

1 **Phylogeny of the *Formicoxenus* genus-group (Hymenoptera: Formicidae) reveals isolated**
2 **lineages of *Leptothorax acervorum* in the Iberian Peninsula predating the Last Glacial**
3 **Maximum**

4
5
6 Dario I. Ojeda^{1*}, Max John², Robert L. Hammond², Riitta Savolainen³, Kari Vepsäläinen⁴ and
7 Torstein Kvamme⁵

8
9
10 ¹Department of Forest Genetics and Biodiversity. Norwegian Institute of Bioeconomy Research
11 (NIBIO), 1430 Ås, Norway

12 ²Department of Genetics & Genome Biology. University of Leicester, University Road, Leicester,
13 LE1 7RH, UK

14 ³Department of Biosciences. University of Helsinki, P.O. Box 65, 00014, Finland

15 ⁴Retuperä Biological Station. 02880 Nuuksio, Finland

16 ⁵Department of Invertebrate Pests and Weeds in Forestry, Agriculture and Horticulture. Norwegian
17 Institute of Bioeconomy Research (NIBIO), 1430 Ås, Norway

18
19
20
21
22
23
24
25
26
27
28
29
30
31

32 * Corresponding author: dario.alayon@nibio.no, Tel: +47 47633227

33 **Abstract**

34 The *Formicoxenus* genus-group comprises six genera within the tribe Crematogastrini. The
35 group is well known for repeated evolution of social parasitism among closely related taxa and
36 cold-adapted species with large distribution ranges in the Nearctic and Palearctic regions. Previous
37 analyses based on nuclear markers (ultraconserved elements, UCEs) and mitochondrial genes
38 suggest close relationship between *Formicoxenus* Mayr, 1855, *Leptothorax* Mayr, 1855 and
39 *Harpagoxenus* Forel, 1893. However, scant sampling has limited phylogenetic assessment of these
40 genera. Also, previous phylogeographic analyses of *L. acervorum* (Fabricius, 1793) have been
41 limited to its West-Palearctic range of distribution, which has provided a narrow view on
42 recolonization, population structure and existing refugia of the species. Here, we inferred the
43 phylogenetic history of genera within the *Formicoxenus* genus-group and reconstructed the
44 phylogeography of *L. acervorum* with more extensive sampling. We employed four datasets
45 consisting of whole genomes and sequences of the COI. The topologies of previous nuclear and
46 our inferences based on mitochondrial genomes were overall congruent. Further, *Formicoxenus*
47 may not be monophyletic. We found several monophyletic lineages that do not correspond to the
48 current species described within *Leptothorax*, especially in the Nearctic region. We identified a
49 monophyletic *L. acervorum* lineage that comprises both Nearctic and Palearctic locations. The
50 most recent expansion within *L. acervorum* probably occurred within the last 0.5 Ma with isolated
51 populations predating the Last Glacial Maximum (LGM), which are localized in at least two
52 refugial areas (Pyrenean and Northern plateau) in the Iberian Peninsula. The patterns recovered
53 suggest a shared glacial refugium in the Iberian Peninsula with cold-adapted trees that currently
54 share high-altitude environments in this region.

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69 **Key words:** Phylogenomics, mitogenomes, ant, *Leptothorax*, *Formicoxenus*, Last Glacial
70 Maximum refugia, biogeography, phylogeography.

71 **1. Introduction**

72 Several invertebrate groups have species with Holarctic distribution, including beetles
73 (Larson and Nilsson, 1985), Lepidoptera (Landry et al., 2013), spiders (Marusik and Koponen,
74 2005) and ants (Schär et al., 2018). Among ant species, only three species (*Camponotus*
75 *herculeanus* Linnaeus, 1758, *Formica gagatoides* Ruzsky, 1904 and *Leptothorax acervorum*
76 Fabricius, 1793) are known to have monophyletic lineages with a Holarctic distribution (Schär et
77 al., 2018). The latter of these ant species belongs to the tribe Crematogastrini Emery, 1914 within
78 the subfamily Myrmicinae, with Crematogastrini comprising some 6,630 species (Blaimer et al.,
79 2018; Ward et al., 2015). There, recent phylogenomic analyses based on ultraconserved elements
80 (UCEs) have consistently recovered a monophyletic lineage of six genera (*Vombisidris* Bolton
81 1991, *Gauromyrmex* Menozzi, 1993, *Harpagoxenus* Forel, 1893, *Formicoxenus* Mayr, 1855,
82 *Temnothorax* Mayr, 1861 and *Leptothorax* Mayr, 1855) referred to informally as the
83 *Formicoxenus* genus-group (Blaimer et al., 2018; Branstetter et al., 2017). These analyses have
84 recovered a close relationship between *Formicoxenus* and *Leptothorax*. However, all these studies
85 have been focused on higher taxonomic relationships and few studies have included a
86 comprehensive sampling of species within each genus (Heinze and Gratiashvili, 2015; Prebus,
87 2017; Schär et al., 2018). Among the *Formicoxenus* genus-group, *Leptothorax* is the second largest
88 genus with an estimated 20 species (AntWeb ver. 8.42, <https://www.antweb.org>, accessed 29
89 October 2020). The genus has a Holarctic distribution and it has been inferred to have originated in
90 the Nearctic and dispersed in the Palearctic within the last 2 Ma (Schär et al., 2018). Relationships
91 within *Leptothorax* have not been fully resolved and previous phylogenetic analyses indicate the
92 presence of multiple undescribed and non-monophyletic taxa, particularly in the Nearctic (Heinze
93 and Gratiashvili, 2015; Schär et al., 2018). At least seven species have been described in the
94 Palearctic region, most of them with limited distribution and only *L. acervorum* with a distribution
95 both in the Nearctic and the Palearctic regions (Schär et al., 2018).

96 Taxa that span large geographic regions in both the Nearctic and Palearctic are expected to
97 have experienced variable connectivity because of the fluctuating presence of the land bridge of
98 Beringia between Eurasia and North America (DeChaine, 2008). Also, climate oscillations during
99 the Quaternary (last 2 Ma), characterized by pronounced cycles between cold glacial (ca. 100,000
100 years) and warm interglacial periods (ca. 20,000 years) during the last 700,000 years, altered the
101 geographic distribution of species in the northern hemisphere (Nearctic and Palearctic) (Hewitt,
102 2000). During these glacial episodes, species ranges contracted to refugia in suitable areas in the
103 southern part of their distribution. As the climate warmed and glaciers retreated, species with a

104 temperate range of distribution expanded and reconnected. In contrast, the range of distribution for
105 boreal cold-adapted species likely reduced and fragmented (Hewitt, 1996; Petit et al., 2003).
106 Evidence from several ant species have suggested the presence of several refugia during the
107 Pleistocene (2.58 - 0.012 Ma) in the southern Mediterranean peninsulas, the Caspian-Caucasus
108 region and further east in southern East Asia (Beibl et al., 2007; Goropashnaya et al., 2004;
109 Leppänen et al., 2013, 2011; Pusch et al., 2006; Schlick-Steiner et al., 2007). In addition,
110 populations of cold-adapted ant species could have also survived in more northerly refugia near the
111 permafrost (Leppänen et al., 2013, 2011). Indeed, *L. acervorum* is among the very few cold-
112 adapted species that extend their distribution above the polar circle both in the Nearctic and the
113 Palearctic (Berman et al., 2010; Heinze et al., 1998, 1996). In the Palearctic, this species occurs in
114 the boreal zone from the Atlantic Ocean to Japan, and in the mountains of southern Europe, the
115 Caucasus, and the Tien-Shan and Pamir (Czechowski et al., 2012; Seifert, 2018). Thus, the
116 climatic fluctuations of the Quaternary have likely played a significant role in shaping its current
117 distribution, connectivity, and genetic diversity. Populations located near the permafrost and those
118 located on the southern range of its distribution were likely affected differently.

119 The most recent phylogenetic analysis of *Leptothorax* indicate that *L. acervorum* originated
120 about 2 Ma, with the least diverged populations located in the Nearctic region. Within the
121 Palearctic, populations situated in the Iberian Peninsula were inferred to be the less divergent
122 among the specimens included (Schär et al., 2018), which might have been located in refugia
123 during the glacial cycles. In addition, more detailed analyses based on mitochondrial DNA (COI-
124 3P region) and microsatellites (SSRs) have been used to infer the phylogeography and population
125 structure of this species in the western part of its distribution (West Palearctic). These analyses
126 have found generally less population structure in *L. acervorum* compared to other species closely
127 related species within *Leptothorax*, e.g. *L. muscorum* (Nylander, 1846) and *Harpagoxenus*
128 *sublaevis* (Nylander, 1849) (Brandt et al., 2007; Foitzik et al., 2009; Trettin et al., 2016), but also
129 evidence of divergent haplotypes have been found in the Pyrenees and Southern France (Trettin et
130 al., 2016). Given the large distribution range across the Holarctic and the extensive variation in the
131 latitudinal range in western Europe (from the Iberian Peninsula to North Cape in Norway) (Heinze
132 and Holldobler, 1994), approaches that combine analyses at different taxonomic levels with
133 extensive sampling are necessary to understand the phylogenetic relationships and evolutionary
134 history of *Leptothorax* species. Here we present the most comprehensive sampling of members of
135 the *Formicoxenus* genus-group with an emphasis on the phylogenetic relationships within
136 *Leptothorax* and the biogeography of *L. acervorum* across its range of distribution in the Holarctic

137 region. Our specific objectives are: 1) to infer relationships among the six genera of the
138 *Formicoxenus* genus-group using whole mitochondrial genomes and assess their correspondence
139 with previous topologies obtained with nuclear genes, 2) to clarify the relationships of *Leptothorax*
140 species and the timing of divergence of the Palearctic species, 3) to determine if the populations of
141 *L. acervorum* situated in the Iberian Peninsula survived in different refugia during the glacial
142 cycles of the Quaternary.

143

144 **2. Materials and methods**

145 *2.1 Taxon sampling and datasets*

146 The sampling strategy used in this study was developed to represent the
147 *Formicoxenus* genus-group at three different hierarchical levels. The first dataset consisted
148 of 49 specimens representing all six genera (*Temnothorax*, *Leptothorax*, *Formicoxenus*,
149 *Harpagoxenus*, *Gauromyrmex* and *Vombisidris*) currently recognized within this group
150 (Blaimer et al., 2015; Prebus, 2017), representing 17% of genera within the
151 Crematogastrini. We also included outgroups from Myrmicinae (all tribes), Dolichoderinae
152 and Ponerinae. In this data set we used whole mitochondrial genomes to explore the major
153 relationships within *Formicoxenus* genus-group at the genus level. We used only one
154 representative specimen per species within each genus, except for *L. acervorum*, where we
155 included multiple samples (Table S1). In the second dataset, we gathered specimens
156 representing eight out of the 20 *Leptothorax* species currently recognized (AntWeb ver.
157 8.42, <https://www.antweb.org>, accessed 29 October 2020), two *Formicoxenus* species and
158 *H. sublaevis*. In this dataset we sequenced the section of the mitochondrial cytochrome c
159 oxidase (COI-5P region, 658 bp) in 96 specimens (Table S2). The third dataset consisted of
160 113 specimens of *L. acervorum* across its distribution range in the Holarctic region, where
161 we sequenced the same gene region (COI-5P region, 658 bp) as the previous dataset. This
162 dataset was complemented with available sequences from public repositories (Table S3).
163 Finally, dataset four consisted of the mitochondrial region COI-*trnL2*-COII (2,293 bp) and
164 it was obtained from a total of 80 *L. acervorum* specimens from six populations in Spain
165 and the UK (Table S3).

166

167 *2.2. Whole mitochondrial sequencing and assembly*

168 Mitochondrial genomes were newly generated for six specimens of *L. acervorum*
169 from six different populations in Spain and the UK (Spain: Valdelinares (V), Orihuela del

170 Tremedal (OT), Larra (L), Niela Refuge (NR), Pla de la Font (PF); UK: Santon Downham
171 (SD) (Table S3). A *de novo* mitochondrial genome was identified as part of a whole
172 genome sequencing project from a single adult male (PF population, sample:
173 PF18_15_M1) using 10x linked reads assembled with Supernova 2.1.1 (Weisenfeld et al.,
174 2018). The scaffold containing the mtDNA genome was identified by a BLASTn query of
175 the assembled genome with two published *L. acervorum* mtDNA sequences (query 1:
176 COXI – tRNA - Leu - COXII: GenBank: KU245569 (Trettin et al., 2016); query 2: COB:
177 GenBank: HQ259995 (Gill et al., 2009). These two sequences, located ~6Kb apart in the
178 canonical hymenopteran mtDNA genome, were used to minimize erroneous matches to
179 nuclear genomic scaffolds containing translocated mtDNA (NUMTs). Only two scaffolds
180 (102,807 and 104,071) showed convincing matches to both query sequences (E value = 0,
181 bit scores > 1000). However, mapping re-sequenced samples (see below) showed scaffold
182 102,807 had 40 times higher coverage (200x-400x) than 104,071 (~5x-10x coverage) with
183 the latter having similar coverage to the rest of the presumed nuclear genome. Furthermore,
184 scaffold 102,807 was 17Kb in length (the expected size of the mtDNA genome) whereas
185 scaffold 104,071 was longer than expected at 24Kb. These lines of evidence clearly show
186 scaffold 102,807 contains the *L. acervorum* mtDNA genome whereas scaffold 104,071 is a
187 transposition of mtDNA sequences to the nuclear genome (a NUMT).

188 To genotype single individuals in the six populations (V, OT, L, NR, PF, and SD),
189 short-read sequence data (Illumina HiSeq 2x150bp paired-end reads) were, after quality
190 control steps, aligned to the draft genome with Bowtie2 2.3.5 (Langmead and Salzberg,
191 2014) and processed with SAMtools (Li et al., 2009) to produce bam files. Bam files were
192 then subset to only include the identified mtDNA scaffold (scaffold: 102,807) with
193 SAMtools. These mtDNA alignments were converted to mpileup with BCFtools (--max-
194 depth 1000) and BCFtools call used to produce vcf files. Vcf files were indexed and
195 normalized and variants within 5bp of any indels removed with BCFtools. Finally, a fasta
196 file for each alignment was produced with BCFtools consensus (Supplementary
197 Information, Online Methods).

198 In addition, mitochondrial genomes of the taxa within the *Formicoxenus* genus-
199 group were extracted and assembled from ultra-conserved elements (UCE) libraries from
200 previous studies (Branstetter et al., 2017; Prebus, 2017) using MitoFinder (Allio et al.,
201 2020). Outgroup species within subfamilies Myrmicinae, Dolichoderinae and Ponerinae
202 were downloaded from Genbank, previously published in several studies (Cicconardi et al.,
203 2020; Du et al., 2019; Duan et al., 2016; Gotzek et al., 2010; Hasegawa et al., 2011; Liu et

204 al., 2016; Park et al., 2021, 2020b, 2020a, 2019; Rodovalho et al., 2014). SRA sequences
205 and the assembled mitochondrial genomes of *L. acervorum* are deposited in the NCBI
206 Bioproject PRJNA634471.

207

208 *2.3 Phylogenomic analyses using whole mitochondrial genomes*

209 All 49 whole mitochondrial genomes were aligned using MAFFT ver. 7.310 (Katoh
210 and Kuma, 2002) with default parameters. Visual inspection and further adjustment were
211 performed with AliView (Larsson, 2014) and summary statistics of the alignment were
212 obtained with AMAS (Borowiec, 2016). Phylogenetic analysis was performed with
213 maximum likelihood (ML) as implemented in IQ-TREE 1.6.1 (Nguyen et al., 2015) with
214 ultrafast likelihood bootstrap with 1000 replicates. The final tree was visualized and edited
215 with FigTree (Rambaut, 2016).

216

217 *2.4 DNA extraction, PCR amplification and sequencing of cytochrome c oxidase (COI)*

218 We collected either pupae or adults of workers, males, or queens from different
219 colonies of *H. sublaevis*, *L. acervorum*, *L. kutteri* (Buschinger, 1965) and *L. muscorum*
220 (Table S2). DNA was extracted from legs or whole specimens using the salt extraction
221 method (Aljanabi and Martinez, 1997) and we sequenced the portion of the mitochondrial
222 COI using previous primers and PCR conditions (Folmer et al., 1994). The sequences
223 obtained were edited, visually inspected using Sequencher (Gene Codes), and aligned with
224 AliView (Larsson, 2014). COI sequences are deposited in the NCBI Bioproject
225 PRJNA634471, under accessions XXXXXXXXX-XXXXXXX.

226

227 *2.5 Phylogenetic analyses and dating estimation within Formicoxenus-Leptothorax based* 228 *on the COI gene region*

229 To infer the phylogenetic relationships within the *Formicoxenus-Leptothorax*, we
230 used the 5' region of COI (658 bp, ranging from 5442-6601 in the *L. acervorum*
231 mitochondrial genome assembly). First, we performed an explorative analysis based on a
232 comprehensive sampling from this gene region using the sequences generated in this study
233 and available sequences from GenBank and BOLD: The Barcode of Life Data System
234 (www.barcodinglife.org). The matrix was aligned and manually edited with AliView
235 (Larsson, 2014), with summary statistics obtained with AMAS (Borowiec, 2016). This
236 preliminary analysis was based in a total of 747 specimens of *Formicoxenus* (2 spp.),
237 *Leptothorax* (8 spp.) and specimens of *H. sublaevis* (outgroup). We used maximum

238 likelihood (ML) as implemented in IQ-TREE 1.6.1 (Nguyen et al., 2015) with ultrafast
239 likelihood bootstrap with 1000 replicates (Minh et al., 2013). Based on the results from this
240 analysis (data not shown), we selected representative specimens of the major Nearctic
241 lineages identified (>80 bootstrap support) and all the specimens in the Palearctic lineage
242 of *L. acervorum*. This final dataset consisted of 96 specimens representing eight
243 *Leptothorax* spp., specimens labelled as *Leptothorax* sp., *L. muscorum* complex,
244 *Leptothorax* sp. AF CAN, two *Formicoxenus* spp. and *H. sublaevis* as an outgroup. This
245 dataset consisted of the newly generated sequences in this study and available sequences
246 from previous publications (Hebert et al., 2016; Prebus, 2017; Schär et al., 2018; Smith et
247 al., 2009; Stahlhut et al., 2013) (Table S2). The best nucleotide substitution model and ML
248 analysis were inferred with IQ-TREE 1.6.1. Clade support was assessed with ultrafast
249 likelihood bootstrap with 1000 replicates. In addition, we also performed a Bayesian
250 inference (BI) as implemented in MrBayes 3.2.6 (Huelsenbeck and Ronquist, 2001;
251 Ronquist et al., 2012) with four chains, two runs of 20 million generations with the
252 invgamma rate of variation, the GTR+ Γ model of nucleotide substitution and a sample
253 frequency of 1000. We used TRACER 1.7 (Rambaut et al., 2018) to verify whether effective
254 samples sizes (ESS values) were higher than 200 for all parameters.

255 To estimate divergence times among the lineages in *Formicoxenus-Leptothorax*, we
256 used a simplified dataset representing the same number of species as above, but fewer
257 specimens (63) of these two genera. We used BEAST 1.10.4 (Bouckaert et al., 2014;
258 Suchard et al., 2018) with a strict clock model and a constant population size under a
259 coalescence model. We employed the divergence time estimated in the Formicinae
260 (Blaimer et al., 2015) by placing a prior in the divergence estimate of *Harpagoxenus* and
261 *Formicoxenus-Leptothorax* of 8.89 (13.89-3.89) Ma. We ran two independent runs of 50
262 million generations each, sampling values every 1,000 steps. Output files were analyzed
263 with TRACER 1.7 to assess chain convergence and LogCombiner 1.10.4 was used to combine
264 independent runs. Finally, we used Treeannotator 1.10.4 to generate the maximum-clade-
265 credibility tree. ML and BI consensus trees were visualized and edited with FigTree
266 (Rambaut, 2016).

267

268 2.6 Phylogeography and genetic diversity of *Leptothorax acervorum*

269 To gain further insights into the geographic distribution of genetic diversity of *L.*
270 *acervorum* across its Holarctic distribution, we first determined the number of haplotypes,
271 haplotype diversity (*Hd*) and defined haplotypes with DnaSP ver. 6.12 (Rozas et al., 2017).

272 Then, we reconstructed the haplotype network of all 113 specimens (Table S3) using the
273 COI-5P gene region (647 bp, dataset 3) with the statistical parsimony network using TCS
274 (Clement et al., 2002) as implemented in popart ver. 1.7 (Leigh and Bryant, 2015). Given
275 that our interest was focused on the populations distributed in the Iberian Peninsula, we
276 explored in more detail five populations from this region and one population from England
277 (Table S3). In these analyses we used the mitochondrial region encompassing COI, *trnL2*
278 and COII comprising 2,293 bp (dataset 4). Haplotype (*h*), polymorphic sites and
279 nucleotide diversities (π) were calculated using the program DnaSP ver. 6.12. We
280 employed Fu's *F_s* and Tajima's *D* tests of selective neutrality to determine whether *L.*
281 *acervorum* populations from the Iberian Peninsula and England could have experienced
282 recent expansions. Fu's *F* is based in the infinite-site model and a population expansion
283 increase of rare alleles in the population, leading to negative values (Fu, 1997). Tajima's *D*
284 uses the frequency of segregating nucleotide sites and the average number of nucleotide
285 differences obtained from pairwise comparisons (Tajima, 1989). Deviations from neutrality
286 could indicate the effect of selection and/or population size changes. Population expansions
287 will increase rare alleles in the population, leading to values < 0 , while population
288 contractions (bottlenecks) will increase intermediate variants in the population. COI, *trnL2*
289 and COII sequences are deposited in the NCBI Bioproject PRJNA634471, under
290 accessions XXXXXXXXX-XXXXXXX.

291

292 **3. Results**

293 *3.1. Phylogenomic analyses of the Formicoxenus genus-group*

294 The final alignment of the mitochondrial genomes consisted of 14,351 bp with 13.32 %
295 missing data, 69% of sites variable and 59% of sites parsimony informative (Supplementary
296 Information, Data S1). We recovered monophyletic lineages for all the tribes, except Attini Smith,
297 1858 within Myrmicinae, with most branches having moderate ($>75\%$) to high ($>85\%$) bootstrap
298 support. Our phylogenetic analysis recovered all six genera of the *Formicoxenus* genus-group as a
299 monophyletic lineages within Crematogastrini, with *Formicoxenus* as the most closely related
300 genus to *Leptothorax*. The most closely related tribe was Solenopsidini Forel, 1893 (Fig.1).

301

302 *3.2 Relationships within Leptothorax and divergence times of the Palearctic species*

303 The alignment matrix of the COI-5P region (dataset 2) consisted of 658 bp with 200
304 variable sites (30.4 %), 148 parsimony informative sites (22.5%) and 1.42% of missing data. We
305 recovered the species of *Formicoxenus* on different lineages within *Leptothorax*, suggesting that

306 the former genus might not represent a monophyletic lineage. All the three Palearctic *Leptothorax*
307 species we included (*L. muscorum*, *L. gredleri* Mayr, 1855 and *L. kutteri*) represent monophyletic
308 lineages, whereas specimens assigned to *L. muscorum* from the Nearctic region represent several
309 undescribed taxa (Fig. 2 and Fig. S1). Similarly, we found non-monophyletic lineages for the other
310 Nearctic species *L. canadensis* Provancher, 1887 and *L. calderoni* Creighton, 1950, but not for *L.*
311 *retractus* Francoeur, 1986. Our divergence estimate suggests that the stem age lineages of
312 Palearctic taxa (*L. gredleri*, *L. muscorum*, *L. kutteri* and *L. acervorum*) ranges between 1-1.6 Ma
313 (Fig. 3). The crown age of the Palearctic lineage of *L. acervorum* was estimated at 0.56 Ma, with
314 specimens from the Iberian Peninsula ranging in age between 0.1 and 0.5 Ma. The most recent
315 derived lineage (0.30 Ma) within *L. acervorum* comprises both specimens from the Nearctic and
316 Palearctic distribution, including specimens at high latitudes mainly from the West Nearctic (Fig.
317 4).

318 3.3 Phylogeography of *L. acervorum* in the Holarctic region and genetic diversity in the 319 populations of the Iberian Peninsula

320 Our analyses based on the COI-5P region (dataset 3, excluding gaps and missing data),
321 recovered 21 variable sites with 19 haplotypes ($Hd = 0.826$) among the 113 specimens examined
322 (Table S3). Only two populations from the Iberian Peninsula (Larra and Pla de la Font) shared
323 haplotypes with the rest of the populations in the West Palearctic, East Palearctic and the Nearctic.
324 We also found a unique haplotype (H4) shared between the population in England and
325 Switzerland, and the presence of unique haplotypes in Bulgaria and Kamchatka. The most widely
326 distributed haplotype (H11) was shared across the entire geographic distributional range (Table 1),
327 and it likely represents the most recent expansion across the distribution of *L. acervorum*. The
328 haplotype network indicates that most populations in the Iberian Peninsula have been isolated from
329 the remaining range of distribution in the West Palearctic, with a recent expansion of the haplotype
330 H11 into this region (Fig. 5). Our examination of the genetic diversity in the populations in the
331 Iberian Peninsula based on a larger segment of the mitochondrial COI-*trnL2*-COII (dataset 4)
332 suggests higher genetic diversity in Niela refuge and Larra than in the other Iberian populations
333 (Table 2).

334 4. Discussion

335 4.1 Congruence between phylogenomic inferences based on mitochondrial genomes and UCEs

336 Recent analyses using large sequence datasets from UCEs have been employed to resolve
337 relationships among ant subfamilies (Blaimer et al., 2015; Branstetter et al., 2017; Li et al., 2018;
338 Longino and Branstetter, 2021). Phylogenomic inferences using mitochondrial genomes have also

339 been used in several ant subfamilies with mostly congruent topologies between nuclear (UCEs)
340 and mitochondrial genomes (Allio et al., 2020). Although we did not have a comprehensive
341 mitochondrial representation within each of the six tribes currently recognized within Myrmicinae
342 (Borowiec et al., 2020), the topology we recovered from this analysis (Fig. 1) is congruent with the
343 known relationships among these tribes inferred from nuclear genes (Ward et al., 2015;
344 https://antwiki.org/wiki/Phylogeny_of_Myrmicinae). Crematogastrini harbors 40% of all
345 Myrmicinae species and 45% of the genera belong to this tribe (Blaimer et al., 2018). Recent
346 phylogenomic analyses based on UCEs (Blaimer et al., 2018; Prebus, 2017), as well as
347 mitochondrial genomes (Park et al., 2021) have been used to increase resolution among genera of
348 of Crematogastrini. Analyses using UCEs have recovered eight clades with high support within
349 Crematogastrini, which have been treated as informal genus-groups.

350 The *Formicoxenus* genus group consists of six genera (*Formicoxenus*, *Leptothorax*,
351 *Vombisidris*, *Gauromyrmex*, *Harpagoxenus* and *Temnothorax*) and the relationships among the
352 genera are relatively well established (Blaimer et al., 2018; Prebus, 2017), but less attention has
353 been paid to the *Leptothorax* genus group, which consists of *Formicoxenus*, *Leptothorax* and
354 *Harpagoxenus*. Species within these genera are prone to develop social parasitism, a set of
355 interrelated lifestyles where the parasitic species depend upon a free-living host to complete their
356 life cycle (Beibl et al., 2005; Heinze, 1995). These three genera, together with *Temnothorax*, are
357 considered a hot spot for the evolution of social parasitism, where it has evolved at least five times
358 among closely related taxa (Beibl et al., 2005; Jongepier et al., 2021; Prebus, 2017).
359 *Formicoxenus* comprises about seven species of Nearctic or Palearctic distribution, with some
360 species (“guest ants”) living in the nest of *Formica* Linnaeus, 1758, *Myrmica* Latreille, 1804 or
361 *Manica* Jurine, 1807 species. Workers of *Formicoxenus* species beg for food from their host
362 species, but rear their own brood in own chambers within the host nest (Franceur et al., 1985;
363 Heinze, 1995). *Leptothorax* contains some species that are workless social parasites that tolerate
364 (inquilines, e.g., *L. kutteri*, *L. pacis* (Kutter, 1945) or kill the host queen (murder parasites, e.g., *L.*
365 *goesswaldi* Kutter, 1967, *L. wilsoni* Heinze, 1989) of other closely related *Leptothorax* species
366 (Foitzik et al., 2009; Heinze and Ortius, 1991). In contrast, *Harpagoxenus* taxa are slave-making
367 (dulotic) ants whose queens kill or expel all adult residents after invading *Leptothorax* spp. (Brandt
368 et al., 2007; Fischer-Blass et al., 2006; Heinze and Ortius, 1991; Pusch et al., 2006).

369
370
371

372 4.2 Evidence on non-monophyletic taxa within *Leptothorax* and monophyletic lineage within *L.*
373 *acervorum*

374 Mitochondrial genes (cytochrome b and cytochrome oxidase subunit 1) alone or in
375 combination with other nuclear markers have been previously used in phylogenetic inferences in
376 *Leptothorax* (Baur et al., 1996, 1995; Beibl et al., 2005; Heinze and Gratiashvili, 2015; Schär et
377 al., 2018), but with limited representation of its species or without using other genera within the
378 *Formicoxenus* genus-group. Our more comprehensive sampling (42% of *Leptothorax* species,
379 AntWeb ver. 8.42, <https://www.antweb.org>, accessed 29 October 2020 and multiple accessions of
380 *L. acervorum*) supports a close relationship between *Formicoxenus* and *Leptothorax* (Fig. 2),
381 similar to previous analyses based on limited taxa of this group (Blaimer et al., 2018; Prebus,
382 2017; Ward et al., 2015). As it has been previously suggested (Heinze and Gratiashvili, 2015;
383 Heinze and Ortius, 1991; Schär et al., 2018), some taxa within the genus *Leptothorax*, particularly
384 the Nearctic ones, represent species groups that deserve taxonomic adjustments. Our analyses
385 suggest the presence of at least four Nearctic lineages (monophyletic groups with moderate to high
386 support) comprising taxa currently assigned to *L. muscorum* (Nearctic), *L. canadensis*, *L.*
387 *calderoni*, *L. sphagnicola* and specimens assigned to the *L. muscorum* complex. This latter
388 complex seems to consist of a species group of three to four different taxa from the Nearctic that
389 display a set of similar morphological characters and chromosome numbers (Brown, 1955; Heinze,
390 1991, 1989; Loiselle et al., 1990). One of these lineages also includes a specimen of *L. acervorum*
391 (LT977594), which is more closely related to the Palearctic taxa *L. muscorum*, *L. gredleri* and *L.*
392 *kutteri* (Fig. 2). This latter lineage deserves further exploration, as it involves determining whether
393 *L. acervorum* in the Nearctic represents a separate lineage from the remaining samples we
394 included.

395 In contrast, the Palearctic species *L. gredleri*, *L. muscorum* and the inquiline *L. kutteri* most
396 likely represent monophyletic lineages (Fig. 2). The lineage of *L. acervorum* comprises both
397 specimens from the Palearctic and Nearctic regions and our divergence age estimates suggest that
398 this clade likely represents the most recent diversification event, within the last 0.5 Ma (Fig. 3).
399 Despite the high support values observed for most of the lineages in the Palearctic region, the
400 support values for the mutual relationships of the lineages were low, and therefore more
401 informative regions will be necessary to determine their relationships.

402
403

404 4.3 Evidence of isolated populations in the Iberian Peninsula with limited contribution to the most
405 recent expansion of *L. acervorum*

406 *Leptothorax acervorum* is one of the only three ant species with Holarctic distribution
407 (Schär et al., 2018); however, all current phylogeographic analyses of this species based on
408 alloenzymes, SSRs, mtDNA and nuclear markers have included only Palearctic specimens (Brandt
409 et al., 2007; Foitzik et al., 2009; Gill et al., 2009; Heinze et al., 1994; Stille and Stille, 1993;
410 Trettin et al., 2016), thus providing an incomplete picture of the patterns of recolonization and
411 populations structure. Previous phylogeographic analyses within *L. acervorum* species have been
412 based on the COI 3'P region and indicate substantial genetic diversity within the species (Brandt et
413 al., 2007; Foitzik et al., 2009; Trettin et al., 2016). The most recent analyses based on SSRs and
414 mtDNA (COI 3'P region) have found the existence of multiple refugia in SW-Europe, and
415 evidence of spatial genetic structure across the sampled area (Trettin et al., 2016). Our most
416 extensive data set, including specimens we identified in the previous phylogenetic analyses (Fig. 2)
417 from the Nearctic region and several populations from the Iberian Peninsula (IP), suggests that the
418 IP populations represent the less derived lineages and that they might have experienced
419 fragmentation and isolation from the remaining Holarctic distribution (Fig. 5). *Leptothorax*
420 *acervorum* is a cold-adapted species that in the IP inhabits mostly mountainous pinewoods and
421 pine-dominated forest (*Pinus sylvestris*) above 1500 m.a.s.l (Felke and Buschinger, 1999; Gill et
422 al., 2009). Our results indicate that all populations we included within the IP, except Larra and Pla
423 de la Font, seem to have been more isolated from the remaining range of distribution of this
424 species (Fig. 4), supporting previous evidence based on SSRs, which have found evidence of
425 bottlenecks and varying levels of connectivity in this area (Trettin et al., 2016). However, there are
426 only a few mutations separating even the most divergent haplotypes among these populations, but
427 these divergent haplotypes in the IP tend to be found in altitudinally restricted populations. In
428 contrast, the most common haplotype (H11) is found in locations where *L. acervorum* is not
429 altitudinally restricted (Fig. 5) and there is greater connectivity of suitable habitat. The lack of
430 spatial genetic structure previously reported within *L. acervorum* using mtDNA (Brandt et al.,
431 2007; Foitzik et al., 2009; Trettin et al., 2016) might be explained by the limited sampling outside
432 the West-Palearctic regions in previous studies, as well as that this lineage represents the group
433 with the most recent expansion (Fig. 3). Additional sampling across the Holarctic distribution with
434 denser sampling among populations, together with the inclusion of additional markers, would be
435 required to further expand the phylogeographic signal we recovered in our analyses.

436 Several refugia areas have been identified in the IP based on the ant species in this region
437 (Tinaut and Ruano, 2021), and our results suggests that only populations from the Pyrenean

438 refugia might have more recent connection with the rest of the West Palearctic range of
439 distribution (Fig. 4). In contrast, the populations located in the Cantabric and the Northern Plateau
440 (Tinaut and Ruano, 2021) were likely more isolated from the rest of the populations. Cold-adapted
441 species (boreal) with wide distribution in the Palearctic could survive in periglacial areas during
442 the periods of maximum glacial expansion (e.g., during the LGM, 23-18 ka BP), expanding their
443 range into southern areas. During periods of postglacial warming, southern populations of these
444 species became isolated in mountainous regions (Schmitt, 2009; Schmitt and Varga, 2012),
445 surviving in southern refugia (Stewart et al., 2010). There is extensive evidence of the glacial-
446 interglacial cycles during the Quaternary having influenced the individual genetic diversity and
447 population structure of plants and animals in the West Palearctic (Bennett et al., 1991; Morales-
448 Barbero et al., 2018; Schmitt and Varga, 2012; Stewart et al., 2010), including the presence of
449 several periglacial and southern refugia of cold-tolerant of *Myrmica* (Leppänen et al., 2013, 2011)
450 and *Formica* ant species (Goropashnaya et al., 2007, 2004). Emerging evidence seems to indicate
451 that these glacial-interglacial cycles could also have shared refugia; for example, the congruent
452 phylogeographic signal between *Myrmica* ants and *Betula* species (Leppänen et al., 2011;
453 Maliouchenko et al., 2007), the leaf beetle *Gonioctena intermedia* and its boreal-temperate host
454 trees *Prunus padus* and *Sorbus aucuparia* (Quinzin et al., 2017), and the similar patterns of
455 isolated populations in the Iberian Peninsula observed between *L. acervorum* (Trettin et al., 2016)
456 and *Pinus sylvestris* (Dering et al., 2017; Tyrmi et al., 2020).

457

458 **Acknowledgements**

459 The phylogenomic analyses based on whole mitochondrial genomes were performed on resources
460 provided by UNINETT Sigma2 - the National Infrastructure for High Performance Computing and
461 Data Storage in Norway. Funding was provided by NIBIO under the ForGeBiM project. MJ and
462 RLH's contribution were funded by BBSRC MIBTP PhD studentship to MJ supervised by RLH.
463 We thank Mari Mette Tollefsrud for drawing the map in Fig. 4.

464

465 **Author Contribution**

466 DIO, TK, JM and RB conceived the idea; DIO, KV, RS and JM performed the laboratory and
467 analyses; MJ and RLH assembled the *L. acervorum* mtDNA genome, identified the mtDNA
468 genomic scaffold and genotyped re-sequenced samples; DIO wrote the manuscript with
469 contributions from all authors.

470

471

472 **References**

- 473 Aljanabi, S.M., Martinez, I., 1997. Universal and rapid salt-extraction of high quality genomic
474 DNA for PCR-based techniques. *Nucleic Acids Res.* 25, 4692–4693.
- 475 Allio, R., Schomaker-Bastos, A., Romiguier, J., Prosdocimi, F., Nabholz, B., Delsuc, F., 2020.
476 MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target
477 enrichment phylogenomics. *Mol. Ecol. Resour.* 20, 892–905.
- 478 Baur, A., Chalwatzis, N., Buschinger, A., Zimmermann, F., 1995. Mitochondrial DNA sequences
479 reveal close relationships between social parasitic ants and their host species. *Curr. Genet.* 28,
480 242–247.
- 481 Baur, A., Sanetra, M., Chalwatzis, N., Buschinger, A., Zimmermann, F., 1996. Sequence
482 comparisons of the internal transcribed spacer region of ribosomal genes support close
483 relationships between parasitic ants and their respective host species (Hymenoptera:
484 Formicidae). *Insectes Soc.* 43, 53–67.
- 485 Beibl, J., Buschinger, A., Foitzik, S., Heinze, J., 2007. Phylogeny and phylogeography of the
486 Mediterranean species of the parasitic ant genus *Chalepoxenus* and its *Temnothorax* hosts.
487 *Insectes Soc.* 54, 189–199.
- 488 Beibl, J., Stuart, R.J., Heinze, J., Foitzik, S., 2005. Six origins of slavery in formicoxenine ants.
489 *Insectes Soc.* 52, 291–297.
- 490 Bennett, K.D., Tzedakis, P.C., Willis, K.J., 1991. Quaternary Refugia of North European Trees. *J.*
491 *Biogeogr.* 18, 103–115.
- 492 Berman, D.I., Alfimov, A.V., Zhigul'skaya, Z.A., Leirikh, A.N., 2010. Overwintering and cold-
493 hardiness of ants in the northeast of Asia. Pensoft, Sofia, Moscow.
- 494 Blaimer, B.B., Brady, S.G., Schultz, T.R., Lloyd, M.W., Fisher, B.L., Ward, P.S., 2015.
495 Phylogenomic methods outperform traditional multi-locus approaches in resolving deep
496 evolutionary history: A case study of formicine ants. *BMC Evol. Biol.* 15, 1–14.
- 497 Blaimer, B.B., Ward, P.S., Schultz, T.R., Fisher, B.L., Brady, S.G., 2018. Paleotropical
498 diversification dominates the evolution of the hyperdiverse ant tribe *Crematogastrini*
499 (Hymenoptera: Formicidae). *Insect Syst. Divers.* 2.
- 500 Borowiec, M., Moreau, C., Rabeling, C., 2020. Ants: phylogeny and classification, in: Starr, C.

- 501 (Ed.), *Encyclopedia of Social Insects*. Springer, Switzerland, pp. 1–18.
- 502 Borowiec, M.L., 2016. AMAS: a fast tool for alignment manipulation and computing of summary
503 statistics. *PeerJ* 4, e1660.
- 504 Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut,
505 A., Drummond, A.J., 2014. BEAST 2: A software platform for Bayesian evolutionary
506 analysis. *PLoS Comput. Biol.* 10, e1003537.
- 507 Brandt, M., Fischer-Blass, B., Heinze, J., Foitzik, S., 2007. Population structure and the co-
508 evolution between social parasites and their hosts. *Mol. Ecol.* 16, 2063–2078.
- 509 Branstetter, M.G., Longino, J.T., Ward, P.S., Faircloth, B.C., 2017. Enriching the ant tree of life:
510 enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera.
511 *Methods Ecol. Evol.* 8, 768–776.
- 512 Brown, W., 1955. The ant *Leptothorax muscorum* (Nylander) in North America. *Entomol. News*.
- 513 Cicconardi, F., Krapf, P., D’Annessa, I., Gamisch, A., Wagner, H.C., Nguyen, A.D., Economo,
514 E.P., Mikheyev, A.S., Guenard, B., Grabherr, R., Andesner, P., Wolfgang, A., Di Marino, D.,
515 Steiner, F.M., Schlick-Steiner, B.C., 2020. Genomic signature of shifts in selection in a
516 subalpine ant and its physiological adaptations. *Mol. Biol. Evol.* 37, 2211–2227.
- 517 Clement, M., Snell, Q., Walke, P., Posada, D., Crandall, K., 2002. TCS: estimating gene
518 genealogies, in: *Proceeding 16th International Parallel Distributed Processing Symposium*. p.
519 184.
- 520 Czechowski, W., Radchenko, A., Czechowska, W., Vepsäläinen, K., 2012. The ants of Poland
521 with reference to the myrmecofauna of Europe. *Muzeum i Instytut Zoologii Polskiej*
522 *Akademii Nauk*.
- 523 DeChaine, E.G., 2008. A bridge or a barrier? Beringia’s influence on the distribution and diversity
524 of tundra plants. *Plant Ecol. Divers.* 1, 197–207.
- 525 Dering, M., Kosiński, P., Wyka, T.P., Pers-Kamczyc, E., Boratyński, A., Boratyńska, K., Reich,
526 P.B., Romo, A., Zadworny, M., Żytkowiak, R., Oleksyn, J., 2017. Tertiary remnants and
527 Holocene colonizers: Genetic structure and phylogeography of Scots pine reveal higher
528 genetic diversity in young boreal than in relict Mediterranean populations and a dual
529 colonization of Fennoscandia. *Divers. Distrib.* 23, 540–555.

- 530 Du, Y., Song, X., Yu, H., Lu, Z., 2019. Complete mitochondrial genome sequence of *Tapinoma*
531 *melanocephalum* (Hymenoptera: Formicidae). *Mitochondrial DNA Part B Resour.* 4, 3448–
532 3449.
- 533 Duan, X.Y., Peng, X.Y., Qian, Z.Q., 2016. The complete mitochondrial genomes of two globally
534 invasive ants, the Argentine ant *Linepithema humile* and the little fire ant *Wasmannia*
535 *auropunctata*. *Conserv. Genet. Resour.* 8, 275–277.
- 536 Feldmeyer, B., Elsner, D., Alleman, A., Foitzik, S., 2017. Species-specific genes under selection
537 characterize the co-evolution of slavemaker and host lifestyles. *BMC Evol. Biol.* 17, 1–11.
538 <https://doi.org/10.1186/s12862-017-1078-9>
- 539 Felke, M., Buschinger, A., 1999. Social organization, reproductive behavior and ecology of
540 *Leptothorax acervorum* (Hymenoptera, Formicidae) from the Sierra de Albarracin in central
541 Spain. *Insectes Soc.* 46, 84–91.
- 542 Fischer-Blass, B., Heinze, J., Foitzik, S., 2006. Microsatellite analysis reveals strong but
543 differential impact of a social parasite on its two host species. *Mol. Ecol.* 15, 863–872.
- 544 Foitzik, S., Bauer, S., Laurent, S., Pennings, P.S., 2009. Genetic diversity, population structure and
545 sex-biased dispersal in three co-evolving species. *J. Evol. Biol.* 22, 2470–2480.
- 546 Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification
547 of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol.*
548 *Mar. Biol. Biotechnol.* 3, 294–299.
- 549 Franceur, A., Loiselle, R., Buschinger, A., 1985. Biosystématique de la tribu Leptothoracini
550 (Formicidae, Hymenoptera). 1. Le genre *Formicoxenus* dans la région holarctique. *Nat. Can.*
551 112, 343–403.
- 552 Fu, X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and
553 background selection. *Genetics* 147, 915–925.
- 554 Gill, R.J., Arce, A., Keller, L., Hammond, R.L., 2009. Polymorphic social organization in an ant.
555 *Proc. R. Soc. B Biol. Sci.* 276, 4423–4431.
- 556 Goropashnaya, A., Fedorov, V., Seifert, B., Pamilo, P., 2007. Phylogeography and population
557 structure in the ant *Formica exsecta* (Hymenoptera, Formicidae) across Eurasia as reflected
558 by mitochondrial DNA variation and microsatellites. *Ann. Zool. Fenn.* 44, 462–474.

- 559 Goropashnaya, A., Fedorov, V., Seifert, B., Pamilo, P., 2004. Limited phylogeographical structure
560 across Eurasia in two red wood ant species *Formica pratensis* and *F. lugubris* (Hymenoptera,
561 Formicidae). *Mol. Ecol.* 13, 1849–1858.
- 562 Gotzek, D., Clarke, J., Shoemaker, D., 2010. Mitochondrial genome evolution in fire ants
563 (Hymenoptera: Formicidae). *BMC Evol. Biol.* 10, 300.
- 564 Guénard, B., Weiser, M.D., Gómez, K., Narula, N., Economo, E.P., 2017. The Global Ant
565 Biodiversity Informatics (GABI) database: Synthesizing data on the geographic distribution
566 of ant species (Hymenoptera: Formicidae). *Myrmecological News* 24, 83–89.
- 567 Hasegawa, E., Kobayashi, K., Yagi, N., Tsuji, K., 2011. Complete mitochondrial genomes of
568 normal and cheater morphs in the parthenogenetic ant *Pristomyrmex punctatus*
569 (Hymenoptera: Formicidae). *Myrmecological News* 15, 85–90.
- 570 Hebert, P.D.N., Ratnasingham, S., Zakharov, E. V., Telfer, A.C., Levesque-Beaudin, V., Milton,
571 M.A., Pedersen, S., Jannetta, P., Dewaard, J.R., 2016. Counting animal species with DNA
572 barcodes: Canadian insects. *Philos. Trans. R. Soc. B Biol. Sci.* 371, 20150333.
- 573 Heinze, J., 1991. Biochemical studies on the relationship between socially parasitic ants and their
574 hosts. *Biochem. Syst. Ecol.* 19, 195–206.
- 575 Heinze, J., 1989. A biochemical approach toward the systematics of the *Leptothorax*
576 “*muscorum*” Group in North America (Hymenoptera: Formicidae). *Biochem. Syst. Ecol.* 17,
577 595–601.
- 578 Heinze, J., Foitzik, S., Kipyatkov, V.E., Lopatina, E.B., 1998. Latitudinal variation in cold
579 hardiness and body size in the boreal ant species *Leptothorax acervorum* (Hymenoptera,
580 Formicidae). *Entomol. Gen.* 22, 305–312.
- 581 Heinze, J., Gratiashvili, N., 2015. High skew in the Caucasus: functional monogyny in the ant
582 *Leptothorax scamni*. *Insectes Soc.* 62, 385–392.
- 583 Heinze, J., Holldobler, B., 1994. Ants in the cold. *Memorab. Zool.* 48, 99–108.
- 584 Heinze, J., Lipski, N., Schlehmeier, K., Hölldobler, B., 1994. Colony structure and reproduction in
585 the ant, *Leptothorax acervorum*. *Behav. Ecol.* 4, 359–367.
- 586 Heinze, J., Ortius, D., 1991. Social organization of *Leptothorax acervorum* from Alaska
587 (Hymenoptera: Formicidae). *Psyche (New York)* 98, 227–240.

- 588 Heinze, J., Stahl, M., Hölldobler, B., 1996. Ecophysiology of hibernation in boreal *Leptothorax*
589 ants (Hymenoptera: Formicidae). *Écoscience* 3, 429–435.
- 590 Heinze, J.ürgen, 1995. The origin of workerless parasites in *Leptothorax* (s.str) (Hymenoptera:
591 Formicidae). *Psyche* (Stuttg). 102, 195–214.
- 592 Hewitt, G., 2000. The genetic legacy of the quaternary ice ages. *Nature* 405, 907–913.
- 593 Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and
594 speciation. *Biol. J. Linn. Soc.* 58, 247–276.
- 595 Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny.
596 *Bioinformatics* 17, 754–755.
- 597 Jongepier, E., Séguret, A., Labutin, A., Feldmeyer, B., Gstöttl, C., Foitzik, S., Heinze, J.,
598 Bornberg-Bauer, E., 2021. Convergent loss of chemoreceptors across independent origins of
599 slave-making in ants. *bioRxiv* 2021.05.11.443570.
- 600 Katoh, M., Kuma, M., 2002. MAFFT: a novel method for rapid multiple sequence alignment based
601 on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.
- 602 Landry, J.F., Nazari, V., Dewaard, J.R., Mutanen, M., Lopez-Vaamonde, C., Huemer, P., Hebert,
603 P.D.N., 2013. Shared but overlooked: 30 species of Holarctic Microlepidoptera revealed by
604 DNA barcodes and morphology. *Zootaxa* 3749, 1–93.
- 605 Langmead, B., Salzberg, S., 2014. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9,
606 357359.
- 607 Larson, D.J., Nilsson, A.N., 1985. The holarctic species of *agabus* (sensu lato) leach (coleoptera:
608 Dytiscidae). *Can. Entomol.* 117, 119–130.
- 609 Larsson, A., 2014. AliView: A fast and lightweight alignment viewer and editor for large datasets.
610 *Bioinformatics* 30, 3276–3278.
- 611 Leigh, J.W., Bryant, D., 2015. POPART: Full-feature software for haplotype network construction.
612 *Methods Ecol. Evol.* 6, 1110–1116.
- 613 Leppänen, J., Vepsäläinen, K., Anthoni, H., Savolainen, R., 2013. Comparative phylogeography of
614 the ants *Myrmica ruginodis* and *Myrmica rubra*. *J. Biogeogr.* 40, 479–491.
- 615 Leppänen, J., Vepsäläinen, K., Savolainen, R., 2011. Phylogeography of the ant *Myrmica rubra*
616 and its inquiline social parasite. *Ecol. Evol.* 1, 46–62.

- 617 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G.,
618 Durbin, R., 2009. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*
619 25, 2078–2079.
- 620 Li, H., Sosa-Calvo, J., Horn, H.A., Pupo, M.T., Clardy, J., Rabeling, C., Schultz, T.R., Currie,
621 C.R., 2018. Convergent evolution of complex structures for ant-bacterial defensive symbiosis
622 in fungus-farming ants. *Proc. Natl. Acad. Sci. U. S. A.* 115, 10720–10725.
- 623 Liu, N., Duan, X.Y., Qian, Z.Q., Wang, X.Y., Li, X.L., Ding, M.Y., 2016. Characterization of the
624 complete mitochondrial genome of the myrmicine ant *Vollenhovia emeryi* (Insecta:
625 Hymenoptera: Formicidae). *Conserv. Genet. Resour.* 8, 211–214.
- 626 Loiselle, R., Francoeur, A., Fischer, K., Buschinger, A., 1990. Variations and taxonomic
627 significance of the chromosome numbers in the Nearctic species of the genus *Leptothorax*
628 (s.s.) (Formicidae: Hymenoptera). *Caryologica* 43, 321–334.
- 629 Longino, J.T., Branstetter, M.G., 2021. Integrating UCE phylogenomics with traditional taxonomy
630 reveals a trove of New World *Syscia* species (Formicidae: Dorylinae). *Insect Syst. Divers.* 5,
631 1–51.
- 632 Maliouchenko, O., Palmé, A.E., Buonamici, A., Vendramin, G.G., Lascoux, M., 2007.
633 Comparative phylogeography and population structure of European *Betula* species, with
634 particular focus on *B. pendula* and *B. pubescens*. *J. Biogeogr.* 34, 1601–1610.
- 635 Marusik, Y.M., Koponen, S., 2005. A survey of spiders (Araneae) with holarctic distribution. *J.*
636 *Arachnol.* 33, 300–305.
- 637 Minh, B.Q., Nguyen, M.A.T., Von Haeseler, A., 2013. Ultrafast approximation for phylogenetic
638 bootstrap. *Mol. Biol. Evol.* 30, 1188–1195.
- 639 Morales-Barbero, J., Martinez, P.A., Ferrer-Castán, D., Olalla-Tárraga, M., 2018. Quaternary
640 refugia are associated with higher speciation rates in mammalian faunas of the Western
641 Palearctic. *Ecography (Cop.)*. 41, 607–621.
- 642 Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A fast and effective
643 stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32,
644 268–274.
- 645 Park, Jonghyun, Kwon, W., Park, Jongsun, 2019. The complete mitochondrial genome of
646 *Ectomomyrmex javanus* Mayr, 1867 (Hymenoptera: Formicidae). *Mitochondrial DNA Part B*

- 647 Resour. 4, 1636–1637.
- 648 Park, Jonghyun, Xi, H., Park, Jongsun, 2021. Complete mitochondrial genome of the acrobat ant
649 *Crematogaster teranishii* Santschi, 1930 (Formicidae; Hymenoptera). Mitochondrial DNA
650 Part B Resour. 6, 593–595.
- 651 Park, Jonghyun, Xi, H., Park, Jongsun, 2020a. The complete mitochondrial genome of
652 *Aphaenogaster famelica* (Smith, 1874) (Hymenoptera: Formicidae). Mitochondrial DNA Part
653 B 5, 492–494.
- 654 Park, Jonghyun, Xi, H., Park, Jongsun, 2020b. The complete mitochondrial genome of *Ochetellus*
655 *glaber* (Mayr, 1862) (Hymenoptera:Formicidae). Mitochondrial DNA Part B 5, 147–149.
- 656 Petit, R., Aguinagalde, I., De Beaulieu, J., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R.,
657 Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Müller-Starck, G., Demesure-Musch, B.,
658 Palmé, A., Martín, J.P., Rendell, S., Vendramin, G.G., 2003. Glacial refugia: Hotspots but not
659 melting pots of genetic diversity. *Science* (80-.). 300, 1563–1565.
- 660 Prebus, M., 2017. Insights into the evolution , biogeography and natural history of the acorn ants ,
661 genus *Temnothorax* Mayr (Hymenoptera: Formicidae). *BMC Evol. Biol.* 17, 250.
- 662 Pusch, K., Seifert, B., Foitzik, S., Heinze, J., 2006. Distribution and genetic divergence of two
663 parapatric sibling ant species in Central Europe. *Biol. J. Linn. Soc.* 88, 223–234.
- 664 Quinzin, M.C., Normand, S., Dellicour, S., Svenning, J.C., Mardulyn, P., 2017. Glacial survival of
665 trophically linked boreal species in northern Europe. *Proc. R. Soc. B Biol. Sci.* 284.
- 666 Rambaut, A., 2016. FigTree ver. 1.4.3.
- 667 Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarisation
668 in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 0, 1–5.
- 669 Rodvalho, C.D.M., Lyra, M.L., Ferro, M., Bacci, M., 2014. The mitochondrial genome of the
670 leaf-cutter ant *Atta laevigata*: A mitogenome with a large number of intergenic spacers. *PLoS*
671 *One* 9, 1–9.
- 672 Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu,
673 L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic
674 inference and model choice across a large model space. *Syst. Biol.* 61, 539–542.
- 675 Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins,

- 676 S.E., Sanchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large
677 data sets. *Mol. Biol. Evol.* 34, 3299–3302.
- 678 Schär, S., Talavera, G., Espadaler, X., Rana, J.D., Andersen Andersen, A., Cover, S.P., Vila, R.,
679 2018. Do Holarctic ant species exist? Trans-Beringian dispersal and homoplasmy in the
680 Formicidae. *J. Biogeogr.* 45, 1917–1928.
- 681 Schlick-Steiner, B.C., Steiner, F.M., Sanetra, M., Seifert, B., Christian, E., Stauffer, C., 2007.
682 Lineage specific evolution of an alternative social strategy in *Tetramorium* ants
683 (Hymenoptera: Formicidae). *Biol. J. Linn. Soc.* 91, 247–255.
- 684 Schmitt, T., 2009. Biogeographical and evolutionary importance of the European high mountain
685 systems. *Front. Zool.* 6, 1–10.
- 686 Schmitt, T., Varga, Z., 2012. Extra-Mediterranean refugia: The rule and not the exception? *Front.*
687 *Zool.* 9, 1–12.
- 688 Seifert, B., 2018. The ants of central and northern Europe. Lutra, Germany.
- 689 Smith, M.A., Fernandez-Triana, J., Roughley, R., Hebert, P.D.N., 2009. DNA barcode
690 accumulation curves for understudied taxa and areas. *Mol. Ecol. Resour.* 9, 208–216.
- 691 Stahlhut, J.K., Fernández-Triana, J., Adamowicz, S.J., Buck, M., Goulet, H., Hebert, P.D.N.,
692 Huber, J.T., Merilo, M.T., Sheffield, C.S., Woodcock, T., Smith, M.A., 2013. DNA barcoding
693 reveals diversity of Hymenoptera and the dominance of parasitoids in a sub-arctic
694 environment. *BMC Ecol.* 13, 2.
- 695 Stewart, J.R., Lister, A.M., Barnes, I., Dalén, L., 2010. Refugia revisited: Individualistic responses
696 of species in space and time. *Proc. R. Soc. B Biol. Sci.* 277, 661–671.
- 697 Stille, M., Stille, B., 1993. Intrapopulation nestclusters of maternal mtDNA lineages in the
698 polygynous ant *Leptothorax acervorum* (Hymenoptera: Formicidae). *Insect Mol. Biol.* 1,
699 117–121.
- 700 Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian
701 phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* 4, vey016.
- 702 Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA
703 polymorphism. *Genetics* 123, 585–595.
- 704 Tinaut, A., Ruano, F., 2021. Biogeography of Iberian ants (Hymenoptera: Formicidae). *Diversity*

705 13, 1–25.

706 Trettin, J., Agrawal, S., Heinze, J., 2016. Phylogeography of social polymorphism in a boreo-
707 montane ant. *BMC Evol. Biol.* 16, 1–14.

708 Tyrmi, J., Vuosku, J., Acosta, J.J., Li, Z., Sterck, L., Savolainen, O., Pyhäjärvi, T., 2020.
709 Genomics of clinal local adaptation in *Pinus sylvestris* under continuous environmental and
710 spatial genetic setting. *G3 Genes Genomes Genet.* 10, 2638–2696.

711 Ward, P.S., Brady, S.G., Fisher, B.L., Schultz, T.R., 2015. The evolution of myrmicine ants:
712 Phylogeny and biogeography of a hyperdiverse ant clade (Hymenoptera: Formicidae). *Syst.*
713 *Entomol.* 40, 61–81.

714 Weisenfeld, N.I., Kumar, V., Shah, P., Church, D.M., Jaffe, D.B., 2018. Direct determination of
715 diploid genome sequences. *Genome Res.* 28, 757–767.

716

717 **Supporting Information**

718 **Figures**

719 **Fig. S1.** Consensus tree obtained with Bayesian inference as implemented in MrBayes on the
720 dataset of COI-5P region. The black circle indicates the monophyletic lineage of *L. acervorum*
721 (plus specimens of *L. muscorum* complex and *Leptothorax* sp.) that were later used in the
722 phylogeographic analysis. Values next to the branches represent Bayesian support.

723

724 **Tables**

725 **Table S1.** List of specimens used in the phylogenomic analysis of the *Formicoxenus* genus-group
726 using whole mitochondrial genomes. The specimens represent the six genera currently recognized
727 in the group and the outgroup species.

728 **Table S2.** List of specimens employed in the phylogenetic analyses of *Formicoxenus* -
729 *Leptothorax* using the cytochrome COI-5P region (658 bp).

730 **Table S3.** Specimens of *L. acervorum* used in the phylogeographic and genetic diversity analyses
731 across its distribution range in the Holarctic region.

732 **Data S1.** Final alignment of the whole mitochondrial genomes of ant species used in the
733 phylogenomic analysis of the *Formicoxenus* genus-group.

734

735 **Supplementary Methods.** Methods for the whole mitochondrial genome assembly.

736

737 **Tables**

738 **Table 1.** Details of the sampling localities included in the phylogeographic analysis of *L.*
739 *acervorum* across its Holarctic distribution and the haplotypes observed in each population. N =
740 number of individuals in each locality, BC = British Columbia, NWT = Northwest Territories.

741

Country	Locality	N	Haplotypes
Spain	Cadí-Moixeró N.P.	1	H1
	Niela refuge	16	H14, H15, H18, H17, H19
	Valdelinares	14	H13, H12
	Orihuela de Tremedal	19	H19
	Aragon	1	H16
	Pla de la Font	12	H5
	Larra	12	H11, H8, H7
England	Santom Downham	7	H4, H11
Switzerland	Lausen, Wettingen	2	H4
Denmark	Nærum	1	H11
Norway	Østfold	1	H11
Sweden	Abisco	1	H11
Bulgaria	Panichishte	1	H9
Finland	Turku	1	H11
	Hikia	1	H11
	Lammi	1	H11
	Tvarminne	1	H11
Estonia	Liiva	1	H2
Russia	Kamchatka	2	H3
	Moscow	1	H10
Canada	Nearctic-West (Yukon, BC, NWT)	10	H11
	Nearctic-East (Manitoba)	5	H11

742

743

744

745

746

747

748

749

750

751

752

753 **Table 2.** Estimates of genetic diversity and neutrality tests obtained in the seven populations of *L.*
754 *acervorum* analyzed in detail.

755

Population	N	Ncol	Polymorphic sites	PI	No. haplotypes	<i>Hd</i>	π	Tajima's D	Fu's <i>F</i>
Spain									
Larra	12	7	5	3	6	0.803	0.00072	-0.0093	- 1.847
Niela refuge	16	13	16	15	6	0.817	0.00228	0.3364	2.118
Orihuela de Tremedal	19	11	8	0	2	0.105	0.00037	-2.1619	2.452
Pla de la Font	12	11	0	0	-	-	0	-	-
Valdelinares	14	12	1	1	2	0.363	0.00016	0.3244	0.643
England									
Santon Dunham	7	3	15	14	3	0.600	0.00340	1.1546	4.116

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

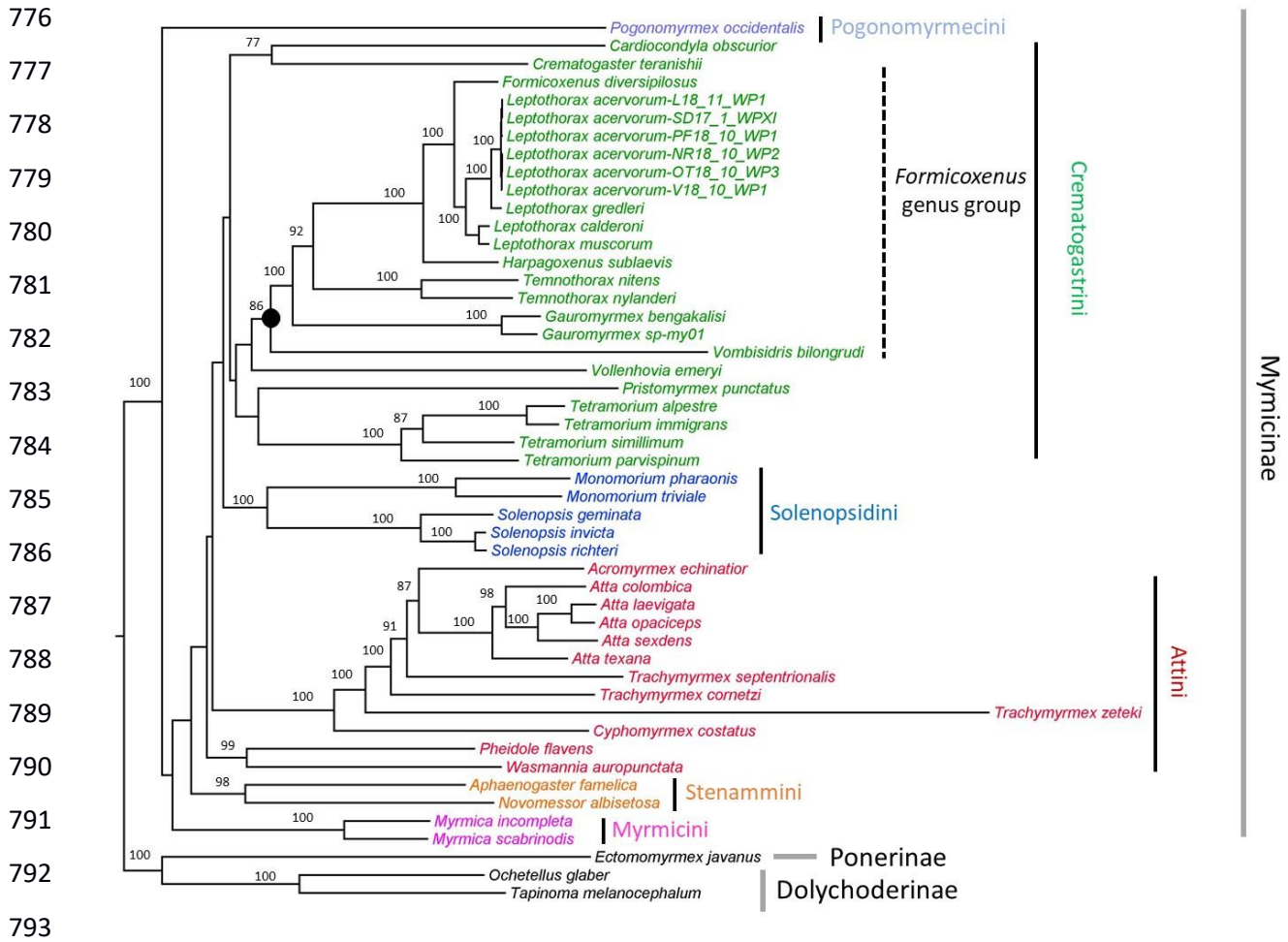
771

772

773

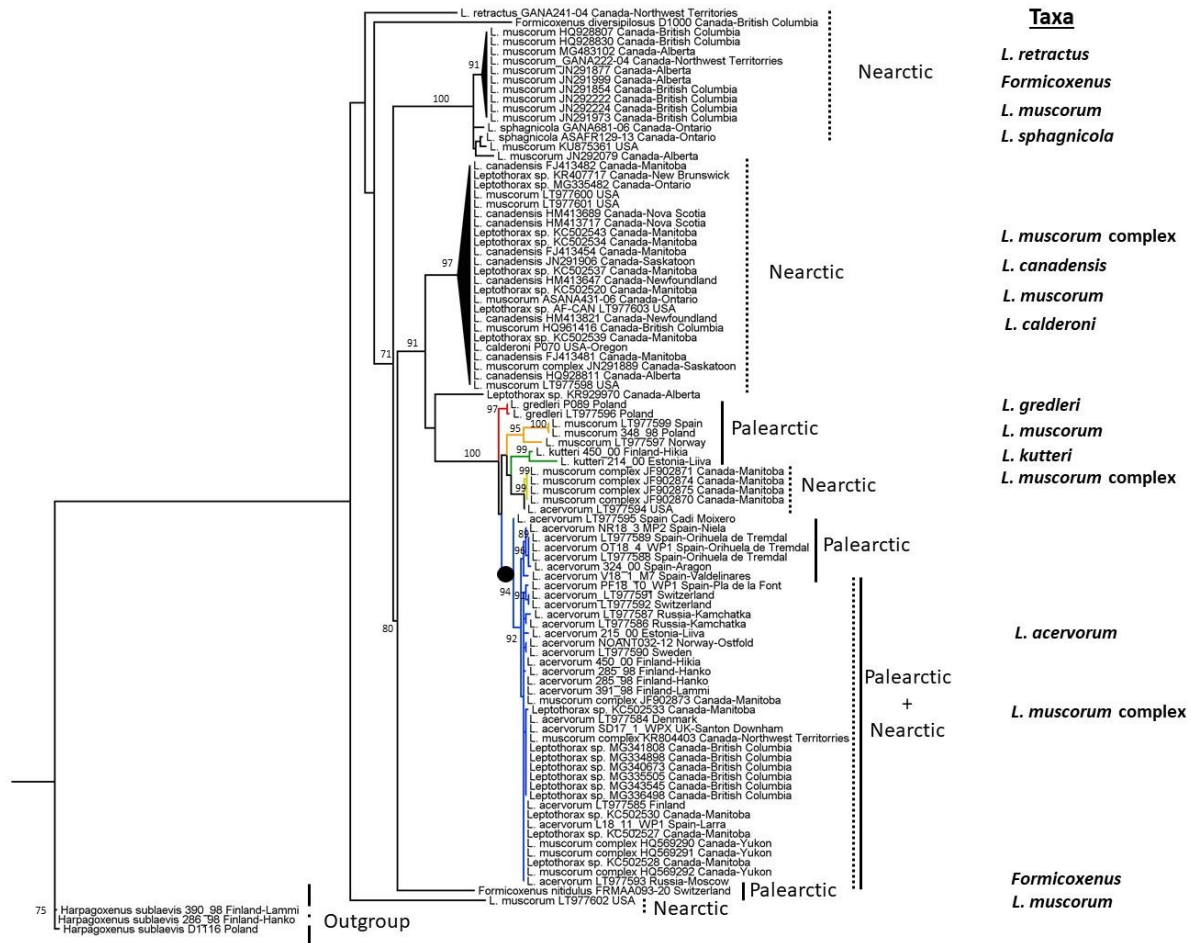
774

775



794 **Fig. 1.** Best tree obtained in the phylogenetic analysis inferred with whole mitochondrial genomes
 795 of the *Formicoxenus* genus-group using ML as implemented in IQ-TREE. Values next to the
 796 branches represent bootstrap support (1000 bs replicates). Only branches with support over 70 are
 797 labelled. The black circle indicates the node with the genera of the *Formicoxenus* genus-group.
 798 Different colors indicate tribes within the Myrmecinae.

799
800
801
802
803
804
805
806
807
808



809

810 **Fig. 2.** Best tree recovered of the analysis of *Formicoxenus-Leptothorax* using ML as implemented
 811 in IQ-TREE using the dataset of COI-5P region. The different colours of the branches indicate
 812 monophyletic lineages recovered on the species with Palearctic distribution. The black circle
 813 indicates the monophyletic lineage of *L. acervorum* (plus specimens of *L. muscorum* complex and
 814 *Leptothorax* sp.) that were later used in the phylogeographic analysis. Species currently recognized
 815 in these genera are indicated next to the lineages (taxa). Values next to the branches represent
 816 bootstrap support (1000 bs replicates). Only branches with >70 bootstrap support are labelled.

817

818

819

820

821

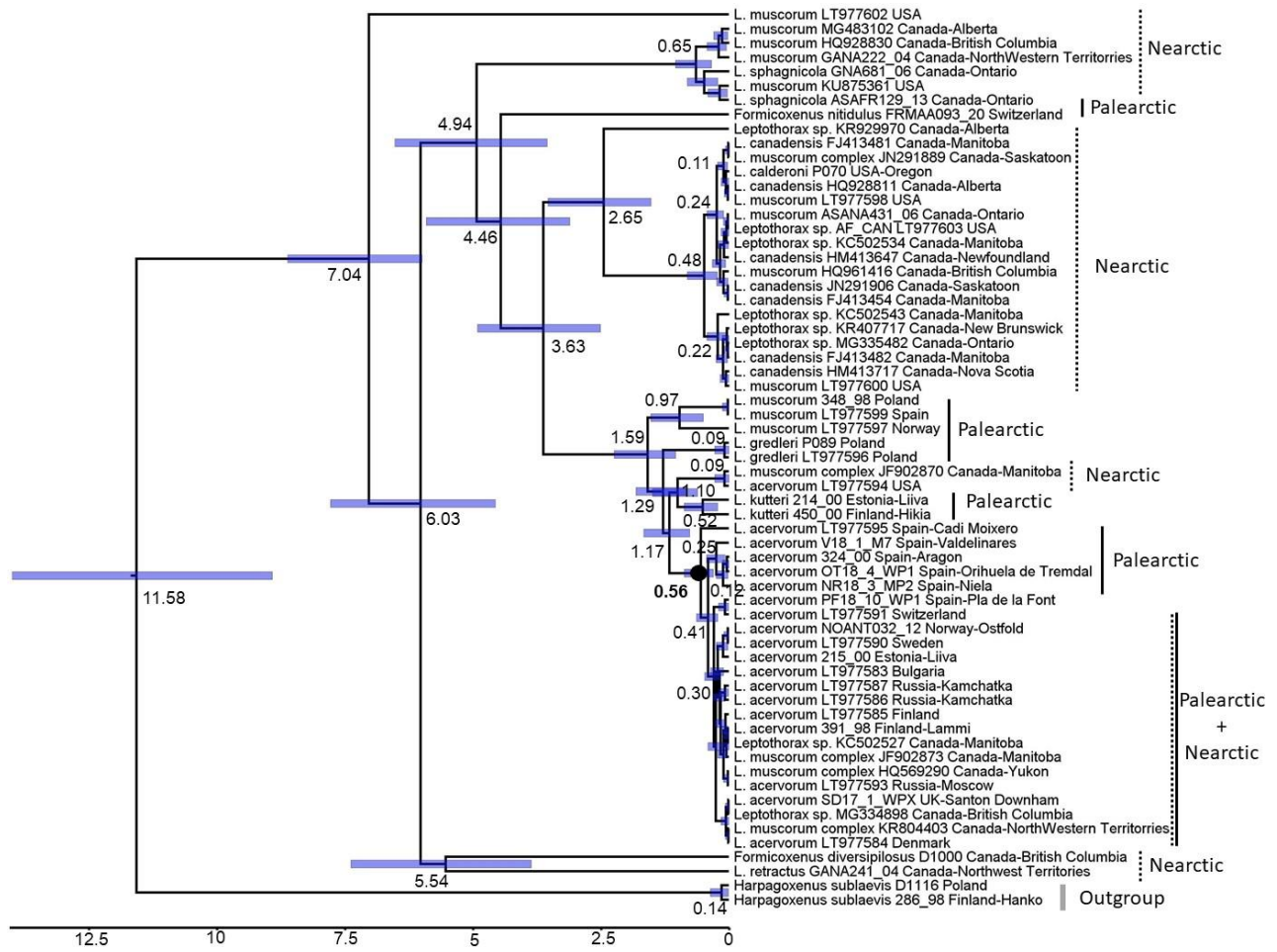
822

823

824

825

826



827

828

829 **Fig. 3.** Chronogram of the divergence times estimated in the *Formicoxenus* - *Leptothorax* lineages
 830 obtained with Beast. The black circle indicates the monophyletic lineage of *L. acervorum*. Values
 831 next to the branches indicate stem ages with blue columns displaying 95% confidence intervals.

832

833

834

835

836

837

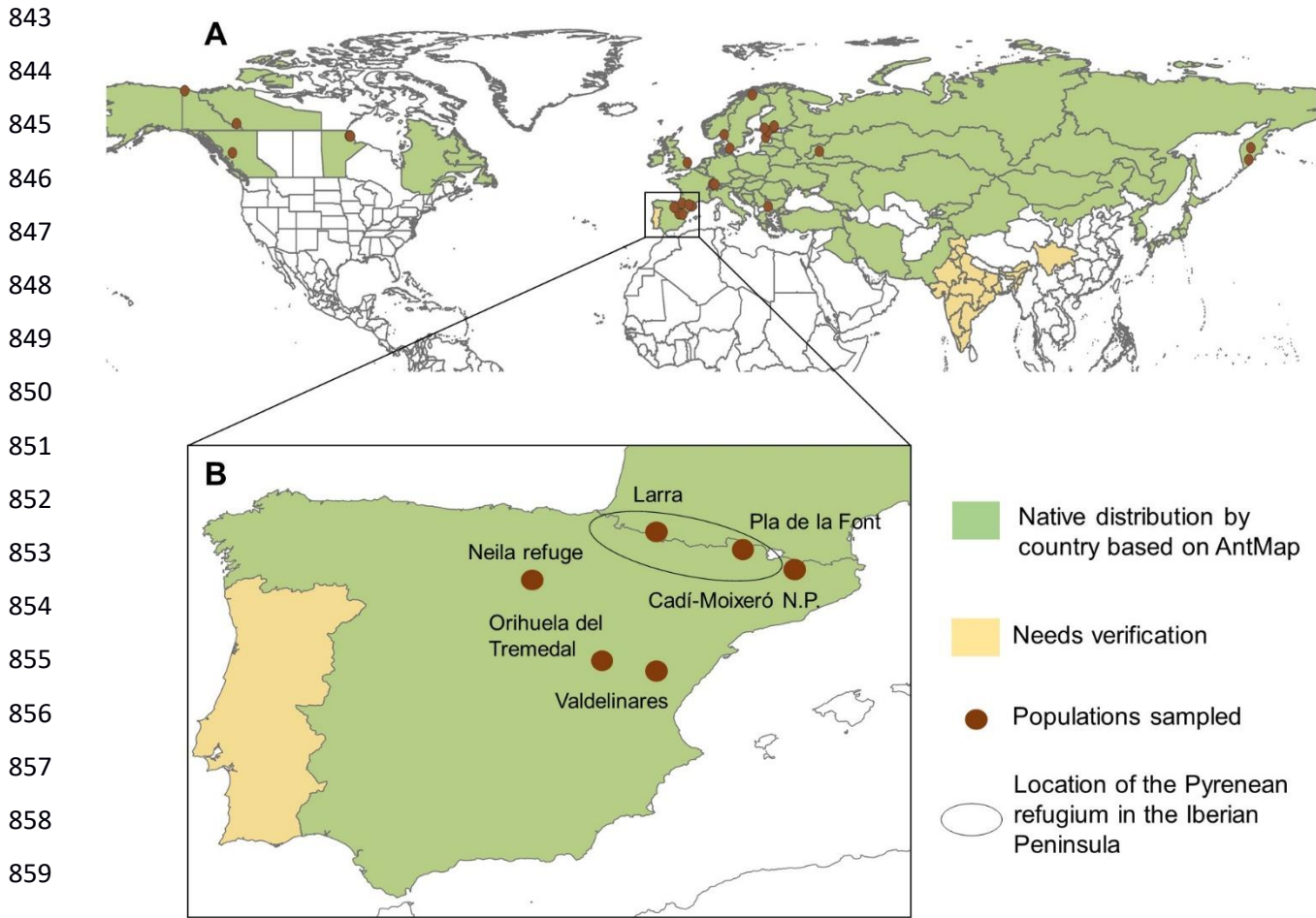
838

839

840

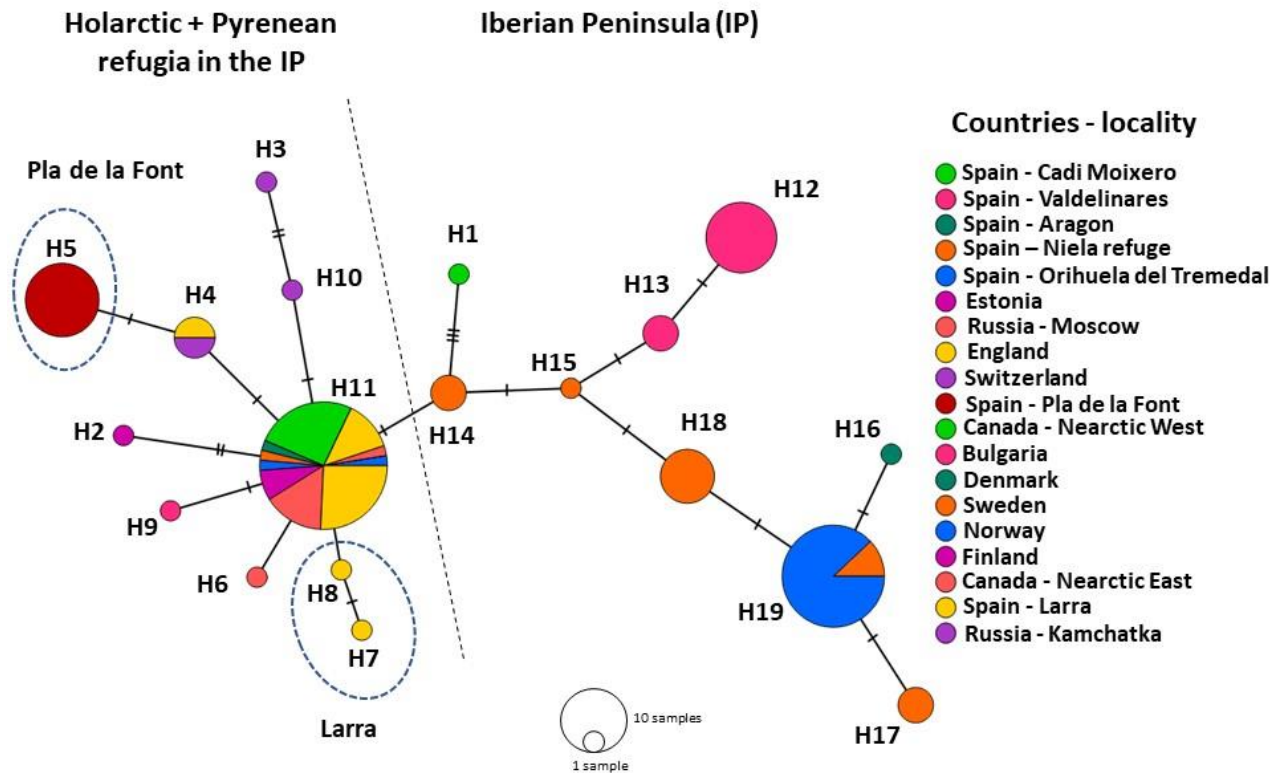
841

842



862 **Fig. 4.** A) Distribution map of *L. acervorum* across the Holarctic region based on the Global Ant
863 Biodiversity Informatics (GABI) database (Guénard et al., 2017). Locations of the different
864 populations included in the phylogeographic analysis are indicated with red dots. B) Close-up of
865 the populations located in the Iberian Peninsula, indicated with blue circles those populations that
866 contributed to the most recent expansion in the Holarctic. The black oval indicates the location of
867 the Pyrenean refugium (Tinaut and Ruano, 2021). N.P. = National Park.

868
869
870
871
872
873
874
875
876



877

878

879 **Fig. 5.** Haplotype network of *L. acervorum* across its Holarctic distribution range. Codes next to

880 the circle indicate the haplotype classification and their distribution. Hatch marks represent

881 mutation differences among the haplogroups. The dash line separates most of the Iberian Peninsula

882 populations from the rest of distribution. Pla de la Font and Larra (in dash circles) are both located

883 in the Pyrenean refugia.

884

885

886

887

888

889