



**HAL**  
open science

## Implications of a phylogeographic approach for the selection of *Ceutorhynchus assimilis* as a potential biological control agent for *Lepidium draba*

Vincent Lesieur, Jean-François Martin, Hinz L Hariet, Boris B. Fumanal, Sobhian Rouhollah, Marie-Claude Bon

### ► To cite this version:

Vincent Lesieur, Jean-François Martin, Hinz L Hariet, Boris B. Fumanal, Sobhian Rouhollah, et al.. Implications of a phylogeographic approach for the selection of *Ceutorhynchus assimilis* as a potential biological control agent for *Lepidium draba*. *Biological Control*, 2018, 123, pp.43-52. 10.1016/j.biocontrol.2018.05.001 . hal-02270252

**HAL Id: hal-02270252**

**<https://institut-agro-montpellier.hal.science/hal-02270252v1>**

Submitted on 27 Aug 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Implications of a phylogeographic approach for the selection of *Ceutorhynchus assimilis* as a**  
2 **potential biological control agent for *Lepidium draba***

3

4 **Authors:**

5 Vincent Lesieur<sup>1,2\*</sup>, Jean-François Martin<sup>2</sup>, Harriet L. Hinz<sup>3</sup>, Boris Fumanal<sup>4</sup>, Rouhollah Sobhian<sup>1</sup> and  
6 Marie-Claude Bon<sup>1</sup>

7

8 **Affiliations:**

9 <sup>1</sup> USDA ARS, European Biological Control Laboratory, 810 avenue du Campus Agropolis, 34980  
10 Montferrier sur Lez, France

11 <sup>2</sup> Montpellier-SupAgro, UMR CBGP, 755 avenue du Campus Agropolis, 34980 Montferrier sur Lez,  
12 France

13 <sup>3</sup> CABI Switzerland, Rue des Grillons 1 CH-2800 Delémont, Switzerland

14 <sup>4</sup> Université Clermont Auvergne, INRA, PIAF, 63000 Clermont-Ferrand, France

15

16 \*Corresponding author. Tel.: +33 499623041; fax: +33 499623049

17 E-mail address: [vincent.lesieur@supagro.fr](mailto:vincent.lesieur@supagro.fr)

18

19 **Abstract**

20 The root-galling weevil, *Ceutorhynchus assimilis* Paykull (Coleoptera, Curculionidae), has  
21 been selected as one of the candidates to control *Lepidium draba* L. (Brassicaceae), a highly invasive  
22 weed in western North America. Originally been described as oligophagous, also attacking crop  
23 species, previous studies indicated that a specific host race to *L. draba* exists in *C. assimilis*. We  
24 therefore explored the evolutionary history and phylogeographic structure of this weevil by sampling  
25 individuals from 67 populations, spanning 12 countries and five different host plant species including  
26 *L. draba*. To explore the genetic diversity of *C. assimilis*, we analysed the *COI* gene sequences of 458  
27 individuals. Analysis revealed three distinct evolutionary lineages, one of which, the so-called  
28 *Lepidium* host race, was only found on *L. draba* and appears to occur in a restricted geographic area,

29 ranging between northern Spain and northern Italy. These results allow us to targeting sites for  
30 collection of the *Lepidium* host race and justify the prioritization of *C. assimilis* *Lepidium* host race as  
31 a candidate for the *L. draba* biological control.

32

### 33 **Keywords**

34 Invasive weed; Heart-podded hoary cress; Root-galling weevil; Host race; Genetic diversity; *COI* gene

35

### 36 **1. Introduction**

37

38 Heart-podded hoary cress, *Lepidium draba* L. ssp. *draba* [= *Cardaria draba* (L.) Desv.]  
39 (Brassicaceae), is a perennial rhizomatous weed, introduced to the US as a contaminant of seed  
40 shipments from Eurasia in the late 19th century (Gaskin et al., 2005; Hinz et al., 2012). The species  
41 has since spread throughout North America and is especially problematic in the West, where it is a  
42 declared noxious weed in 15 US states, three Canadian provinces, and also in Mexico (Rice, 2014;  
43 USDA-NRCS, 2016). *Lepidium draba* competes with native plants, decreasing local biodiversity  
44 (Francis and Warwick, 2008; Puliafico et al., 2011). Besides its ecological impact, it causes economic  
45 losses by invading several crops (McInnis et al., 2003; Mulligan and Findlay, 1974) where it also  
46 poses a dual problem: as well as being a weed, it provides an alternative food source for major crop  
47 pests, such as the cabbage seedpod weevil, *Ceutorhynchus obstrictus* (Marsham) (Doddall and Moisey,  
48 2004; Fox and Doddall, 2003). Furthermore, the weed, which is well adapted to sub-irrigated pastures  
49 and rangelands, is toxic to cattle due to alkaloid compounds, thus becoming a serious problem for  
50 livestock and forage production (Francis and Warwick, 2008; McInnis et al., 2003).

51 The invasive success of *L. draba* is related to its vegetative reproduction by an extensive  
52 rhizome system as well as its high seed production (Francis and Warwick, 2008; McInnis et al., 2003;  
53 Mulligan and Findlay, 1974). The use of herbicides is costly, not always feasible and mechanical  
54 control is not effective due to the extensive underground rhizome system of the weed (Miller et al.,  
55 1994). These limitations have motivated the exploration of alternative methods, and a classical  
56 biological control project was started in 2001 (Anonymous, 2004).

57           Among the known natural enemies of *L. draba* in its native range in Eurasia, *Ceutorhynchus*  
58 *assimilis* Paykull, 1792 [syn. *Ceutorhynchus pleurostigma* (Marsham 1802)] (Coleoptera,  
59 Curculionidae) (Colonnelli, 1993) has been prioritized as a of candidates for biological control  
60 (Fumanal et al., 2004a; 2004b; Hinz et al., 2013; Hinz et al., 2016; Virag et al., 2016). Adults lay most  
61 of their eggs in young and soft roots of the plant, causing the development of galls in which the three  
62 larval instars of the weevil develop before larvae leave the galls and pupate in the soil (Hoffman,  
63 1954). Across its Eurasian distribution, the weevil has been assumed to display a broad host range and  
64 is listed as a pest of more than 13 plant species including crops (Dennis, 1987; Hoffman, 1954;  
65 Jourdheuil, 1963). However, Fumanal et al. (2004b) proposed that *C. assimilis* is probably a complex  
66 of morphocryptic entities that differ in their host-plant spectrum within the family Brassicaceae and  
67 one of these is a specialist on *L. draba* with a distribution in the warmer climatic regions of southern  
68 Western Europe (Fumanal et al., 2004a). Whilst the host-specificity and life-cycle of this specialist  
69 entity have been studied (Fumanal et al., 2004b; Hinz et al., 2016; Virag et al., 2016), population  
70 genetic structure and the evolutionary relationships between it and the others remain poorly depicted.  
71 Investigating the population genetic structure of potential candidate agents at the onset of biological  
72 control programs would streamline host-specificity testing of differentiated genetic groups of the agent  
73 and hence can help improve the safety and success of biological control of weeds (De Biase et al.,  
74 2016; Gaskin et al., 2011; Rauth et al., 2011; Rector et al., 2010).

75           The aims of the current study were to further elucidate the genetic diversity of *C. assimilis*  
76 across its native range and between different host plants, to better characterize the distribution of the  
77 different lineages or host races and to reconstruct its evolutionary history. To do so, extensive  
78 sampling was carried out across the weevil's native range, with samples collected from different  
79 known host-plants of *C. assimilis*. Preliminary sampling (the one used in Fumanal et al., 2004a) was  
80 enriched with new populations collected in eastern and continental Europe, but also with more  
81 individuals analyzed per population, especially those associated with *L. draba*.

82           The samples were used to analyze the genetic structure of *C. assimilis* by examining the  
83 mitochondrial cytochrome *c* oxidase subunit I (*COI*) gene, a marker that has proven its usefulness and  
84 is widely used in biological control programs (Cheyppé-Buchmann et al., 2011; De Biase et al., 2016;

85 Nadel et al., 2012; Rauth et al., 2011; Tavares et al., 2015). The results were expected to provide  
86 essential information which would help to make decision regarding the suitability of the weevil as a  
87 biocontrol agent of *L. draba*, and in particular for selecting areas where the most-appropriate *L. draba*  
88 host race should be collected for further host specificity studies.

89

## 90 **2. Materials and Methods**

91

### 92 *2.1. Insect sampling*

93

94 Galls containing mature larvae of *C. assimilis* were collected in the field between 2001 and  
95 2014 from the five taxa within the Brassicaceae known to be the most commonly attacked by the  
96 weevil: *L. draba* (n = 366), *Brassica oleracea* L. (n = 52), *Brassica napus* L. (n = 7), *Diplotaxis*  
97 *erucoides* Dc. (n = 35) and *Sinapis arvensis* L. (n = 18). Plants were sampled at least 1 m apart and  
98 only one larva per gall and per plant was collected. Some of the larvae were reared to adult for  
99 morphological confirmation by Enzo Colonnelli (University of Rome, Italy). The others were stored in  
100 absolute ethanol at -20°C until DNA extraction. Sampling covered a large part of the Eurasian  
101 distribution of *C. assimilis* and collection sites were distributed across 12 countries (Table 1). By  
102 combining host-plants and geographic collection sites, the study included no less than 67 populations.  
103 Between one and ten individuals from each population were analyzed. However, for some populations  
104 of interest, up to 28 individuals were analyzed.

105

### 106 *2.2. Molecular characterization protocol*

107

108 Genomic DNA was extracted from 458 individuals using a modified cetyltrimethyl  
109 ammonium bromide (CTAB) method (Doyle and Doyle, 1987) and more recently using Qiagen  
110 DNeasy Blood & Tissue kit following manufacturer's protocol.

111 To explore intraspecific diversity, the *COI* mitochondrial gene was partially amplified  
112 (831 bp) with the primers C1-J-2183 (5'-CAACATTTATTTTGATTTTTTGG-3') and TL2-N-3014

113 (5'-TCCAATGCACTAATCTGCCATATTA-3') (Simon et al., 1994). The PCR reaction mix  
114 contained, 1X Qiagen buffer, 0.2  $\mu$ M of each dNTP, 0.3  $\mu$ M of each primer and 2 U of *Taq*  
115 polymerase and 2  $\mu$ L of genomic DNA diluted at 1/5 in a final volume of 25  $\mu$ L. PCR were performed  
116 using the following program: 94°C for 3 min, followed by 35 cycles of 92°C for 30 s, 52°C for 30 s  
117 and 62°C for 1 min and finished by a final elongation at 62°C for 7 min. The amplicons were purified  
118 and sequenced in both directions by Genoscreen (Lille, France) using an ABI PRISM 3730XL DNA  
119 sequencer. Sequences were assembled into consensus contigs with CodonCode Aligner  
120 ([www.codoncode.com](http://www.codoncode.com)) and then aligned using CLUSTAL W. In addition, sequences of closely  
121 related weevil species (*i.e.* other *Ceutorhynchus* species) were identified by a NCBI BLAST search in  
122 GenBank with one of the newly derived COI sequences of *C. assimilis*. Three *COI* sequences of two  
123 species (*C. obstrictus*, GenBank accession number: DQ058695 and JN163956; *C. gallorhenanus*,  
124 GenBank accession number: DQ058700) were included as outgroups in the phylogenetic analysis.

125

### 126 2.3. Phylogenetic analysis and divergence time estimation

127

128 Relationships among *C. assimilis* from different host plants and geographical regions were  
129 reconstructed using both maximum likelihood and Bayesian inferences. First, jModelTest2 (Darriba et  
130 al., 2012) was used to test for the best-fit model of sequence evolution. The best-fit model was  
131 selected using the corrected Akaike Information Criterion (AICc) (Darriba et al., 2012). Maximum  
132 likelihood analyses were conducted with PhyML v3.0 (Guindon et al., 2010). Nodal support was  
133 assessed using non-parametric bootstrapping (1,000 replicates). Bayesian phylogeny was  
134 reconstructed using the software BEAST v1.8.0 (Drummond et al., 2012). We also used this software  
135 to estimate divergence time of lineages characterized within *C. assimilis*. In the absence of geological  
136 and/or fossil calibration points to estimate divergence times, we dated the clock-like *COI* phylogeny  
137 by applying a global mutation rate of 3.54% Myr<sup>-1</sup>, a substitution rate calibrated for tenebrionid beetles  
138 (Papadopoulou et al., 2010). We used an uncorrelated lognormal (UCLN) relaxed clock under  
139 exponential growth population models. The analyses were performed using two independent runs of  
140 20 million generations and trees were sampled every 1,000 iterations. Input files were generated with

141 BEAUTi (Drummond et al., 2012) v1.8.0. After checking for the effective sampling sizes (ESS) of  
142 parameters (a threshold of 200 was used) with Tracer v1.6 (available from  
143 <http://tree.bio.ed.ac.uk/software/tracer>), the first 25% of each run was discarded as burn-in phase for  
144 the estimation of the consensus topology and the computation of the posterior probability (PP) for  
145 each node. The intraspecific relationships might be better visualized with networks, allowing a more  
146 detailed display of population information than on evolutionary trees (Posada and Crandall, 2001)  
147 although both may be informative. A haplotype network was constructed in PopART (Leigh and  
148 Bryant, 2015) using TCS network. To visualize the spatial distribution of the different entities (*e.g.*  
149 *Lepidium* host race *versus* the other lineages), the genetic diversity of the populations was plotted as  
150 pie charts on a map.

151 To estimate the number of molecular operational taxonomic units (MOTUs) and to explore  
152 molecular species delimitation, two different methods were used. First, we used Automatic Barcode  
153 Gap Discovery (ABGD) developed by Puillandre et al. (2012). This tool uses pairwise distance to  
154 detect a barcode gap, this barcode gap is used as a threshold to separate the sequences into putative  
155 species (*i.e.* genetically homogeneous groups). The procedure is recursively applied to the previously  
156 obtained groups until no more partitioning can be done. The analysis was conducted through  
157 uploading a sequence alignment on the ABGD web server at  
158 <http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html>. We computed a matrix of pairwise distances  
159 using the K2P model (Kimura, 1980) and ABGD was run with the default settings (Pmin = 0.001,  
160 Pmax = 0.1, Steps = 10, X (relative gap width) = 1.5, Nb bins = 20). The Poisson tree processes (PTP)  
161 model for species delimitation (Zhang et al., 2013) was also used. This model uses coalescence theory  
162 and estimates the speciation rate directly from the number of substitutions. It assumes that each  
163 substitution has a small probability of generating a speciation event. Consequently, the number of  
164 substitutions between species is expected to be significantly higher than within species. The ML  
165 phylogeny obtained with PhyML was used as input. The analysis was run on a web server for PTP  
166 (available at <http://species.h-its.org/ptp/>) with 200,000 MCMC generations, a thinning value of 100  
167 and a burn-in of 25%. As recommended by the developers, the convergence of the MCMC chain was  
168 confirmed visually.

169 To estimate divergence between the main genetic entities defined by the resulting  
170 phylogenetic tree, we used Mega version 7.0 (Kumar et al., 2016). We calculated the pairwise genetic  
171  $p$ -distances among the phylogenetic entities. To get a rough overview of the historical demography for  
172 each lineage, the frequency-based indicators of a population expansion (or selection in non-neutral  
173 markers) Fu's  $F_s$  (Fu, 1997) were calculated with Arlequin 3.5 (Excoffier and Lischer, 2010). Fu's  
174 statistic is considered highly sensitive to detect population expansion (Fu, 1997; Ramos-Onsins and  
175 Rozas, 2002). The values of Fu's  $F_s$  are expected to have significantly negative values for population  
176 expansion.

177 The estimation of gene diversity ( $H_d$ ) and nucleotide diversity ( $\pi$ ) for each of the sampling  
178 locality was conducted with Arlequin 3.5 (Excoffier and Lischer, 2010) and allelic richness ( $r$ ) was  
179 computed using the rarefaction method proposed by Petit et al. (1998) using Contrib  
180 (<http://www.pierroton.inra.fr/genetics/labo/Software/Contrib/>).

181 Finally, to analyze the genetic structure among populations and to identify potentially distinct  
182 units within the restricted area where the weevil *Lepidium* host race was found (see Results), we used  
183 a spatial analysis of molecular variance (SAMOVA) with SAMOVA 1.0 (Dupanloup et al., 2002).  
184 This approach was thus only used for individuals collected on *L. draba*. It defines groups of  
185 populations that are geographically homogeneous and maximally differentiated. The program was run  
186 for two to 10 differentiated groups ( $K = 2$  to  $K = 10$ ) using 10,000 permutations from 100 random  
187 initial conditions.

188

### 189 **3. Results**

190

191 The 458 sequenced individuals from 67 populations yielded a final alignment of 774 bp long  
192 uncovering a total of 66 haplotypes. All haplotype sequences were deposited in GenBank (accession  
193 numbers: KY486382-KY486447). From the final alignment, a total of 59 (7.6%) polymorphic  
194 nucleotide sites, of which 46 were parsimony-informative. Total gene diversity was  $0.929 \pm 0.006$  and  
195 nucleotide diversity was  $0.013 \pm 0.003$ .



196 The most appropriate model of *COI* sequence evolution was TN93+I+G, and this model was  
197 employed for further analyses. Phylogenetic reconstructions using both ML and BI revealed three  
198 divergent genetic entities within *C. assimilis* (Fig. 1). The first of them, referred to hereafter as the  
199 ‘*Lepidium* host race’, was defined with strong support (BS = 94 and PP = 1.00). This evolutionary  
200 lineage, composed of 12 haplotypes, was represented by weevils sampled only on *L. draba* ssp. *draba*  
201 within a restricted area comprising northern Spain, southern France and the northwestern part of Italy  
202 (Figs. 2 and 3). The second evolutionary lineage, called Lineage 2 (support values: BS = 83 and  
203 PP = 0.99), was composed of seven haplotypes (Fig. 1). These haplotypes were recovered in southern  
204 France, northern Italy and Greece from individuals collected on three different host plants: *D.*  
205 *erucoides*, *S. arvensis* and *B. oleracea* (Figs. 2 and 3). This lineage was never found on *L. draba* in our  
206 study. The remaining haplotypes clustered in Lineage 1 (Fig. 1). This evolutionary lineage was widely  
207 distributed across the entire sampled area and from the whole spectrum of host plants considered in  
208 this study (Fig. 3). On internal nodes, phylogeny is not fully resolved with low support values or  
209 discrepancy between ML and BI (Fig. 1).

210 As observed on the phylogenetic tree, the haplotype network also revealed a clear split  
211 between the *Lepidium* host race and the two other lineages (Fig. 3). The haplotype network showed  
212 that the *Lepidium* host race and Lineage 1 differed from each other by at least 14 mutation steps  
213 (Fig. 2).

214 The genetic divergence between sequences obtained from the *Lepidium* host race and the two  
215 other lineages were 2.41% and 2.21%, respectively for Lineage 1 and Lineage 2 (Table 2). This is  
216 much lower than divergence values obtained with the *Ceutorhynchus* species used as outgroups (11.38  
217 – 12.95%). From the ABGD analysis, the partitions 1-3 (P = 0.001000-0.002783) matched the  
218 phylogenetic tree and partitioned the dataset into three molecular operational taxonomic units  
219 (MOTUs) corresponding to Lineage 1, Lineage 2 and *Lepidium* host race. However, ABGD results did  
220 not fully support a marked barcoding gap (Supplementary material Fig. 1). The partition 4  
221 (P = 0.004642) only resulted in two MOTUs while the partition 5 (P = 0.007743) suggested that *C.*  
222 *assimilis* was a single species. The PTP analysis resulted in partitioning *C. assimilis* into two MOTUs,  
223 the *Lepidium* host race on the one hand and the combination of the Lineages 1 and 2 on the other hand.

224 This analysis suggested that these MOTUs may represent potential cryptic species or subspecies  
225 within *C. assimilis*.

226 Estimation of time divergence was made for the two major splits within the *C. assimilis* tree  
227 (Fig. 1). Divergence between the *Lepidium* host race and the other lineages of *C. assimilis* (labeled  
228 with Roman numeral I on Fig. 1) was estimated to have occurred 1.17 Mya (95% HPD: 0.70–1.76),  
229 this estimation dates the split between the lower and the mid-Pleistocene. The Lineage 1 and  
230 Lineage 2 (referred on Fig. 1 as the Roman numeral II), diverged from each other around 0.87 Mya  
231 (95% HPD: 0.50–1.26). Demographic analyses suggest a population expansion in the three lineages  
232 (Table 2), although for the Lineage 2, negative value of Fu's  $F_s$  was not significant.

233 Most populations within sampling localities had one to three haplotypes. Haplotype H 21,  
234 belonging to Lineage 1, was the most common and the most widespread haplotype. It was found in 83  
235 individuals from 17 sampling sites, distributed from France to Armenia. In the restricted area where  
236 the *Lepidium* host race was found, the most geographically widespread haplotype corresponded to H 1,  
237 and was found in 10 sampling sites ( $n = 35$  individuals), and only on *L. draba*. Both genetic diversity  
238 and nucleotide diversity varied considerably from one sampling site to another in the restricted area of  
239 the *Lepidium* host race as well as in the other parts of Europe (Table 1).

240 Within the restricted area where the *Lepidium* host race lineage was found, and taking into  
241 account only the weevils collected on *L. draba*, the SAMOVA analysis suggested the presence of five  
242 clusters in the dataset. This configuration was retained because the largest increase in  $F_{CT}$  values was  
243 observed when the geographic sampling area was partitioned into five clusters (Supplementary  
244 material Fig. 2), although the  $F_{CT}$  values still slowly increased after  $K = 5$ , probably a response to the  
245 continuous decrease in  $F_{SC}$  until all sampling sites are separated (Dupanloup et al., 2002). Moreover,  
246 for  $K > 5$ , the newly formed groups only contained single populations, which indicates that the group  
247 structure disappeared (Magri et al., 2006). The five groups were geographically consistent, except  
248 group 4 that showed a disjunctive distribution, occurring mainly in France and Northern Spain but also  
249 in Northern Italy (Fig. 4). The population A84, located in Central Spain was mostly composed of  
250 individuals harboring the haplotype H55 that is only found in this sampling site ( $n = 26$ ). This  
251 population was segregated from the other Spanish populations and considered as a distinct group

252 (group 5). Group 1, located in the northern continental range, clustered the sites A83 and A69. Those  
253 two sites are predominantly composed of individuals from the generalist Lineage 1 (Fig. 3), 14  
254 individuals among 20 for site A83 and 12 individuals among 14 for site A69, respectively.

255

## 256 **4. Discussion**

257

### 258 *4.1. Evolutionary history of C. assimilis*

259

260 The results presented here, confirm the proposal (Fumanal et al., 2005; Fumanal et al., 2004b)  
261 that the weevil populations developing on *L. draba* in a restricted area of the western Mediterranean  
262 region should be considered as a *Lepidium* host race. Present results met most of the criteria proposed  
263 by Drès and Mallet (2002) in their definition of host race including genetically differentiated  
264 sympatric populations living on different hosts. Indeed, the lineage referred to here as the *Lepidium*  
265 host race represents a divergent evolutionary lineage strictly associated with *L. draba*. The genetic  
266 divergence between the host race and the other lineages of *C. assimilis* (2.21% with Lineage 1 and  
267 2.41% with Lineage 2) corresponds to values lower than those usually observed between  
268 *Ceutorhynchus* species (Laffin et al., 2005a) as confirmed by the divergence observed with the  
269 congeneric species used as outgroups. However, the level of divergence detected here is in the same  
270 order of magnitude as the one observed between two sister species, *C. erysimi* and *C. contractus*  
271 (Stepanović et al., 2015). Likewise, although the ABGD procedure tends to consider the three different  
272 lineages as a single species, the PTP analysis suggests that the *Lepidium* host race may potentially  
273 represent cryptic species or subspecies. The potential for hybridization between the different genetic  
274 entities has never been investigated in the field, and only partly in lab conditions (Fumanal et al.  
275 2004a).

276 Preliminary crossing experiments between individuals of *C. assimilis* from the *Lepidium* host  
277 race and individuals reared from *L. draba* but belonging to the generalist Lineage 1 showed an  
278 absence of prezygotic barriers but suggested partial postzygotic barriers. Indeed, fertilization seemed  
279 possible and egg laying from females was observed in some cases but it was not possible to test for

280 fertility of the resulting outcrossed progenies as they died before they reproduced. Crossing of  
281 individuals belonging to the *Lepidium* host race with individuals developing on other host plants  
282 (Lineage 2) resulted only in abnormal small galls and abortion of neonate larvae (Fumanal et al.,  
283 2004a). Taken as a whole, the available data (*i.e.* ecological, biological and genetic data) point to a  
284 limited gene flow between the *Lepidium* host race and the two other lineages. However, the extent of  
285 hybridization and gene flow between the different entities remain to be found under natural conditions.  
286 This could be tested in the Rhone Valley, where the *Lepidium* host race and more generalist entities  
287 co-occur on *L. draba*. The recent development of microsatellite markers for *C. assimilis* (Lesieur et al.,  
288 2016) will enable investigation of the gene flow between the different entities identified here.

289         The split between the *Lepidium* host race and the two other lineages was estimated to occur  
290 during the Pleistocene and, diversification within each lineage also started during this period. The  
291 Pleistocene and its climatic cycles, with the succession of glacial and interglacial periods, have had a  
292 great impact on species evolution and have shaped the current distribution of numerous organisms  
293 (Hewitt, 2000; Médail and Diadema, 2009; Schmitt, 2007). The western Mediterranean distribution of  
294 the *Lepidium* host race suggests one (or more) refugia in Spain or the south of France, a region  
295 identified as glacial refugium for several other species (Médail and Diadema, 2009).

296         Over its wide distribution, no clear geographic structure within Lineage 1 was found in the  
297 analysis. However, several glacial refugia for *C. assimilis* and further expansion of both lineages are  
298 suggested. In particular, neutrality tests support the hypothesis of a recent population expansion. The  
299 current distributions of Lineage 1 and the *Lepidium* host race reveal a secondary contact of those  
300 lineages in the Rhone Valley. Except for the sampling site A83 located in the Massif Central, where  
301 the *Lepidium* host race is not the main genetic lineage, the distribution of the *Lepidium* host race in  
302 France matches the Mediterranean biogeographical area (Condé et al., 2002). This would suggest that  
303 past but also present climatic conditions may shape the distribution of this lineage. The geographic  
304 distribution of the Lineage 2 (*i.e.* Greece, Southern France and also three individuals in Italy) is also  
305 consistent with southern refugia (Médail and Diadema, 2009). This lineage, only found on  
306 *D. eruroides*, *S. arvensis* and *B. oleracea* was never associated with *L. draba*. It may be more  
307 widespread than illustrated by our study, but this lack of representation may be linked with the lower

308 sampling effort on these plants at the global scale. Moreover, the moderate support values for  
309 Lineages 1 and 2 need further considerations for better resolving the evolutionary history of these two  
310 lineages.

311         Within the area where the *Lepidium* host race is present, Médail and Diadema (2009) defined  
312 several ‘phylogeographical hotspots’ favorable to the evolutionary divergence of different plant  
313 species. In line with this, the weevil’s genetic divergence correlates with the differentiation of the  
314 European *L. draba* ssp *draba*. Fumanal et al. (2004a) found two different plant genetic lineages within  
315 the European *L. draba*, one restricted to Southern Europe and one widely distributed in Europe. As for  
316 *C. assimilis* in which the lineages were morphologically similar, a comparison of several plant traits  
317 (such as shoot density, individual shoot height) of European populations of *L. draba* from several  
318 regions did not reveal any differences (Hinz et al., 2012). However, the authors did not include in their  
319 study any population of the weed from the area where the *C. assimilis* *Lepidium* host race occurs. In  
320 our study, the divergence found between the *Lepidium* host race and the other genetic entities of *C.*  
321 *assimilis* can also be linked with biological features of weevil populations. Differences in phenology,  
322 sex ratio, cold hardiness and overwintering survival have been observed between the different entities  
323 (Fumanal et al., 2004b; Virag et al., 2016). For instance, Fumanal et al. (2004b) observed in the area  
324 where the *Lepidium* host race is found that sympatric weevil populations on *L. draba* developed faster  
325 and pupation started up to two months earlier than individuals on *S. arvensis*. Phenological differences  
326 are an important factor in sympatric speciation or, at least, host plant or habitat specialization (Drès  
327 and Mallet, 2002; Roy et al., 2016; Santos et al., 2007). These biological differences could partly  
328 explain the genetic divergence between sympatric populations living on different host- plants as  
329 observed in the two sites in southern France (*i.e.* Q-VV-VVS and A49-L-B (Table 1)) where the  
330 weevils were collected synchronously.

331         Despite numerous taxonomic revisions of the genus (Colonnelli, 1993; Colonnelli, 2004;  
332 Hoffman, 1954; Tempère and Péricart, 1989), taxonomical and nomenclatural issues represent the  
333 main constraints in studying this important weevil group, as highlighted by Stepanović et al. (2015). In  
334 this context, although it was not the main aim of this study, we confirmed the distinction between the  
335 cabbage seedpod weevil, *C. obstrictus* (Marsham) [= *C. assimilis* (Paykull)], a serious European pest

336 of canola or oil seed rape, which is invasive in North America (Dosdall and Mason, 2010; Laffin et al.,  
337 2005b), and *C. assimilis* [= *C. pleurostigma*]. Colonnelli (1993) determined that *C. obstrictus* is the  
338 correct name for the cabbage seedpod weevil, however confusion still exists in some publications  
339 where *C. obstrictus* is still cited as *C. assimilis*. Moreover, *C. assimilis* has been listed as a pest of  
340 several cultivated crops such as cabbage or rapeseed (Hoffman, 1954; Jourdheuil, 1963).

341

#### 342 4.2. Implications for biological control

343

344 The present distinction between the specialist (*i.e.* the *Lepidium* host race) and the other  
345 lineages has serious implications from a biological control perspective. Indeed, the *Lepidium* host race  
346 should be considered as beneficial and a potential biological control agent for *L. draba* whereas the  
347 others, particularly those developing on economic crops, should still be regarded as pests. First, the  
348 results allow targeting precise sites for collection of the *Lepidium* host race. We showed that the  
349 *Lepidium* host race of *C. assimilis* is restricted to a relatively small geographic area. On the 122  
350 individuals collected on *L. draba* within a triangular area including Valls in Cataluna (Spain), San  
351 Remo in Liguria (Italy) and Montélimar in the Rhone Valley (France), 116 individuals (ca. 95%)  
352 corresponded to the *Lepidium* host race.

353 Secondly, within the restricted distribution area where the *Lepidium* host race occurs in  
354 sympatry with the other lineages, Lineage 1 was observed on *L. draba* at a low frequency and Lineage  
355 2 has never been found on *L. draba* but on other Brassicaceae. Moreover, when sampling was carried  
356 out synchronously on several host plant species, the *Lepidium* host race has never been found on  
357 plants other than *L. draba*, thus reducing the risk of collecting generalist individuals. Indeed, these  
358 other plant species do not constitute acceptable hosts for the *Lepidium* host race (Fumanal et al.,  
359 2004b). Thus, when restricting sampling to sites/areas where the *Lepidium* host race predominates and  
360 only collecting galls/adults from *L. draba*, one can be close to certain to only collect individuals of the  
361 specialist *Lepidium* host race. Nevertheless, as previously mentioned, the different entities are  
362 morphologically identical, discriminating the *Lepidium* host race from the other lineages remains an  
363 issue. However, Fumanal et al. (2005) developed a non-invasive technique based on direct-PCR of

364 fecal DNA combined with double strand conformation polymorphism (DSCP) typing. This method  
365 allows to quickly discriminating the *Lepidium* host race from the other entities without killing the  
366 insects and could be adapted for systematic screenings with current metabarcoding methods.

367 Thirdly, based on our results, the split between the *Lepidium* host race and the other more  
368 generalist lineages is relatively old (1.17 Mya, Pleistocene) suggesting that the *Lepidium* host race  
369 belongs to a long-standing *L. draba*-associated entity rather than a lineage prone to a recent host shift.  
370 As such it should remain specific to *L. draba*, which justifies its prioritization as a candidate for  
371 biological control of *L. draba*.

372 Even if the results presented here justify the prioritization of the *Lepidium* host race as a  
373 candidate for the *L. draba* biological control, studies on the biology and the specificity of the host race  
374 must be undertaken to evaluate its potential host spectrum in North America. Host specificity tests,  
375 including North American endemic plant species, are currently conducted by CABI at Délemont  
376 (Switzerland).

377 Genetic diversity within the *Lepidium* host race was highly variable from one sites to another,  
378 and in some sampling sites the genetic diversity was very low (*e.g.* sites A55; A69; A72 (Table 1)). At  
379 the sampling site scale, only one to three haplotypes were observed, even when the sampling size was  
380 large ( $n > 25$ ). Collecting individuals from several nearby sites should ensure maximizing genetic  
381 diversity. Preserving this diversity in released populations is of importance as genetic variation in  
382 ecologically important traits is hypothesized to be important in successful establishment, allowing  
383 adaptation to the novel environment (Dlugosch et al., 2015; Szucs et al., 2014).

384 *Lepidium draba* is invasive in a large part of North America, including areas with a wide  
385 spectrum of climatic conditions (Rice, 2014; USDA-NRCS, 2016). A climate match model suggests  
386 that the *Lepidium* host race may establish and persist in some areas where the weed is problematic,  
387 while establishment in more continental climates could be possible, but more difficult (Virag et al.,  
388 2016). However, relatively rapid adaptation after release may occur as observed for the flea beetle  
389 *Longitarsus jacobaeae* (Waterhouse) (Szucs et al., 2012). Preserving the genetic variation of the  
390 *Lepidium* host race in released populations may increase their adaptive potential in the new

391 environment (Phillips et al., 2008), and therefore also potential adaptation to more continental climate,  
392 such as eastern Colorado, Wyoming and central Montana.

393 The *L. draba* populations that invaded North America displayed genetic differentiation and  
394 originated from multiple Eurasian origins including the southern European region where the weevil  
395 host race occurs (Gaskin et al., 2005). Likewise, within its restricted distribution area, the genetic  
396 diversity of the *Lepidium* host race appears as different units as highlighted by the SAMOVA.  
397 Different populations of potential biological control agents may have different abilities to develop and  
398 damage different introduced genotypes of the target weed, as it was observed for the eriophyid mite,  
399 *Floracarus perrepae* Knihinicki and Boczek, to control the fern, *Lygodium microphyllum* (Cav.) R.  
400 Br. in Florida (Goolsby et al., 2006). Because of the common pattern of differentiation between the  
401 *Lepidium* host race of *C. assimilis* and *L. draba*, future investigations should evaluate the impact of the  
402 different units of the *Lepidium* host race on different populations of *L. draba*, including North  
403 American populations.

404

#### 405 **Acknowledgements**

406

407 We thank the Director of EBCL, Lincoln Smith, for reviewing the manuscript and providing  
408 useful comments. We thank Camille Rouvière for her help in extracting DNA and sequencing. We  
409 also thank Fatiha Guermache and Mélanie Jeanneau (USDA-ARS-EBCL) for technical assistance and  
410 field collections. We are also indebted to many collaborators who provided samples. Enzo Colonnelli  
411 (University of Roma, Italy) is gratefully acknowledged for identification of the weevils. We would  
412 also like to thank Dr. G. Kergoat (INRA, UMR CBGP) for helpful discussions on data analyses and  
413 interpretation. Mention of any trade names or commercial products in this publication is done solely  
414 for the purpose of providing specific information, and does not imply recommendation or endorsement  
415 by the USDA; the USDA is an equal opportunity provider and employer. This research did not receive  
416 any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

417

#### 418 **References**



419

420 Anonymous, 2004. Whitetop / Hoary Cress. In: Coombs, E.M., Clark, J.K., Piper, G.L., Cofrancesco  
421 Jr., A.F. (Eds.), Biological Control of Invasive Plants in the United States. Oregon State  
422 University Press. Oregon State University Press, pp. 467.

423 Cheyppé-Buchmann, S., Bon, M.C., Warot, S., Jones, W., Malausa, T., Fauvergue, X., Ris, N., 2011.  
424 Molecular characterization of *Psytalia lounsburyi*, a candidate biocontrol agent of the olive  
425 fruit fly, and its *Wolbachia* symbionts as a pre-requisite for future intraspecific hybridization.  
426 *Biocontrol* 56, 713-724.

427 Colonnelli, E., 1993. The Ceutorhynchinae types of I.C. Fabricius and G. von Paykull (Coleoptera:  
428 Curculionidae). *Koleopterologische Rundschau* 63, 299-310.

429 Colonnelli, E., 2004. Catalogue of Ceutorhynchinae of the world, with a key to genera. (Insecta:  
430 Coleoptera: Curculionidae).

431 Condé, S., Richard, D., Liamine, N., Leclère, A.-S., Framstad, E., Erikstad, K.E., Yoccoz, N., 2002.  
432 The Mediterranean biogeographical region. European Environment Agency, Sweden.

433 Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics  
434 and parallel computing. *Nature methods* 9, 772-772.

435 De Biase, A., Colonnelli, E., Belvedere, S., La Marca, A., Cristofaro, M., Smith, L., 2016. Genetic and  
436 morphological studies of *Trichosirocalus* species introduced to North America, Australia and  
437 New Zealand for the biological control of thistles. *Bulletin of entomological research* 106, 99-  
438 113.

439 Dennis, S., 1987. Agricultural insect pest of temperate regions and their control. Cambridge University  
440 Press, London, England.

441 Dlugosch, K.M., Anderson, S.R., Braasch, J., Cang, F.A., Gillette, H.D., 2015. The devil is in the  
442 details: genetic variation in introduced populations and its contributions to invasion.  
443 *Molecular Ecology* 24, 2095-2111.

444 Dosdall, L.M., Mason, P.G., 2010. Key pests and parasitoids of oilseed rape or canola in North  
445 America and the importance of parasitoids in integrated management. In: Williams, I.H.,  
446 (Ed.), *Biocontrol-based integrated management of oilseed rape pests*. Springer, Dordrecht, The  
447 Netherlands, pp. 167-213.

448 Dosdall, L.M., Moisey, D.W., 2004. Developmental biology of the cabbage seedpod weevil,  
449 *Ceutorhynchus obstrictus* (Coleoptera: Curculionidae), in spring canola, *Brassica napus*, in  
450 western Canada. *Annals of the Entomological Society of America* 97, 458-465.

451 Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf  
452 tissue. *Phytochemistry Bulletin* 19, 11-15.

453 Drès, M., Mallet, J., 2002. Host races in plant-feeding insects and their importance in sympatric  
454 speciation. *Philosophical Transactions of the Royal Society B-Biological Sciences* 357, 471-  
455 492.

456 Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti  
457 and the BEAST 1.7. *Molecular biology and evolution* 29, 1969-1973.

458 Dupanloup, I., Schneider, S., Excoffier, L., 2002. A simulated annealing approach to define the  
459 genetic structure of populations. *Molecular Ecology* 11, 2571-2581.

460 Excoffier, L., Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform  
461 population genetics analyses under Linux and Windows. *Molecular ecology resources* 10,  
462 564-567.

463 Fox, A.S., Dosdall, L.M., 2003. Reproductive biology of *Ceutorhynchus obstrictus* (Coleoptera:  
464 Curculionidae) on wild and cultivated Brassicaceae in southern Alberta. *Journal of*  
465 *Entomological Science* 38, 533-544.

466 Francis, A., Warwick, S., 2008. The biology of Canadian weeds. 3. *Lepidium draba* L., *L. chalepense*  
467 L., *L. appelianum* Al-Shehbaz (updated). *Canadian Journal of Plant Science* 88, 379-401.

468 Fu, Y.-X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and  
469 background selection. *Genetics* 147, 915-925.

470 Fumanal, B., Martin, J.F., Sobhian, R., Gaskin, J., Bon, M., 2004a. Evolutionary biology as a tool  
471 towards a more customized biological control strategy of weeds: *Lepidium draba* as a case  
472 study. AFPP, XII Colloque international sur la Biologie des mauvaises herbes, Dijon, France,  
473 pp. 421-426.

- 474 Fumanal, B., Martin, J.F., Bon, M.C., 2005. High through-put characterization of insect morphocryptic  
475 entities by a non-invasive method using direct-PCR of fecal DNA. *Journal of Biotechnology*  
476 119, 15-19.
- 477 Fumanal, B., Martin, J.F., Sobhian, R., Blanchet, A., Bon, M.C., 2004b. Host range of *Ceutorhynchus*  
478 *assimilis* (Coleoptera : Curculionidae), a candidate for biological control of *Lepidium draba*  
479 (Brassicaceae) in the USA. *Biological Control* 30, 598-607.
- 480 Gaskin, J.F., Bon, M.-C., Cock, M.J., Cristofaro, M., De Biase, A., De Clerck-Floate, R., Ellison,  
481 C.A., Hinz, H.L., Hufbauer, R.A., Julien, M.H., 2011. Applying molecular-based approaches  
482 to classical biological control of weeds. *Biological Control* 58, 1-21.
- 483 Gaskin, J.F., Zhang, D.Y., Bon, M.C., 2005. Invasion of *Lepidium draba* (Brassicaceae) in the western  
484 United States: distributions and origins of chloroplast DNA haplotypes. *Molecular Ecology*  
485 14, 2331-2341.
- 486 Goolsby, J.A., De Barro, P.J., Makinson, J.R., Pemberton, R.W., Hartley, D.M., Frohlich, D.R., 2006.  
487 Matching the origin of an invasive weed for selection of a herbivore haplotype for a biological  
488 control programme. *Molecular Ecology* 15, 287-297.
- 489 Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New  
490 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the  
491 performance of PhyML 3.0. *Systematic biology* 59, 307-321.
- 492 Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405, 907-913.
- 493 Hinz, H., Blair, L., Closca, C., Diaconu, A., 2013. Biological control of whitetops, *Lepidium draba*, *L.*  
494 *chalepense* and *L. appelianum*. Annual Report 2012. CABI Switzerland.
- 495 Hinz, H.L., Cloșca, C., Castellan, I., Heijs, W., Morelon, S., Stahlke, A., 2016. Biological control of  
496 whitetops, *Lepidium draba*, *L. chalepense* and *L. appelianum*. Annual Report 2015. CABI  
497 Switzerland.
- 498 Hinz, H.L., Schwarzländer, M., McKenney, J.L., Cripps, M.G., Harmon, B., Price, W.J., 2012.  
499 Biogeographical comparison of the invasive *Lepidium draba* in its native, expanded and  
500 introduced ranges. *Biological Invasions* 14, 1999-2016.
- 501 Hoffman, A., 1954. Faune de France. Office Central de Faune publ., Paris, France.
- 502 Jourdheuil, P., 1963. *Ceutorhynchus pleurostigma* Marsham. In: Balachowsky, A.S., (Ed.),  
503 Entomologie appliquée à l'agriculture (Coléoptères). Masson et Cie.
- 504 Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through  
505 comparative studies of nucleotide sequences. *Journal of molecular evolution* 16, 111-120.
- 506 Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis  
507 version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33, 1870-1874.
- 508 Laffin, R., Dossdall, L., Sperling, F., 2005a. Population structure and phylogenetic relationships of  
509 *Ceutorhynchus neglectus* (Coleoptera: Curculionidae). *The Canadian Entomologist* 137, 672-  
510 684.
- 511 Laffin, R.D., Dossdall, L.M., Sperling, F.A.H., 2005b. Population structure of the cabbage seedpod  
512 weevil, *Ceutorhynchus obstrictus* (Marsham) (Coleoptera Curculionidae): Origins of North  
513 American introductions. *Environmental Entomology* 34, 504-510.
- 514 Leigh, J.W., Bryant, D., 2015. POPART: full-feature software for haplotype network construction.  
515 *Methods in Ecology and Evolution* 6, 1110-1116.
- 516 Lesieur, V., Jeanneau, M., Martin, J., Bon, M., 2016. Development and characterization of 11  
517 microsatellite markers in the root-gall-forming weevil, *Ceutorhynchus assimilis* (Coleoptera:  
518 Curculionidae). *Applied Entomology and Zoology*, 1-6.
- 519 Magri, D., Vendramin, G.G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., Latałowa, M., Litt,  
520 T., Paule, L., Roure, J.M., 2006. A new scenario for the Quaternary history of European beech  
521 populations: palaeobotanical evidence and genetic consequences. *New Phytologist* 171, 199-  
522 221.
- 523 McInnis, M.L., Kiemnec, G.L., Larson, L.L., Carr, J., Sharratt, D., 2003. Heart-podded hoary cress.  
524 *Rangelands Archives* 25, 18-23.
- 525 Médail, F., Diadema, K., 2009. Glacial refugia influence plant diversity patterns in the Mediterranean  
526 Basin. *Journal of Biogeography* 36, 1333-1345.
- 527 Miller, R.F., Svejcar, T.J., Rose, J.A., McInnis, M.L., 1994. Plant development, water relations, and  
528 carbon allocation of heart-podded hoary cress. *Agronomy Journal* 86, 487-491.

529 Mulligan, G.A., Findlay, J.N., 1974. The biology of Canadian weeds. 3. *Cardaria draba*, *C.*  
530 *chalepensis*, and *C. pubescens*. Canadian Journal of Plant Science 54, 149-160.

531 Nadel, R.L., Wingfield, M.J., Scholes, M.C., Lawson, S.A., Noack, A.E., Neser, S., Slippers, B., 2012.  
532 Mitochondrial DNA diversity of *Cleruchoides noackae* (Hymenoptera: Mymaridae): a  
533 potential biological control agent for *Thaumastocoris peregrinus* (Hemiptera:  
534 Thaumastocoridae). Biocontrol 57, 397-404.

535 Papadopoulou, A., Anastasiou, I., Vogler, A.P., 2010. Revisiting the insect mitochondrial molecular  
536 clock: the mid-Aegean trench calibration. Molecular Biology and Evolution 27, 1659-1672.

537 Petit, R.J., El Mousadik, A., Pons, O., 1998. Identifying populations for conservation on the basis of  
538 genetic markers. Conservation biology 12, 844-855.

539 Phillips, C., Baird, D., Iline, I., McNeill, M., Proffitt, J., Goldson, S., Kean, J., 2008. East meets west:  
540 adaptive evolution of an insect introduced for biological control. Journal of Applied Ecology  
541 45, 948-956.

542 Posada, D., Crandall, K.A., 2001. Intraspecific gene genealogies: trees grafting into networks. Trends  
543 in Ecology & Evolution 16, 37-45.

544 Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, Automatic Barcode Gap  
545 Discovery for primary species delimitation. Molecular ecology 21, 1864-1877.

546 Puliafico, K.P., Schwarzländer, M., Price, W.J., Harmon, B.L., Hinz, H.L., 2011. Native and exotic  
547 grass competition with invasive hoary cress (*Cardaria draba*). Invasive Plant Science and  
548 Management 4, 38-49.

549 Ramos-Onsins, S.E., Rozas, J., 2002. Statistical properties of new neutrality tests against population  
550 growth. Molecular biology and evolution 19, 2092-2100.

551 Rauth, S.J., Hinz, H.L., Gerber, E., Hufbauer, R.A., 2011. The benefits of pre-release population  
552 genetics: a case study using *Ceutorhynchus scrobicollis*, a candidate agent of garlic mustard,  
553 *Alliaria petiolata*. Biological Control 56, 67-75.

554 Rector, B.G., De Biase, A., Cristofaro, M., Primerano, S., Belvedere, S., Antonini, G., Sobhian, R.,  
555 2010. DNA fingerprinting to improve data collection efficiency and yield in an open-field  
556 host-specificity test of a weed biological control candidate. Invasive Plant Science and  
557 Management 3, 429-439.

558 Rice, P., 2014. Invaders database system. <http://invader.dbs.umt.edu/>

559 Roy, L., Bon, M.C., Cesarini, C., Serin, J., Bonato, O., 2016. Pinpointing the level of isolation  
560 between two cryptic species sharing the same microhabitat: a case study with a scarabaeid  
561 species complex. Zoologica Scripta 45, 407-420.

562 Santos, H., Rousselet, J., Magnoux, E., Paiva, M.R., Branco, M., Kerdelhue, C., 2007. Genetic  
563 isolation through time: allochronic differentiation of a phenologically atypical population of  
564 the pine processionary moth. Proceedings of the Royal Society B-Biological Sciences 274,  
565 935-941.

566 Schmitt, T., 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends.  
567 Frontiers in zoology 4, 1.

568 Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and  
569 phylogenetic utility of mitochondrial gene sequences and a compilation of conserved  
570 polymerase chain reaction primers. Annals of the entomological Society of America 87, 651-  
571 701.

572 Stepanović, S., Kosovac, A., Krstić, O., Jović, J., Toševski, I., 2015. Morphology versus DNA  
573 barcoding: two sides of the same coin. A case study of *Ceutorhynchus erysimi* and *C.*  
574 *contractus* identification. Insect science 23, 638-648.

575 Szucs, M., Melbourne, B.A., Tuff, T., Hufbauer, R.A., 2014. The roles of demography and genetics in  
576 the early stages of colonization. Proceedings of the Royal Society B-Biological Sciences 281.

577 Szucs, M., Schaffner, U., Price, W.J., Schwarzaender, M., 2012. Post-introduction evolution in the  
578 biological control agent *Longitarsus jacobaeae* (Coleoptera: Chrysomelidae). Evolutionary  
579 Applications 5, 858-868.

580 Tavares, C., Roques, A., Courtial, B., Branco, M., Jactel, H., Lopez-Vaamonde, C., 2015.  
581 Phylogeography of the ladybird *Iberorhynchobius rondensis*, a potential biological control agent  
582 of the invasive alien pine bast scale *Matsucoccus feytaudi*. Biocontrol 60, 59-69.

- 583 Tempère, G., Péricart, J., 1989. Faune de France, Coléoptères Curculionidae Quatrième partie.  
584 Fédération Française des Sociétés de Sciences Naturelles, Paris, France.
- 585 USDA-NRCS, 2016. The plants database. United States Department of Agriculture, National  
586 Resources Conservation Service.
- 587 Virag, A., Bon, M., Cloșca, C., Diaconu, A., Haye, T., Weiss, R., Müller-Schärer, H., Hinz, H., 2016.  
588 Phenology and temperature-dependent development of *Ceutorhynchus assimilis*, a potential  
589 biological control agent for *Lepidium draba*. Journal of Applied Entomology 141, 219-230.
- 590 Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A., 2013. A general species delimitation method with  
591 applications to phylogenetic placements. Bioinformatics 29, 2869-2876.  
592  
593

594 **Figures and Tables Caption**

595

596 Fig 1. Bayesian phylogenetic tree from analysis of the *COI* sequences of *Ceutorhynchus assimilis*.  
597 Maximum likelihood bootstrap values (> 80%) are listed above the branches, followed by Bayesian  
598 posterior probabilities (> 0.75). Branch lengths represent expected substitutions per site. The scale bar  
599 indicates the expected number of substitutions per site. *Ceutorhynchus obstrictus* and *C.*  
600 *gallorhenanus* were used as outgroup. Roman numerals refer to estimated divergence times given in  
601 the Results section. Grey vertical bars represent the results of the species delimitation methods (*i.e.*  
602 ABGD (Puillandre et al., 2012) and PTP (Zhang et al., 2013) procedures).

603

604

605 Fig 2. *COI* mitochondrial haplotype network of *Ceutorhynchus assimilis*. Each circle corresponds to  
606 one haplotype; circle size gives the proportion of individuals belonging to the haplotype. The color  
607 inside each circle represents the host and indicates the proportion of individuals sampled in the  
608 different hosts. Small crossing lines indicate the number of mutations separating haplotypes. Black  
609 circles represent missing intermediate haplotypes.

610

611

612 Fig 3. Geographic distribution of the different lineages of *Ceutorhynchus assimilis*. Pie chart sizes are  
613 proportional to sample size.

614

615

616 Fig 4. Geographical distribution of the five SAMOVA derived groups identified within the 23  
617 populations of the *Lepidium* host race of *Ceutorhynchus assimilis* in the western Mediterranean region.  
618 Color codes of the populations correspond to the five groups identified by SAMOVA.

619

620

621 Table 1. Sampling details and summary statistics of genetic diversity of the sampled *Ceutorhynchus*  
622 *assimilis* populations used in this study.

623

624

625 Table 2. Estimates of evolutionary divergence over the different genetic entities (*p*-distance, expressed  
626 as a %) and results of neutrality tests for the three lineages identified within *Ceutorhynchus assimilis*.

627

628