phylogenetic and selective implications.

Materials and methods

Study species

Cw, Cc and Pm adults are univoltine, flying from late May to August, and peaking in late June and early July (Cw and Cc) or a little later (Pm). Adults are large (about 25–60 mm) with a

blackish-brown body and show sexual dimorphism. Larvae are xylophagous mainly on oak species, and bore galleries into the wood that may cause huge physiological, mechanical and structural damage to host trees. Cw and Cc adults feed mainly on tree sap and exudates while those of Pm do not feed at all. Mated females lay eggs into bark crevices, pruning cuts and cork stripping wounds. After hatching, neonate larvae bore subcortically into the inner bark and then tunnel increasingly wider and longer galleries into sapwood and heartwood. Larval development usually lasts 2–4 years and pupation occurs in early (Pm) or late summer (Cw and Cc) within a pupal cell in the sapwood. Pupal stage is relatively short and lasts about 1 month. Pm adults leave the tree in the same summer to reproduce, while emerged Cw and Cc adults overwinter protected within the pupal cell in a pre-reproductive stage until the following year. Colonized trees can be identified by the presence of adult exit holes (often with frass presence evidencing larval activity) but hole morphology is not species-specific and lacks taxonomic value. Daily activity of adults (feeding, flight, mating and egg-laying) occurs typically at dusk and early evening (Duffy, 1953; Bense, 1995; Vives, 2000; López-Pantoja *et al.*, 2011; Torres-Vila *et al.*, 2016, 2017a, b).

Study site and insect sampling

The study site was the whole region of Extremadura in southwestern Spain, which extends over 41.634 km² and houses more than one million ha of oak forests, most of them dehesa open woodlands (fig. 1). The climate is typically Mediterranean with dry and hot summers (up to 40°C). Samples of the three cerambycid species were collected in different development stages for several years (2012–2016) on different oak (*Quercus*) species, holm oak (*Q. ilex* Linnaeus, 1753), cork oak (*Q. suber* Linnaeus, 1753), pyrenean oak (*Q. pyrenaica* Willdenow, 1805) and gall oak (*Q. faginea* Lamarck, 1785). Geographical coordinates were taken for each sampled tree with a sub-metric GPS device (Trimble Geoexplorer® 6000, Westminster, USA) and altitudes were obtained from geographical coordinates using the MDT05-LIDAR Digital Elevation Model (CNIG-IGN, 2016). Altitude ranged from 180 m (Guadiana and Tajo river valleys) to about 1200 m (mountains of the Central System in the northern region).

Insects were collected either with traps or dissecting the host trees. Free-living adults were captured in summer during the flight period with baited traps (mostly Cw and Cc adults as Pm adults rarely go to traps) designed to prevent the captured adults from drowning in the bait (Torres-Vila *et al.*, 2017a). Larvae and Cw/Cc overwintering adults were collected inside host trees taking advantage of a parallel study dealing with the larval ecology and assemblages of the target cerambycids (Torres-Vila *et al.*, 2017b). Collected insects were arranged individually in aerated plastic containers and carried to the laboratory. Larvae were fed on artificial diet (Morales-Rodríguez *et al.*, 2015) and reared to adulthood when necessary. Additional samples of eggs and neonate larvae were obtained in the laboratory from field-derived ovipositing females. Insect sampling was often connected with other studies on the reproductive output of the target species (Torres-Vila *et al.*, 2016, 2017a; Torres-Vila, 2017), so that for most sequenced individuals there was available information of their mating history, fertility and/or lineage.

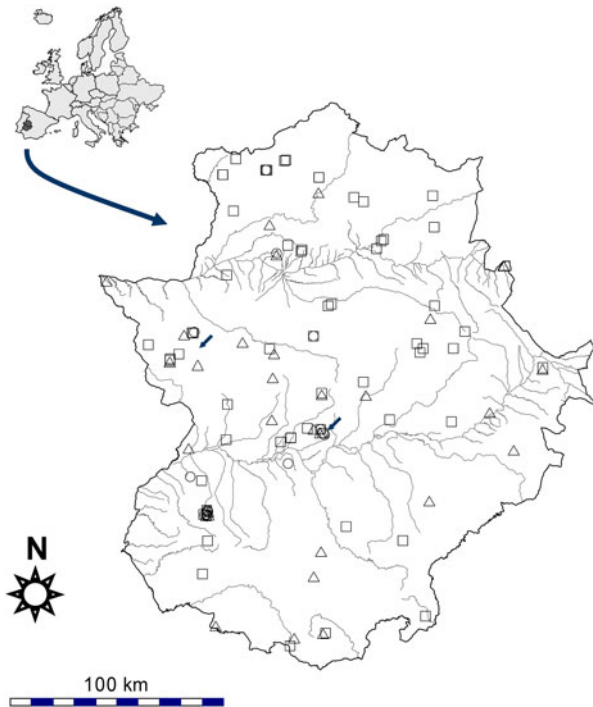


Fig. 1. Sequenced samples of *Cerambyx welensii* (white squares), *Cerambyx cerdo* (white circles), *Prinobius myardi* (white triangles) and *Cerambyx* hybrids (arrows) collected in Extremadura (SW Spain) during five consecutive years (2012–2016). Main river basins are superimposed. See table S1 for a complete key to sampling years, municipalities, sites/dehesas, host trees, UTM coordinates and altitudes.

DNA extraction and gene amplification

Samples to be sequenced were chosen to adequately cover either geographical, altitudinal or host ranges. We included in the analyses a total of 172 individuals from 83 sites located all over Extremadura (fig. 1, table S1). Among these individuals, 137 (60 adults and 77 larvae) were collected in the field and the rest (35 samples) were either adults, eggs or larvae obtained from laboratory rearing. Biological fluids as larval haemolymph and defensive oral secretions (saliva) were also occasionally sampled. A total of 69 adults from either field or laboratory (23 Cc, 27 Cw and 19 Pm) were selected for molecular analyses. These specimens were carefully identified to the species level observing their morphological features under a stereomicroscope and the available barcoding sequences obtained from these adults ($n = 62$, see Results) were used as molecular references for further identification of larvae, eggs and biological samples. For the taxonomy of Cw and Cc adults we used as main diagnostic characters elytral shape and sculpture, elytral apex truncation, apical spine robustness and the hairiness in the second segment underside of the hind tarsus (see the above-cited taxonomic references).

DNA was extracted from insect tissue following the salt extraction protocol (Aljanabi & Martínez, 1997). The body part used for the extraction depended on the development stage of each specimen. In adults, we used one leg or a half antenna. Medium- and large-sized larvae were cut transversally and one to two central segments were taken. In case of eggs and

small larvae, complete samples were processed. After removing/processing each sample, scissors, scalpels and forceps were decontaminated by cleaning and flaming with ethanol. Biological fluids (larval haemolymph and saliva) were sampled with a disposable micro-syringe. Individual samples were transferred to eppendorf tubes under 1.5 ml 96% ethanol and stored until DNA extraction.

For each sample, we amplified a 658 base pairs long fragment of the mitochondrial gene *COI* using the universal primers pair LCOI1490/HCOI2198 (see Folmer *et al.*, 1994 for details on the primer sequences and PCR protocols). However, due to the existence of sequence ambiguities at the beginning/end of some sequences, some base pairs were removed to obtain a 625 base pairs long fragment. Besides, for a selection of 12 adults (six Cc and six Cw), we also sequenced fragments of another two mitochondrial genes (12S rRNA and 16S rRNA) codifying for the small ribosomal unit (lengths 351 and 370 base pairs, respectively) (see Kambhampati & Smith, 1995 for details on the PCR reactions), as well as a fragment 631 base pair long of the D3 fragment of the 28S rRNA nuclear gene (Whiting *et al.*, 1997). We did so to assess whether a multiple gene approach could aid in the molecular diagnosis of the two *Cerambyx* congeneric species, which were expected to diverge much less between them than with Pm. All sequence chromatograms were assembled and edited using Sequencher 4.6 (Gene Codes Corp., Ann Arbor, MI, USA), and edited sequences were aligned by MUSCLE software (Edgard, 2004) as implemented in MEGA7 (Kumar *et al.*, 2016) using the default alignment parameters.

Species determination

The alignment set including all the *COI* sequences was collapsed into unique haplotypes using the online fasta sequence toolbox FaBox (Villesen, 2007). These haplotypes were used to build a Neighbour-Joining tree in MEGA7 (Kumar *et al.*, 2016). The Neighbour-Joining approach (Saitou & Nei, 1987) based on Kimura 2 Parameter model (Kimura, 1980) genetic distances is analogous to that provided in BOLD (Ratnasingham & Hebert, 2007). This analysis allows assessing whether the haplotypes corresponding to individuals morphologically classified within the same species cluster together into discrete groups separated from the others by significant genetic discontinuities (branch lengths). In addition, for the more closely related species (Cc and Cw), we calculated the pairwise distances between all distinct haplotypes using K2P model (Kimura, 1980). For the other three gene fragments (mitochondrial 12S and 16S or nuclear 28S), the K2P genetic distances between the different haplotypes of the two species were also calculated.

To confirm the species distinctiveness of Cc and Cw, we additionally used the GMYC single-locus method (Pons *et al.*, 2006; Fujisawa & Barraclough, 2013) and the BIN approach implemented in BOLD (Ratnasingham & Hebert, 2013) to delimit independently evolving evolutionary units (i.e., OTUs). For the BIN approach, the set of sequences has to be uploaded to an active project in BOLD. The query sequences are then aligned and their genetic divergence compared to assess the number of OTUs using the Refined Single Linkage (RESL) analysis. In this analysis, sequence alignments go through a Markov clustering with an optimality criterion that refines the structure of the OTUs (Ratnasingham & Hebert, 2013). Doing this it is possible to know the number of putative species, and whether they can be assigned

to an already existing OTU in BOLD (with a Barcode Index Number, BIN). These OTUs with BIN sometimes include sequences previously uploaded that correspond to a Linnaean name, thus allowing the new query sequences to be identified to the species level.

The GMYC approach does not make any aprioristic consideration of species membership either. Rather to the contrary, the model delimits a number of independent evolutionary units closely correlated with the number of species (Fujisawa & Barraclough, 2013). On the basis of a clock-constrained phylogenetic tree, GMYC method delimits distinct clusters, each of which would correspond to a different unit (i.e., species). GMYC requires an ultrametric gene tree as input and we derived this from the *COI* sequence data using *Beast* 1.7.5 (Drummond *et al.*, 2012). In this tree, we included all the distinct haplotypes obtained with Cc and Cw in our study site plus all the sequences available at GenBank/BOLD for the species within the Cerambycini tribe to which the genus *Cerambyx* belong.

To select an appropriate partitioning scheme and substitution model for the three codon positions of *COI*, we used Partitionfinder 1.1.1 (Lanfear *et al.*, 2012). We tested between the available models with *Beast* software using the Bayesian Information Criteria (BIC) and between all possible partitioning schemes. Partitionfinder software supported three separate partitions for the codon positions, and three different models were selected according to the BIC scores (TrNef + I, HKY + I and TrN + G for first, second and third codon positions, respectively). We used a strict clock model with a rate fixed to $n = 1$ and a constant size coalescent tree prior, as this could be considered conservative towards the null model when testing against the GMYC model in a likelihood ratio test. The effect of tree reconstruction method and model for the GMYC results have been investigated and, in general, a Bayesian estimation under a coalescent tree prior has performed well in comparisons (Talavera *et al.*, 2013; Tang *et al.*, 2014). Talavera *et al.* (2013) also found no difference in GMYC results between using a strict or relaxed clock model for inferring the input ultrametric tree. We ran two separate Markov chain Monte Carlo (MCMC) runs, each with 10 million generations, sampled every 1000 generations. Convergence of the two runs and effective sample size values for sampled model parameters were monitored in Tracer 1.6 as implemented in *Beast* 1.7.5 software (Drummond *et al.*, 2012). Sampled trees from the two runs were combined using Logcombiner (implemented in *Beast*) removing 25% from each run as burning, and resampling trees at half the original frequency. Treeannotator software (implemented in *Beast*) was used to select the maximum clade credibility tree (MCC tree) from the sampled trees with posterior median values used for node heights.

The MCC tree with branch lengths was imported in R version 3.3.1 (R Core Team, 2016) and the GMYC analysis was conducted using the Splits package 1.0 (Ezard *et al.*, 2009) assisted by the APE package (Paradis *et al.*, 2004). We used the single threshold method, as the multiple threshold method does not seem to improve accuracy (Fujisawa & Barraclough, 2013). The Splits package calculates the likelihood of the tree under a single coalescent (null model) and under a GMYC model with the single threshold at every node in the tree. The threshold at the maximum likelihood solution delimits the number of evolutionary units, which have been shown close correlation with the number of species in the tree (Fujisawa & Barraclough, 2013).

After species delimitation was concluded, we used POPART software (Leigh & Bryant, 2015) to build statistical parsimony networks (Templeton *et al.*, 1992) and so assess the evolutionary relationships among the *COI* haplotypes of each species.

Results

PCR success and sequencing reaction

The overall PCR success rate in our study was quite high with not significant differences between adults (93.1%) and larvae (93.7%) ($\chi^2 = 0.23$, $df = 1$, $P = 0.66$). DNA was also successfully extracted and sequenced from both Cw and Cc eggs (six out of six, 100%) and from samples of larval haemolymph/saliva (two out of three, 66%). The mitochondrial gene *COI* was sequenced in all individuals with successful PCRs, but we later discarded 10% of the sequences that had not met a minimum quality for further analyses. The total number of individuals with sequences for molecular identification was finally $n = 137$ (table S1).

DNA barcoding and species determination

The Neighbour-Joining tree based on *COI* (Kimura 2 Parameter model genetic distances) including all individuals retrieved three clusters that corresponded to Cc, Cw and Pm, respectively, and grouped together adults and larvae of the same species (fig. 2). Genetic variability within each cluster was very low and each species was separated from the remainder by long branches. In Cw and Pm, most individuals had the same haplotype, whereas in Cc, that dominance did not exist (GenBank accession numbers for each haplotype shown in table S1). In Cw and Pm, the statistical parsimony networks were star-like, with one dominant haplotype linked to many rare ones genetically very close (in all cases only one mutation away) (fig. 3).

Interspecific genetic divergence (K2P model) was in all cases higher than the minimum barcoding gap accepted for species identification (Bergsten *et al.*, 2012; Ratnasingham & Hebert, 2013), but in the case of Cw and Cc, it was just slightly over 3% (table 1). The analyses of interspecific differences in two additional mitochondrial genes (12S and 16S) confirmed the genetic closeness of Cc and Cw. In the case of these genes, which exhibit a lower mutation rate than *COI*, the maximum distance between the distinct haplotypes of each species was always below 2% (table 2) (GenBank accession numbers for 12S haplotypes MK084974 to MK084977 and for 16S haplotypes MK088074 to MK088076). In the nuclear 28S gene fragment, interspecific separation was not yet complete, as individuals of Cc and Cw even shared the same haplotype.

DNA taxonomy and species determination

As indicated, the phylogenetic tree was built using GMYC pooling together all the different haplotypes recorded in our study area for Cc (five haplotypes: H_01 to H_05) and for Cw (seven haplotypes: H_06 to H_12) together with all the sequences available at GenBank/BOLD for the Cerambycini tribe. The GMYC analyses delimited 19 distinct evolutionary units (also called GMYC taxa) that matched very well with the taxonomical species determination (Maximum Likelihood of the GMYC model = 327.66; likelihood ratio = 38.57; $P < 0.0001$) (fig. 4). Those clusters grouping several

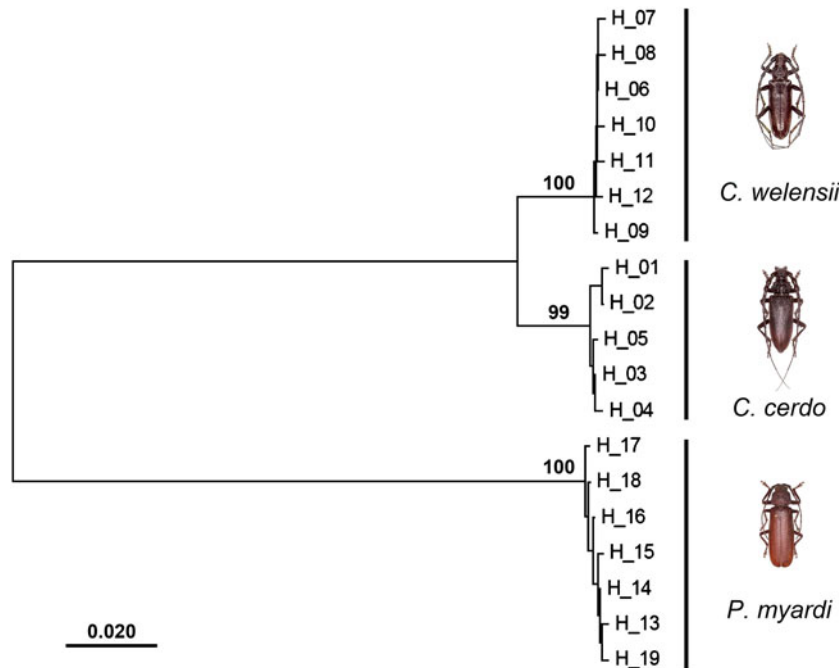


Fig. 2. Neighbour-Joining *COI* gene tree (genetic distance Kimura 2 Parameter model) depicting the relationships between the different haplotypes of *Cerambyx welensii*, *Cerambyx cerdo* and *Prinobius myardi* recorded in Extremadura (SW Spain). The percentage of replicate trees in which the associated taxa clustered together in the non-parametric bootstrap test (500 replicates) (Felsenstein, 1985) is shown next to the branches for each species cluster. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree (based on the Kimura 2 parameter model).

sequences corresponded to individuals previously identified as the same species, with the sole exception of a single specimen of *Neoplocaederus ferrugineus* Linnaeus surely due to a misidentification (fig. 4).

The GMYC species delimitation confirmed the genetic distinctiveness of Cc and Cw, which were classified as different species (fig. 4). The case of Cc showed a noticeable geographic pattern as Iberian haplotypes (H_01 to H_05) formed a cluster separated from that of the remainder European sequences (fig. 4). This fact was not observed in Cw, as the Iberian haplotypes (H_06 to H_12) grouped well with the only sequence available for Cw north of the Pyrenees (fig. 4). The BIN approach gave identical results. The *Cerambyx* sequences belonged to two different OTUs with BINs corresponding to Cc and Cw. Within each of those BINs, intraspecific divergence was below 1%, although the inclusion of the Iberian specimens of Cc increased genetic variability to almost 2%.

Hybridization between Cc and Cw and limits to DNA barcoding

All sequenced individuals were successfully ascribed to morphological species exception found in three *Cerambyx* larvae, which apparently were not adequately barcoded. To ascertain if these results were due either to an experimental error, barcoding inefficacy or other reasons, we studied each case separately.

The first unexpected result came from a laboratory larvae (Tube 79, table S1) barcoded as Cc despite its mother being morphologically Cw. The mother was collected early in the season (4-VI-2012) in a baited trap so their mating history prior to capture was unknown. Given the inconsistency between the maternal and filial phenotypes, the mother was

also barcoded (Tube 98, table S1) and subsequently scored as Cc. A second PCR and sequencing reaction for both larval and adult samples was performed to discard manipulation errors. Female longevity, fecundity and fertility had normal values when compared (controlling for female size) with baseline data from either Cw × Cw or Cc × Cc conspecific crosses (Torres-Vila *et al.*, 2016; Torres-Vila, 2017).

The second unexpected results came from two L2 larvae (Tubes 82 and 83, table S1) barcoded as Cc. Both larvae were selected among the laboratory offspring of a Cw female collected in the field (30-III-2012) overwintering inside its pupal cell (which ensured that female was virgin) in a holm oak bolt for firewood. The female had been allowed to mate just once in the laboratory with a Cw male after which eggs were laid and viable offspring produced. It follows that these two sequenced larvae were full-sibs and were barcoded as Cc despite their mother being morphologically Cw. Given the inconsistency between the maternal and filial phenotypes, the mother was also barcoded (Tube 97, table S1) and subsequently scored as Cc. Male partner was barcoded as well (Tube 99, table S1) and genetically scored as Cw in agreement with its phenotype. As in the above case, all samples were also double-checked. Female longevity and reproductive parameters also have normal values when compared with baseline data as above. These results clearly show that females in both cases were in fact hybrid specimens despite their clear Cw morphological phenotype.

Our data also allowed some rough estimates of hybridization rate. From 43 sequenced adults morphologically determined as either Cc or Cw, the mismatch between morphological and genetic species identification only occurred in the two reported hybrid females (two out of 43,

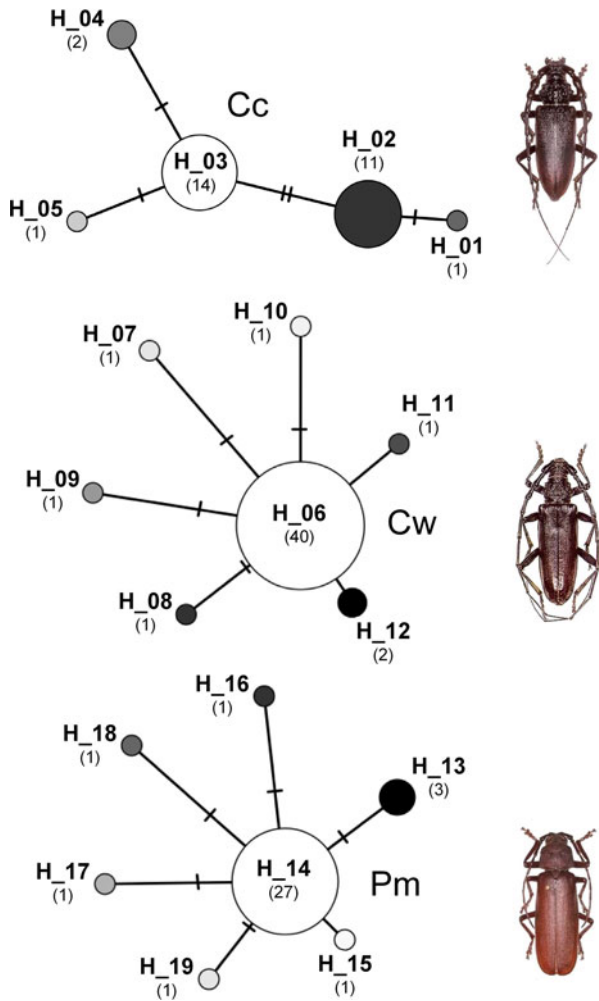


Fig. 3. Statistical parsimony networks of all mtDNA haplotypes (*COI* fragment 625 bp) of *Cerambyx cerdo* (Cc), *Cerambyx welensii* (Cw) and *Prinobius myardi* (Pm). Each circle represents a different haplotype. Haplotype reference is shown together with the number of individuals bearing it within brackets (the size of the circle is proportional to that number). The dashes crossing the connection lines show the number of mutations among haplotypes.

4.7%) (table S1). From 22 sequenced eggs and neonate/small larvae obtained from laboratory crosses between adults morphologically identified to the species level, hybrid offspring was observed in three larvae (two of them full-sibs) from two crosses (two out of 22, 9.1%) (table S1). We are aware that by using mtDNA, the hybrid rate we can detect is underestimated as only in those cases in which the mitochondrial genome does not correspond to the expected species membership morphologically determined (or the mother's species membership in the case of neonate larvae or eggs) we could say that the target individual is a hybrid.

Discussion

Our results clearly showed that Cw, Cc and Pm can be readily distinguished by means of DNA barcoding with an

Table 1. Genetic distance (Kimura 2 Parameter model) between the distinct haplotypes of the mitochondrial gene *COI* recorded for *Cerambyx welensii* (Cw) and *Cerambyx cerdo* (Cc) in Extremadura (SW Spain).

Cw haplotype	Cc haplotype				
	H_01	H_02	H_03	H_04	H_05
H_06	0.041	0.040	0.038	0.040	0.038
H_07	0.043	0.041	0.039	0.041	0.040
H_08	0.043	0.041	0.040	0.041	0.040
H_09	0.041	0.039	0.038	0.039	0.038
H_10	0.041	0.039	0.038	0.039	0.038
H_11	0.043	0.041	0.040	0.041	0.040
H_12	0.040	0.038	0.036	0.038	0.036

Table 2. Genetic distance (Kimura 2 Parameter model) between the distinct haplotypes of the mitochondrial genes 12S rRNA (A) and 16S rRNA (B) recorded for *Cerambyx welensii* and *Cerambyx cerdo* in Extremadura (SW Spain).

(A) Gene 12S rRNA		<i>C. cerdo</i>		
		H_02	H_03	H_04
<i>C. welensii</i>	H_01	0.016	0.014	0.014
(B) Gene 16S rRNA		<i>C. cerdo</i>		
		H_01	H_02	
<i>C. welensii</i>	H_03	0.020	0.017	

overall high success rate. We also show that a relatively simple method of DNA extraction, such as the salt extraction protocol (Aljanabi & Martínez, 1997), produced successful results for molecular analyses irrespective of insect development stage (adult, larvae and eggs) and can be useful for small amounts of biological fluids. This is important since these insects in the field may be collected in different life stages, and their identification allows a fine characterization of the differences in habitat selection among species taking into account the whole life cycle. The Neighbour-Joining tree based on *COI* retrieved three clusters corresponding accurately to the three target cerambycids (Cc, Cw and Pm). As expected, the distance was shorter between the congeneric species Cc and Cw (Cerambycinae), and longer in the case of Pm which belongs to a different subfamily (Prioninae). The clustering of sequences from larvae with those from adults (of known species) allowed a clear taxonomic assignment of the studied larvae. This is especially relevant in the case of Cc and Cw, as their larvae cannot be identified based on morphological traits only.

Besides the *COI*, we sequenced for the first time in Cw and Cc three more genes (the mitochondrial 12S and 16S and the nuclear 28S). The two mitochondrial ones differentiated well both species (interspecific divergence was higher than the intraspecific one), although the interspecific divergence was lower than in the case of *COI*. In the case of the nuclear 28S, there was no differentiation between species, as Cw and Cc shared haplotypes. This gene has also shown little or no variability among closely related species in other insect genera (Ács *et al.*, 2007) due to its low mutation rate. The power of nuclear rDNA expansion segments for species identification increases when more than one of these genes are combined,

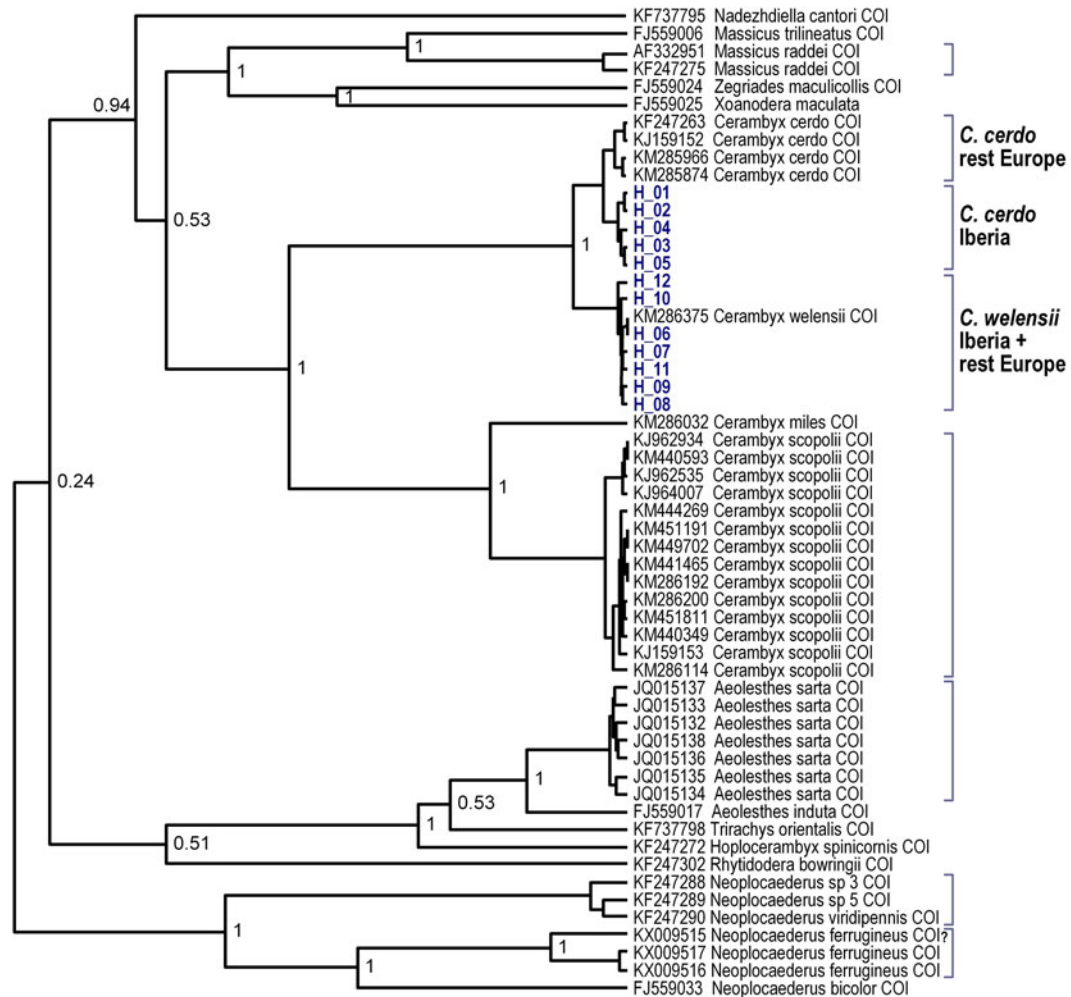


Fig. 4. Bayesian clock-constrained gene tree (COI) used for species delimitation according to the Generalized Mixed Yule Coalescent (GMYC) model. The relationship between the different haplotypes recorded in Extremadura (SW Spain) for *Cerambyx welensii* and *Cerambyx cerdo* (H_01 to H_12) with respect to the rest of the species within the Cerambycini tribe available at GenBank/BOLD is shown. The vertical bars on the right indicate sequence clusters that belong to the same putative species (OTUs). Within *Cerambyx cerdo*, two bars are shown to separate the cluster with Iberian sequences from that grouping DNA barcodes from elsewhere in Europe. The alphanumeric codes besides the scientific names correspond to the accession codes of the sequences downloaded from GenBank/BOLD databases. The model defines independent evolutionary units (i.e., species) represented by individual or multiple sequences. Support (Bayesian Posterior Probability) is shown for all nodes above species level. The specimen of *Neoplocaederus ferrugineus* marked with '?' surely corresponds to a misidentification.

but not when used isolated as in the present study (Raupach *et al.*, 2010). These results confirm that the mitochondrial *COI*, the standard gene used in DNA barcoding (Hebert & Gregory, 2005; Ratnasingham & Hebert, 2007), is also the more accurate for species differentiation between Cw and Cc.

Both the GMYC analyses and the BIN approach matched with the taxonomical species determination in the Cerambycinae tribe and confirmed the genetic distinctiveness of Cc and Cw, which were scored as different species as expected. The ultrametric gene tree showed that they are two nearby species sharing a common ancestor very close in evolutionary time, especially when compared with the common ancestor of the congeneric species *Cerambyx miles* Bonelli, 1823 and *Cerambyx scopolii* Füssly, 1775. The present study is the first providing sequences for Cc and Cw from Iberia, and showing a

dissimilar genetic background between Cc populations from Iberia and those from European countries north of the Pyrenees, with important underlying taxonomic and evolutionary connotations. These two distinct Cc clusters show that the Iberian populations of this species have gone through periods of isolation from the rest. This isolation could date back to the Pleistocene glaciations, a period in which the populations of many species linked to forests remained isolated in different Peninsulas of southern Europe (Schmitt, 2007).

Our phylogenetic data could also contribute to clarify the controversial taxonomic status of Cc at the subspecific level in the western Palaearctic (Vives, 2000; Sama, 2002; Özdikmen & Turgut, 2009; Danilevsky, 2017). This exemplifies how molecular studies may reveal the occurrence of greater biodiversity within species than previously thought from

classic taxonomic studies (Bickford *et al.*, 2007). Moreover, these results stress the problems of a reduced sampling in the reference collections for species identification using DNA barcoding (Bergsten *et al.*, 2012), even if in some insect taxa large geographic distances show small genetic impacts (Huemer *et al.*, 2014). In the case of Cw, the Iberian sequences differed little from the European references available in BOLD, and could thus be matched according to sequence identity. However, this was not the case of Cc, as the genetic distance (K2P) between any of the Iberian sequences and the references from Central Europe was always above 1%, the minimum allowed in BOLD for unequivocal species identification (Ratnasingham & Hebert, 2007). It is thus necessary to add barcodes from different geographic origin, as we have done in this study with Cc, to have a reference collection that successfully identifies query sequences at a wide geographic scale.

The occurrence of natural cross-breeding between Cw and Cc is highly consistent with their recent phylogenetic divergence. Incongruities between morphology/lineage and barcodes in some individuals revealed natural hybridization between Cw and Cc. Interspecific mating between Cw and Cc has been reported in the field (Martínez-García, 2011) and also observed by us in the wild and laboratory. Our barcoding results coupled with laboratory crosses clearly showed that: (1) natural hybridization between Cc and Cw is possible, (2) genetic introgression between both species may occur in the wild, and last but not least, (3) hybrids produced from interspecific crosses may be fertile and able to reproduce. Note that we can identify as hybrids only those individuals that had a mitochondrial *COI* haplotype that did not correspond to its morphological species identity, a fact also reported in other beetle species (Solano *et al.*, 2016). Moreover, our results may have restrictions, as these hybrids could correspond to ancient cross-breeding between species in which the offspring would have subsequently mated with the father's species (Cw in this case, as all the hybrids were morphologically Cw with Cc mitochondrial genes). We miss those cases in which the hybrid offspring mated with the mother's species, as that would leave no trace in mitochondrial DNA. Further analyses using fast-evolving nuclear markers inherited from both parents (e.g., microsatellites or SNPs) will show the full picture of hybridization between these two species.

Our results suggest that hybridization is geographically widespread since the two detected cases come from populations relatively far away (almost 100 km apart) at least for low-dispersal beetles (Torres-Vila *et al.*, 2017a). In spite of the fact that hybridization between Cw and Cc occurs in the wild, hybrid frequency is expected to be low. Pre- and post-mating interspecific barriers are acting against hybridization, among which mate recognition mediated by female-produced short-range sexual pheromones (Allison *et al.*, 2004) is expected to be prominent in *Cerambyx* as suggested by laboratory-controlled mating observations (Torres-Vila *et al.*, unpublished data). Nevertheless, the occurrence of natural cross-breeding between Cw and Cc is consistent with the high conservation in the pheromone chemistry between related species of cerambycids (Allison *et al.*, 2004; Sweeney *et al.*, 2010; Barbour *et al.*, 2011; Mitchell *et al.*, 2013; Hanks & Millar, 2013). Hence, cross-attraction between related species may be prevented or minimized just by the ratio of minor pheromone components that putatively act as synergists for conspecifics and antagonists for heterospecifics (Mitchell *et al.*, 2015). It follows that a similar pheromone blend together with a high variability in EAG

responses among adults depending on their physiological status (Sánchez-Osorio *et al.*, 2009) could contribute to explain the occurrence of interspecific matings and hybridization between Cw and Cc in the wild.

In conclusion, despite its usefulness for quick species identification and biodiversity assessments, DNA barcoding has got a number of limitations (see Introduction). Hybridization is one of them, as the barcode gene (*COI*) is in the mitochondrial genome, which is only maternally inherited. Hybrids are thus identified as the maternal species when it is not completely correct. The problem aggravates when species morphologically distinguishable nowadays interbred in the past and share the same mitochondrial genome (Nicholls *et al.*, 2012). An expected low hybrid rate minimizing the risk of larval misidentification will confirm DNA barcoding as a useful tool for species diagnosis within the *Cerambyx* genus. Our finding about natural hybridization between *Cerambyx* species is important from a phylogenetic and evolutionary perspective, but the extent in which cross-breeding occurs in the wild and the behavioural and ecological factors involved remain to be investigated.

Supplementary material

The supplementary material for this article can be found at <https://doi.org/10.1017/S0007485318000925>

Acknowledgements

The authors are grateful to all colleagues who provided technical assistance in the field and laboratory: Félix Fernández, Paco Ponce, F. Javier Mendiola, Álvaro Sánchez, Rafael López, Yolanda Conejo, Juan Gragera (SSV-Mérida), Carlos Zugasti, Jose M. de Juan (SSV-Cáceres), Juande del Pozo (SSV-Badajoz) and María Santoro (CSIC-UCLM). We are indebted to the owners of dehesas and oak forests for their kind permission to take samples. We also appreciate the constructive criticism of two anonymous reviewers. This research was supported by the Servicio de Sanidad Vegetal (SSV, Junta de Extremadura), and benefited from funding by the projects PPII-2014-01-P from Junta de Comunidades de Castilla-La Mancha/European Social Fund and AGL2014-54739-R from Spanish Ministry of Economy and Competitiveness.

References

- Ács, Z., Melika, G., Péntzes, Z., Pujade-Villar, J. & Stone, G.N. (2007) The phylogenetic relationships between *Dryocosmus*, *Chilaspsis* and allied genera of oak gallwasps (Hymenoptera, Cynipidae: Cynipini). *Systematic Entomology* **32**, 70–80.
- Ahrens, D., Monaghan, M.T. & Vogler, A.P. (2007) DNA-based taxonomy for associating adults and larvae in multi-species assemblages of chafer (Coleoptera: Scarabaeidae). *Molecular Phylogenetics and Evolution* **44**, 436–449.
- Aljanabi, S.M. & Martínez, I. (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* **25**, 4692–4693.
- Allison, J.D., Borden, J.H. & Seybold, S.J. (2004) A review of the chemical ecology of the Cerambycidae (Coleoptera). *Chemoecology* **14**, 123–150.
- Arthofer, W., Avtzis, D.N., Riegler, M. & Stauffer, C. (2010) Mitochondrial phylogenies in the light of pseudogenes and

- Wolbachia*: re-assessment of a bark beetle dataset. *ZooKeys* 56 (Special issue), 269–280.
- Barbour, J.D., Millar, J.G., Rodstein, J., Ray, A.M., Alston, D.G., Rejzek, M., Dutcher, J.D. & Hanks, L.M. (2011) Synthetic 3,5-dimethyldodecanoic acid serves as a general attractant for multiple species of *Prionus* (Coleoptera: Cerambycidae). *Annals of the Entomological Society of America* 104, 588–593.
- Bense, U. (1995) *Longhorn Beetles: Illustrated Key to the Cerambycidae and Vesperidae of Europe*. Weikersheim, Germany, Margraf Verlag.
- Bergsten, J., Bilton, D.T., Fujisawa, T., Elliott, M., Monaghan, M. T., Balke, M., Hendrich, L., Geijer, J., Herrmann, J., Foster, G.N., Ribera, I., Nilsson, A.N., Barraclough, T.G. & Vogler, A.P. (2012) The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology* 61, 851–869.
- Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K. & Das, I. (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* 22, 148–155.
- Bonal, R., Espelta, J.M. & Vogler, A.P. (2011) Complex selection on life-history traits and the maintenance of variation in exaggerated rostrum length in acorn weevils. *Oecologia* 167, 1053–1061.
- Bugalho, M.N., Caldeira, M.C., Pereira, J.S., Aronson, J. & Pausas, J.G. (2011) Mediterranean cork oak savannas require human use to sustain biodiversity and ecosystem services. *Frontiers in Ecology and the Environment* 9, 278–286.
- Buse, J., Ranius, T. & Assmann, T. (2008) An endangered long-horn beetle associated with old oaks and its possible role as an ecosystem engineer. *Conservation Biology* 22, 329–337.
- Carrasco, A. (Ed.) (2009) *Procesos de Decaimiento Forestal (la Seca): Situación del Conocimiento*. Córdoba, Spain, Consejería de Medio Ambiente, Junta de Andalucía.
- CE [Council of Europe] (1979) *The Bern Convention (19 September 1979) on the Conservation of European Wildlife and Natural Habitats, Document 104*. Strasbourg, France, Council of Europe.
- CEC [Council of the European Communities] (1992) Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora [Habitats Directive]. *Official Journal of the European Communities* 35, 7–50.
- CNIG-IGN (2016) *MDT05: Modelo Digital del Terreno con Paso de Malla 5m (SRG: ETRS89)*. Madrid, Spain, Centro Nacional de Información Geográfica, Instituto Geográfico Nacional (CNIG-IGN). Available online at <http://centrodedescargas.cnig.es/CentroDescargas/catalogo.do?Serie=LIDAR> (accessed 24 April 2018).
- Danilevsky, M.L. (2017) *A Check-List of Longicorn Beetles (Coleoptera, Cerambycoidea) of Europe*. Available online at <https://www.zin.ru/Animalia/Coleoptera/rus/danlists.htm> (accessed 24 April 2018).
- Dasmahapatra, K.K. & Mallet, J. (2006) DNA barcodes: recent successes and future prospects. *Heredity* 97, 254–255.
- Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29, 1969–1973.
- Duffy, E.A.J. (1953) *A Monograph of the Immature Stages of British and Imported Timber Beetles (Cerambycidae)*. Norwich, UK, Jarrold & Sons Ltd.
- Edgard, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32, 1792–1797.
- Ezard, T., Fujisawa, T. & Barraclough, T.G. (2009) *SPLITS: Species' Limits by Threshold Statistics*. R Package version 1.0-18/r45. Available online at <http://R-Forge.R-project.org/projects/splits/> (accessed 24 April 2018).
- Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.
- Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* 3, 294–299.
- Fujisawa, T. & Barraclough, T.G. (2013) Delimiting species using single-locus data and the generalized mixed Yule coalescent (GMYC) approach: a revised method and evaluation on simulated datasets. *Systematic Biology* 62, 707–724.
- González-Peña, C.F., Vives-Noguera, E. & de Sousa-Zuzarte, A.J. (2007) *Nuevo Catálogo de los Cerambycidae (Coleoptera) de la Península Ibérica, Islas Baleares e Islas Atlánticas: Canarias, Açores y Madeira, Monografías SEA Vol. 12*. Zaragoza, Spain, Sociedad Entomológica Aragonesa.
- Gossner, M. & Hausmann, A. (2009) DNA barcoding enables the identification of caterpillars feeding on native and alien oak. *Mitteilungen der Münchner Entomologischen Gesellschaft* 99, 135–140.
- Grove, S.J. (2002) Saprophytic insect ecology and the sustainable management of forests. *Annual Review of Ecology and Systematics* 33, 1–23.
- Hanks, L.M. & Millar, J.G. (2013) Field bioassays of cerambycid pheromones reveal widespread parsimony of pheromone structures, enhancement by host plant volatiles, and antagonism by components from heterospecifics. *Chemoecology* 23, 21–44.
- Hebert, P.D.N. & Gregory, T.R. (2005) The promise of DNA barcoding for taxonomy. *Systematic Biology* 54, 852–859.
- Hebert, P.D.N., Cywinska, A., Ball, S.L. & de Waard, J.R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences* 270, 313–321.
- Hernández, J.M. (1991) Estudio de los caracteres del huevo en diversos Cerambycidae Ibéricos y su interés taxonómico (Coleoptera). *Graellsia* 47, 49–59.
- Huemer, P., Mutanen, M., Sefc, K.M. & Hebert, P.D.N. (2014) Testing DNA barcode performance in 1000 species of European Lepidoptera: large geographic distances have small genetic impacts. *PLoS ONE* 9, e115774.
- IUCN [International Union for Conservation of Nature] (2010) *Cerambyx cerdo*. *The IUCN Red List of Threatened Species 2010: e.T4166A10502932*. Cambridge, UK, IUCN. Available online at <http://www.iucnredlist.org/details/4166/1> (accessed 24 April 2018).
- Kambhampati, S. & Smith, P.T. (1995) PCR primers for the amplification of four insect mitochondrial gene fragments. *Insect Molecular Biology* 4, 233–236.
- Karahan, A., Douek, J., Paz, G., Stern, N., Kideys, A.E., Shaish, L., Goren, M. & Rinkevich, B. (2017) Employing DNA barcoding as taxonomy and conservation tools for fish species censuses at the southeastern Mediterranean, a hot-spot area for biological invasion. *Journal for Nature Conservation* 36, 1–9.
- Kelley, M.B., Wingard, S.W., Szalanski, A.L. & Stephen, F.M. (2006) Molecular diagnostics of *Enaphalodes rufulus* (Coleoptera: Cerambycidae). *Florida Entomologist* 89, 251–256.
- Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16, 111–120.

- Kumar, S., Strecher, G. & Tamura, K. (2016) MEGA7: molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870–1874.
- Lanfear, R., Calcott, B., Ho, S.Y.W. & Guindon, S. (2012) Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* **29**, 1695–1701.
- Leigh, J.W. & Bryant, D. (2015) PopART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6**, 1110–1116.
- López-Pantoja, G., Domínguez-Nevaldo, L. & Sánchez-Osorio, I. (2008) Mark-recapture estimates of the survival and recapture rates of *Cerambyx welensii* Küster (Coleoptera cerambycidae) in a cork oak dehesa in Huelva (Spain). *Central European Journal of Biology* **3**, 431–441.
- López-Pantoja, G., Domínguez, L. & Sánchez-Osorio, I. (2011) Analysis of *Prinobius myardi* Mulsant population dynamics in a Mediterranean cork oak stand. *Annales de la Société Entomologique de France (N.S.)* **47**, 260–268.
- López-Pantoja, G., Domínguez-Nevaldo, L. & Sánchez-Osorio, I. (2016) A procedure for calculating the thermal constants associated with the flight period of *Cerambyx welensii* Küster and *Prinobius myardi* Mulsant from field observations. *Agricultural and Forest Entomology* **18**, 82–90.
- Martín, J., Cabezas, J., Buyolo, T. & Patón, D. (2005) The relationship between *Cerambyx* spp. damage and subsequent *Biscogniauxia mediterranea* infection on *Quercus suber* forests. *Forest Ecology and Management* **216**, 166–174.
- Martínez-García, Á. (2011) Cópula mixta entre capricornios de las encinas. *Quercus* **303**, 43–44.
- Micó, E., García-López, A., Sánchez, A., Juárez, M. & Galante, E. (2015) What can physical, biotic and chemical features of a tree hollow tell us about their associated diversity? *Journal of Insect Conservation* **19**, 141–153.
- Miller, S.E. (2007) DNA barcoding and the renaissance of taxonomy. *Proceedings of the National Academy of Sciences* **104**, 4775–4776.
- Mitchell, R.F., Millar, J.G. & Hanks, L.M. (2013) Blends of (R)-3-hydroxyhexan-2-one and alkan-2-ones identified as potential pheromones produced by three species of cerambycid beetles. *Chemoecology* **23**, 121–127.
- Mitchell, R.F., Reagel, P.F., Wong, J.C.H., Meier, L.R., Silva, W.D., Mongold-Diers, J., Millar, J.G. & Hanks, L.M. (2015) Cerambycid beetle species with similar pheromones are segregated by phenology and minor pheromone components. *Journal of Chemical Ecology* **41**, 431–440.
- Montero, G., San Miguel, A. & Cañellas, I. (1998) Sistemas de silvicultura mediterránea. La dehesa. pp. 519–554 in Jiménez-Díaz, R.M. & Lamo de Espinosa, J. (Eds) *Agricultura Sostenible*. Madrid, Spain, Agrofuturo, Life and Mundi-Prensa.
- Morales-Rodríguez, C., Sánchez-González, Á., Conejo-Rodríguez, Y. & Torres-Vila, L.M. (2015) First record of *Beauveria bassiana* (Balsamo) Vuillemin (Ascomycota: Clavicipitaceae) infecting *Cerambyx welensii* (Coleoptera: Cerambycidae) and pathogenicity tests using a new bioassay method. *Biocontrol Science and Technology* **25**, 1213–1219.
- Morcuende, A. & Naveiro, F. (1993) Capturas de cerambycoides con trampa luminosa en una dehesa extremeña durante 1990, 1991 y 1992. *Phytoma-España* **48**, 53–56.
- Naveiro, F. & Morcuende, A. (1994) Observaciones sobre los cerambycoides de la dehesa quercínea en la provincia de Cáceres. *Phytoma-España* **60**, 49–52.
- Nicholls, J.A., Challis, R.J., Mutun, S. & Stone, G.N. (2012) Mitochondrial barcodes are diagnostic of shared refugia but not species in hybridizing oak gallwasps. *Molecular Ecology* **21**, 4051–4062.
- Özdkimen, H. & Turgut, S. (2009) On Turkish *Cerambyx* Linnaeus, 1758 with zoogeographical remarks (Coleoptera: Cerambycidae: Cerambycinae). *Munis Entomology and Zoology* **4**, 301–319.
- Paradis, E., Claude, J. & Strimmer, K. (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics (Oxford, England)* **20**, 289–290.
- Picard, F. (1929) *Coléoptères: Cerambycidae. Faune de France 20*. Paris, France, P. Lechevalier.
- Pons, J., Barraclough, T.G., Gómez-Zurita, J., Cardoso, A., Durán, D.P., Hazell, S., Kamoun, S., Sumlin, W.D. & Vogler, A.P. (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* **55**, 595–609.
- Porter, T.M. & Hajibabaei, M. (2018) Automated high throughput animal COI metabarcoding classification. *Scientific Reports* **8**, 1–10.
- R Core Team (2016) *R: A Language and Environment for Statistical Computing*. Vienna, Austria, R Foundation for Statistical Computing.
- Ramírez-Hernández, A., Micó, E., Marcos-García, M.A., Brustel, H. & Galante, E. (2014) The “dehesa”, a key ecosystem in maintaining the diversity of Mediterranean saproxylic insects (Coleoptera and Diptera: Syrphidae). *Biodiversity and Conservation* **23**, 2069–2087.
- Ratnasingham, S. & Hebert, P.D.N. (2007) BOLD: the Barcode of Life Data system (<http://www.barcodinglife.org>). *Molecular Ecology Resources* **7**, 355–364.
- Ratnasingham, S. & Hebert, P.D.N. (2013) A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. *PLoS ONE* **8**, e66213.
- Ratzeburg, J.T.C. (1839) *Die Forst-Insecten*. Berlin, Germany, N. Buchhandlung.
- Raupach, M.J., Astrin, J.J., Hannig, K., Peters, M.K., Stoeckle, M.Y. & Wägele, J.W. (2010) Molecular species identifications of Central European ground beetles (Coleoptera: Carabidae) using nuclear rDNA expansion segments and DNA barcodes. *Frontiers in Zoology* **7**, 26.
- Saitou, N. & Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406–425.
- Sama, G. (2002) *Atlas of the Cerambycidae of Europe and the Mediterranean Area 1*. Zlín, Czech Republic, Kabourek.
- Sama, G. (2013) *Fauna Europaea: cerambycidae*. in Audisio, P. (Ed.) *Fauna Europaea: Coleoptera 2*. Berlin, Germany, Fauna Europaea, version 2.6. Available online at <http://www.fauna-eu.org/> (accessed 27 October 2016).
- Sánchez-Osorio, I., Tapias, R., Domínguez, L. & López Pantoja, G. (2009) Variabilidad intraespecífica de la respuesta electroantagráfica en *Cerambyx welensii* Küster (Coleoptera, Cerambycidae). Influencia de factores anatómicos, fisiológicos y experimentales. *Investigación Agraria: Sistemas y Recursos Forestales* **18**, 140–151.
- Schmitt, T. (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology* **4**, 11.
- Solano, E., Thomaes, A., Cox, K., Carpaneto, G.M., Cortellessa, S., Baviera, C., Bartolozzi, L., Zilioli, M., Casiraghi, M., Audisio, P. & Antonini, P. (2016) When morphological identification meets genetic data: the case of *Lucanus cervus* and *L. tetradon* (coleoptera, Lucanidae). *Journal of Zoological Systematics and Evolutionary Research* **54**, 197–205.
- Speight, M.C.D. (1989) *Saproxylic Invertebrates and their Conservation*. Nature and Environment Series 46. Strasbourg, France, Council of Europe.

- Starzyk, J.R. & Strojny, W. (1985) The morphological variability of adults of the great capricorn beetle, *Cerambyx cerdo* L. (Coleoptera, Cerambycidae). *Polish Journal of Entomology* **55**, 491–504.
- Švácha, P. & Danilevsky, M.L. (1987) Cerambycid larvae of Europe and Soviet Union (Coleoptera, Cerambycoidea) Part I. *Acta Universitatis Carolinae, Biologica* **30**(1986), 1–176.
- Švácha, P. & Danilevsky, M.L. (1988) Cerambycid larvae of Europe and Soviet Union (Coleoptera, Cerambycoidea) Part II. *Acta Universitatis Carolinae, Biologica* **31**(1987), 121–284.
- Sweeney, J.D., Silk, P.J., Gutowski, J.M., Wu, J., Lemay, M.A., Mayo, P.D. & Magee, D.I. (2010) Effect of chirality, release rate, and host volatiles on response of *Tetropium fuscum* (F.), *Tetropium cinnamopterum* Kirby, and *Tetropium castaneum* (L.) to the aggregation pheromone, fuscumol. *Journal of Chemical Ecology* **36**, 1309–1321.
- Talavera, G., Dinca, V. & Vila, R. (2013) Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in Ecology and Evolution* **4**, 1101–1110.
- Tang, C.Q., Humphreys, A., Fontaneto, D. & Barraclough, T.G. (2014) Effects of phylogenetic reconstruction method on the robustness of species delimitation using single locus data. *Methods in Ecology and Evolution* **5**, 1086–1094.
- Templeton, A.R., Crandall, K.A. & Sing, C.F. (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**, 619–633.
- Torres-Vila, L.M. (2017) Reproductive biology of the great capricorn beetle, *Cerambyx cerdo* (Coleoptera: Cerambycidae): a protected but occasionally harmful species. *Bulletin of Entomological Research* **107**, 799–811.
- Torres-Vila, L.M., Sánchez-González, Á., Ponce-Escudero, F., Martín-Vertedor, D. & Ferrero-García, J.J. (2012) Assessing mass trapping efficiency and population density of *Cerambyx welensii* Küster by mark-recapture in dehesa open woodlands. *European Journal of Forest Research* **131**, 1103–1116.
- Torres-Vila, L.M., Sánchez-González, Á., Merino-Martínez, J., Ponce-Escudero, F., Conejo-Rodríguez, Y., Martín-Vertedor, D. & Ferrero-García, J.J. (2013) Mark-recapture of *Cerambyx welensii* in dehesa woodlands: dispersal behaviour, population density, and mass trapping efficiency with low trap densities. *Entomologia Experimentalis et Applicata* **149**, 273–281.
- Torres-Vila, L.M., Mendiola-Díaz, F.J., Conejo-Rodríguez, Y. & Sánchez-González, Á. (2016) Reproductive traits and number of matings in males and females of *Cerambyx welensii* (Coleoptera: Cerambycidae) an emergent pest of oaks. *Bulletin of Entomological Research* **106**, 292–303.
- Torres-Vila, L.M., Mendiola-Díaz, F.J. & Sánchez-González, Á. (2017a) Dispersal differences of a pest and a protected *Cerambyx* species (Coleoptera: Cerambycidae) in oak open woodlands: a mark-recapture comparative study. *Ecological Entomology* **42**, 18–32.
- Torres-Vila, L.M., Zugasti-Martínez, C., Mendiola-Díaz, F.J., De-Juan-Murillo, J.M., Sánchez-González, Á., Conejo-Rodríguez, Y., Ponce-Escudero, F. & Fernandez-Moreno, F. (2017b) Larval assemblages of large saproxylic cerambycids in Iberian oak forests: wood quality and host preference shape resource partitioning. *Population Ecology* **59**, 315–328.
- Torres-Vila, L.M., Mendiola-Díaz, F.J. & Sánchez-González, Á. (2018) Adult size and sex ratio variation of *Cerambyx welensii* (Coleoptera: Cerambycidae) in Mediterranean oak (Fagaceae) woodlands. *The Canadian Entomologist* **150**, 334–346.
- van Velzen, R., Weitschek, E., Felici, G. & Bakker, F.T. (2012) DNA barcoding of recently diverged species: relative performance of matching methods. *PLoS ONE* **7**, e30490.
- Verdugo, A. (2004) *Los Cerambycidos (Coleoptera, Cerambycidae) de Andalucía*. Córdoba, Spain, Sociedad Andaluza de Entomología.
- Villesen, P. (2007) *FaBox: an Online Fasta Sequence Toolbox*. Available online at <http://www.birc.au.dk/software/fabox> (accessed 24 April 2018).
- Villiers, A. (1946) *Coléoptères Cérambycides de l'Afrique du Nord. Faune de l'Empire Français* 5. Paris, France, ORSC.
- Villiers, A. (1978) *Faune des Coléoptères de France 1. Cerambycidae*. *Encyclopédie Entomologique* 42. Paris, France, Lechevalier.
- Vitali, F. (2001) Description des œufs des espèces françaises du genre *Cerambyx* (première contribution à l'étude des œufs des longicornes) (Coleoptera, Cerambycidae, Cerambycinae). *Les Cahiers Magellanes* **4**, 1–14.
- Vives, E. (2000) *Coleoptera Cerambycidae. Fauna Ibérica Vol. 12*. Madrid, Spain, Museo Nacional de Ciencias Naturales (CSIC).
- Whiting, M.F., Carpenter, J.C., Wheeler, Q.D. & Wheeler, W.C. (1997) The strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology* **46**, 1–68.
- Xamheu, C. (1895) Moeurs et métamorphoses d'insectes (*Cerambyx velutinus*, Brullé) [Tirage à part, pagination spéciale]. *L'Échange Revue Linnéenne* **11**, 59–61.