

Taxonomic revision of the *Sylvarum* group of bumblebees using an integrative approach

Nicolas Brasero, Baptiste Martinet, Denis Michez, Thomas Lecocq, Irena Valterova, Pierre Rasmont

► **To cite this version:**

Nicolas Brasero, Baptiste Martinet, Denis Michez, Thomas Lecocq, Irena Valterova, et al.. Taxonomic revision of the *Sylvarum* group of bumblebees using an integrative approach. *Systematics and Biodiversity*, Taylor & Francis: STM, Behavioural Science and Public Health Titles, 2020, 18 (1), pp.12-28. 10.1080/14772000.2020.1737843 . hal-02906102

HAL Id: hal-02906102

<https://hal.univ-lorraine.fr/hal-02906102>

Submitted on 11 Dec 2020

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 Taxonomic revision of the *Sylvarum* group of bumblebees using an
2 integrative approach

3 Nicolas Brasero^{1*}, Baptiste Martinet¹, Denis Michez¹, Thomas Lecocq^{1,2}, Irena Valterova^{3,4}
4 and Pierre Rasmont¹

5 1: University of Mons, Research institute of Biosciences, Laboratory of Zoology, Place du
6 Parc 20, 7000 Mons, Belgium (brasero.nicolas@gmail.com)

7 2: University of Lorraine, INRA, URAFPA, F-54000 Nancy, France

8 3: Academy of Sciences of the Czech Republic, Institute of Organic Chemistry and
9 Biochemistry, Flemingovo nám 2, CZ-166 10 Prague, Czech Republic

10 4: Czech University of Life Sciences, Faculty of Tropical AgriSciences, Kamýcká 129, CZ-
11 165 21 Prague, Czech Republic

12

13 * corresponding authors

14 **Acknowledgement**

15 This work was supported by the Institute of Organic Chemistry and Biochemistry of the
16 Academy of Sciences of the Czech Republic under grant number 61388963; the FRS-FNRS
17 (Fonds de la Recherche Scientifique) under grant number 2013/V 3/5/307-IB/JN-11888 and
18 the European Community's Seventh Framework Program, STEP Project (Status and Trends
19 of European Pollinators, www.step-project .net, under grant number 244090, FP7/2007-2013).

20

21

22

23

24 **Abstract**

25 Recent improvements in taxonomy consider multiple operational criteria. Integrative
26 taxonomy provides a methodological framework merging these multisource approaches.
27 Bumblebees are considered as a uniform group where their taxonomy remains one of the most
28 difficult. Here, we investigate the taxonomic statuses inside a monophyletic group including
29 six taxa (*B. inexpectatus*, *B. mlokosievitzii*, *B. ruderarius*, *B. sylvarum*, *B. velox* and *B.*
30 *veteranus*) in the most diverse subgenus of bumblebees: *Thoracobombus*. We used an
31 integrative approach based on mitochondrial and nuclear genetic markers and eco-chemical
32 traits commonly used in bumblebee taxonomy. For all species, our study shows a clear
33 differentiation in DNA and eco-chemical traits. However, we conserve the subspecies status
34 of *B. ruderarius simulatilis* and *B. sylvarum daghestanicus* from the east of Turkey and Iran
35 according to their degree of genetic and eco-chemical differentiation.

36 **Keywords:** Bumblebees, eco-chemical traits, integrative taxonomy, mitochondrial genetic
37 maker, nuclear genetic maker, *Thoracobombus*.

38 **Introduction**

39 Species are considered as the fundamental unit in biology and their accurate delimitation is
40 essential respectively in biodiversity classification and in evolutionary and conservation
41 biology (Moraes, 1987; Paquin, Dupérré, Cokendolpher, White & Hedin, 2008; Wheeler,
42 Raven & Wilson, 2004). Taxonomy inside invertebrate groups is mainly based on
43 morphological approach. However, the common strong homogeneity in morphology may be
44 an obstacle for identification, especially in the case of cryptic species (Williams et al., 2012).
45 Moreover, many species exhibit a high intra-specific variability and cases of convergent color
46 pattern were recorded in sympatry (Françoso, Freitas de Oliveira & Arias, 2016; Lecocq et al.,
47 2015a). Therefore, many alternative features and methods such as DNA-based approaches

48 (e.g. pairwise distance thresholds, bootstrap support or reciprocal monophyly), geometric
49 morphometrics of wing shape or semiochemical traits have been used to delineate species
50 (Lecocq et al., 2015b; Martin, Helantera & Drijfhout, 2008; Oleksa & Tofilski, 2015; White,
51 Pilgrim, Boykin, Stein & Mazor, 2014). In order to draw a strongly supported taxonomic
52 hypothesis, recent betterment in taxonomy consider multiple operational criteria from the
53 same sampling (Gibbs, 2009; Lecocq et al., 2011; Martinet et al., 2018). The integrative
54 taxonomy based on the Unified Species Concept (USC) provides a methodological framework
55 merging these multisource approaches (De Queiroz, 2007; Padial, Miralles, De la Riva &
56 Vences, 2010; Schlick-Steiner et al., 2010). The USC argues that all species concepts agree
57 that species exist as separately evolving metapopulation lineages but diverge in their criterion
58 for delimiting species (De Queiroz, 2007). Therefore, ecological differentiation, phylogenetic
59 divergence, morphological characteristic or reproductive isolation could highlight a
60 metapopulation lineage separation by avoiding false taxonomic conclusions driven by a single
61 trait. Moreover, by considering subspecies as a step in the process of allopatric speciation
62 (Mayr, 1942), and because of its important value in conservation and evolutionary biology
63 (Coppée, Terzo, Valterová & Rasmont, 2008; Phillimore & Owens, 2006; Rasmont, Coppée,
64 Michez & De Meulemeester, 2008; Sackett et al., 2014; Terzo, Urbanová, Valterová &
65 Rasmont, 2005), subspecies status has been proposed as solution when differentiation is
66 highlighted by a subset of operational criteria. This proposal aims to reduce the under-
67 estimate risk of a congruence approach (Hawlitschek, Nagy & Glaw, 2012; Padial et al.,
68 2010).

69 Bumblebees are considered as a complex group where their classification remains one of the
70 most difficult in bees (Moure & Sakagami, 1962; Murray, Fitzpatrick, Brown & Paxton,
71 2008). In this group, integrative taxonomy has considerably improved taxonomic studies and
72 has solved many doubtful cases thanks to operational criteria as DNA (mitochondrial and

73 nuclear) and semio-chemicals (Bertsch, Schweer, Titze & Tanaka, 2005; Lecocq et al., 2015a;
74 Martinet et al. 2018).

75 In this paper, we apply an integrative approach to assess the taxonomic status of a group of
76 phylogenetically closely related species in the *Thoracobombus* Dalla Torre, 1880 subgenus:
77 *B. mlokosievitzii* Radoszkowski, 1877, *B. ruderarius* (Müller, 1776), *B. sylvarum* (L., 1761),
78 *B. velox* (Skorikov, 1914), *B. veteranus* (Fabricius, 1793) and *Bombus inexpectatus* (Tkalcù,
79 1963). We revise the taxonomy of this group through a decision framework based on
80 differentiation of three operational and independent criteria commonly used in bumblebee
81 taxonomy: a mtDNA marker (COI), a nuDNA marker (PEPCK) and an eco-chemical trait.

82

83 **Materials and methods**

84 *Studied taxa*

85 *B. ruderarius* (Müller, 1776) is a widespread species distributed from Tunisia in the south to
86 the Arctic Circle in the north (Rasmont et al., 2015a). Four subspecies and one form have
87 been described (Rasmont, 1983): (i) *B. ruderarius ruderarius* (Müller, 1776) is characterized
88 by a reddish tail and the rest of the body is black intermixed with more or less greyish hairs
89 (Fig. 1C). They occurred throughout Europe except the Mediterranean region; (ii) *B.*
90 *ruderarius simulatilis* Radoszkowski, 1888 with a wide white collare and scutellare bordering
91 a black interalar band, the tergite 1 and 2 completely white and with the tergite 3 black, the
92 tail is red (Fig. 1D). This subspecies occurred in mountainous regions of Anatolia as well as
93 in the Caucasus region in north Iran; (iii) *B. ruderarius tunensis* (Tkalcù, 1974) is
94 characterized by three yellow bands and a white tail and is endemic to the east of Algeria and
95 north of Tunisia; (iv) *B. ruderarius sartus* (Skorikov, 1922) with a wide yellowish collare and
96 scutellare bordering a black interalar band, the tergite 1 and 2 completely yellowish and with
97 the tergite 3 black, the tail is red. The eastern limits of the species were known to be the

98 mountains of east Turkestan but other specimens were found in Mongolia (Skorikov, 1922);
99 (v) *B. ruderarius* form *montanus* is characterized by large yellowish bands compared to the
100 dark nominal subspecies (Fig. 1C). They occur from the Cantabrian Mountains to the
101 Pyrenees where they coexist with the nominal subspecies *B. ruderarius ruderarius*.
102 *B. sylvarum* (L., 1761) is also a widespread species distributed in the west-palaeartic region
103 but hardly reaches Mongolia (Rasmont et al., 2015a). We considered four subspecies and one
104 form (Rasmont, 1983): (i) *B. sylvarum sylvarum* (L., 1761) with three large greyish bands
105 (thorax and abdomen) and thin grey bands (abdomen) and a reddish tail intermixed with grey
106 (Fig. 1F). They occur in the west-Palaeartic region; (ii) *B. sylvarum citrinofasciatus* Vogt,
107 1909 is characterized by grey yellowish bands compared to the nominal subspecies.
108 Moreover, the thin grey bands on the abdomen are often absent. They occur in the west of
109 Turkey; (iii) *B. sylvarum daghestanicus* Radoszkowski, 1877 is characterized by white bands
110 and a reddish tail (Fig. 1E). This subspecies is distributed in the east of Turkey, the Caucasus
111 region as well as in Iran; (iv) *B. sylvarum rogenhoferi* Dalla Torre, 1882 with white yellowish
112 bands, a red tail and without thin grey bands on the abdomen is restricted to south Italy and
113 Sicily; (v) *B. sylvarum* form *nigrescens* Pérez, 1879 is considered as the melanic form inside
114 *B. sylvarum* species. They are completely black with a reddish tail and with or without
115 intermixed grey hairs on the thorax (Fig. 1G). They occur in the south-west of France, the
116 Atlantic regions as well as the south of Scandinavia. Currently, all these subspecies and form
117 are considered as *B. sylvarum* (Rasmont et al., 2015a; Williams, 1998).
118 All other species included in this paper form a monophyletic group with *B. ruderarius* and *B.*
119 *sylvarum* (Cameron, Hines & Williams, 2007) and will be treated as comparison group.
120 *Bombus inexpectatus* (Tkalčů, 1963) (Fig. 1A) is a sub-alpine uncommon European species
121 restricted in the Cantabrian Mountains and the Alps from 1200m to 2100m in altitude (Amiet,
122 1996; Rasmont et al., 2015a; Tkalčů, 1963, 1965; Yarrow, 1970). It has been described as a

123 species adapted to parasitic life (Müller, 2006; Yarrow, 1970) with *B. ruderarius* as its only
124 host (Müller, 2006). There is no doubt about its specific status.

125 *B. mlokosievitzii* Radoszkowski, 1877 is a ponto-mediterranean species whose distribution
126 extends from the Balkan Mountains to Caucasian region via the south of the Black Sea
127 (Rasmont & Iserbyt, 2014). Three subspecies diagnosable by their coat color variation has
128 been defined: (i) *B. mlokosievitzii mlokosievitzii* Radoszkowski, 1877 with a wide greyish
129 white collare and scutellare bordering a black interalar band. This subspecies is distributed in
130 the north of Anatolia; (ii) *B. mlokosievitzii convergens* Skorikov, 1908 characterized by a
131 thorax completely greyish white is distributed in the Caucasian region and (iii) *B.*
132 *mlokosievitzii vogtiellus* diagnosable by their thorax completely black (Fig. 1B) and
133 distributed in Greece, Macedonia, Bulgaria and west of Anatolia. The presence of a reddish
134 band on tergite 2 is common to all subspecies. Unfortunately we were not able to collect all
135 subspecies inside this taxon and only the subspecies *mlokosievitzii* was considered in this
136 study.

137 *B. velox* (Skorikov, 1914) is restricted to the north east of Anatolia and the Caucasus region. It
138 is an uncommon species with two greyish bands on the thorax and one on the tergite 1, the
139 rest of the abdomen is yellowish (Fig. 1H). There is no doubt about its specific status.

140 *B. veteranus* (Fabricius, 1793) lives in sparse population in the plains of northern Europe
141 (Rasmont & Iserbyt, 2014). This species presents very little color variation and is characterize
142 by a yellowish coat color on all the body except the presence of a black interalar band (Fig.
143 1I). There is no doubt about its specific status.

144 *Sampling*

145 150 male and female specimens were collected between 2001 and 2017 in the Palaearctic
146 region (supplementary file 1, Tab. 1). All specimens were determined based on their
147 morphology. We attributed a name to the specimens without taxonomic a priori on the species

148 status. The ingroup comprises: *runderarius* from Estonia (n=5), France (n=16), Macedonia
149 (n=1), Serbia (n=1), Slovakia (n=4), Sweden (n=1) and Switzerland (n=5), *montanus* from
150 Andorra (n=3), Spain (n=2) and France (n=5), *simulatilis* from Iran (n=3) and Turkey (n=9),
151 *sartus* from Kyrgyzstan (n=2) and Mongolia (n=2), *sylvarum* from Belgium (n=3), France
152 (=13), Germany (n=4), Italy (n=3), Poland (n=6), Serbia (n=2), Slovakia (n=1), Spain (n=2)
153 and Sweden (n=4), *daghestanicus* from Iran (n=1) and Turkey (n=8), *nigrescens* from Spain
154 (n=1) and Sweden (n=5), *rogenhoferi* from Italy (n=2). The comparison group comprises:
155 *inexpectatus* from Italy (n=9); *mlokosievitzii* from Turkey (n=11), *velox* from Turkey (n=1)
156 and *veteranus* from Belgium (n=4), France (n=4) and the Netherlands (n=4). The
157 phylogenetically related species *B. mucidus* (n=3) was used as outgroup.

158 *Genetic trait analyses*

159 Two genes commonly used in bumblebee phylogeny were sequenced (Cameron et al., 2007;
160 Martinet et al., 2018; Williams et al., 2012): the "barcode" region of the cytochrome oxidase 1
161 (COI) mitochondrial gene was sequenced from 88 specimens, and the nuclear protein-coding
162 gene phosphoenolpyruvate carboxykinase (PEPCK) was sequenced from 83 specimens
163 (supplementary file 1).

164 We carried out polymerase chain reaction amplifications with new design primer pairs
165 CSylG-F/CSylG-R: (5'-TGATCAGGAATAATTGGATCTTCA-3') / (5'-
166 TGATCAGGAATAATTGGATCTTCA-3') for COI and PSylG-F/PSylG-R: (5'-
167 CTTTAAAGGCAATTGTATCAAAGTT-3') / (5'-CTCGCCAATCGGTGATCTAT-3') for
168 PEPCK. We performed sequencing procedures described in Lecocq et al. (2013). Sequences
169 were aligned with BioEdit version 7.2.5 (Hall, 1999). The resulting COI (624bp) and PEPCK
170 (894bp) sequences are available on GenBank (accession numbers in supplementary file 1).

171 We investigated the potential genetic differentiation for each gene individually within the six
172 taxa through haplotype networks analyses. We used the median-joining method with Network

173 4.6.1.1 software (www.fluxus-engineering.com) to produce haplotype networks. We weighted
174 transversions twice as high as transitions to reconstruct the network (Lecocq et al., 2015a;
175 Martinet et al., 2018). In phylogenetic analyses, gaps (deletion or insertion) are usually treated
176 as missing data. Because of the presence of gaps sites in our PEPCK genes, two haplotype
177 networks considering or not the gaps sites will be performed in order to be aware of
178 phylogenetic information provide by these two methods.

179 In phylogenetic analyses, we analyzed each gene independently with maximum likelihood
180 (ML) and Bayesian (MB) methods to explore the genetic divergence and define lineages
181 within the different groups. For ML methods, we partitioned each gene to explore the best
182 substitution model: i) PEPCK into two introns and one exon; ii) COI and PEPCK exons by
183 base position (1st, 2nd and 3rd) and choose the best fitting substitution models with JModelTest
184 Server 2.0 (Posada, 2008) using the Akaike information criterion corrected for small sample
185 sizes (Hurvich & Tsai, 1989). The chosen models were: i) For COI: TIM1 + I (1st), HKY
186 (2nd), and TPM3uf + I (3rd); ii) for PEPCK, exon: F81 (1st), JC (2nd), K80 (3rd); intron 1:
187 TMP3uf + I; intron 2: HKY. Selected models which are not implemented in MrBayes were
188 substituted by the closest over-parameterized model. For each gene, ML analyses were
189 performed with ten independent runs in Garli 2.0 (Zwickl, 2006); the topology and -ln L was
190 the same among replicates. The run with the highest likelihood was retained. Statistical
191 significance in nodes was evaluated with 10000 non-parametric bootstrap replicates.
192 Topologies with bootstrap values $\geq 85\%$ were considered as well supported (Hillis & Bull,
193 1993). For MB method, we performed Bayesian inference analyses with mrbayes 3.1.2
194 (Ronquist & Huelsenbeck, 2003). Ten independent analyses were achieved for each gene (100
195 million generations, four chains with mixed models, default priors, saving trees every 1000
196 generations). Then the first 10 million generations were removed as burn-in procedure. A
197 majority-rule 50% consensus tree was constructed. We considered as statistically significant

198 only branch supports (topologies) with high posterior probabilities (≥ 0.95) (Wilcox, Zwickl,
199 Heath & Hillis, 2002). The phylogenetic analysis based on PEPCK gene fragment considered
200 gaps as missing data. All trees were rooted with the taxon *B. mucidus*.

201 In order to recognize species threshold, we applied the Bayesian implementation of the
202 general mixed Yule-coalescent (bGMYC) method to the COI dataset by using R Package
203 "bGMYC" (Reid & Carstens, 2012). A range of probabilities < 0.05 was considered as strong
204 evidence that taxa was heterospecific while a range of probabilities > 0.95 suggested that taxa
205 was conspecific (Reid & Carstens 2012). BEAST 1.7.4 (Drummond, Suchard, Xie &
206 Ranbaut, 2012) was used to generate ultrametric trees with a phylogenetic clock model in
207 order to calculate a posterior distribution of trees (length of the MCMC chain: 1 billion
208 generations). The first million sampled trees were burned-in using the maximum clade
209 credibility method and setting the posterior probability limit to 0. The bGMYC analysis was
210 based on 1000 trees sampled every 10 000 generations. For each of these 1000 trees, the
211 MCMC was made of 100 000 generations, discarding the first 90 000 as burn-in and sampling
212 every 100 generations.

213 *Eco-chemical trait analyses*

214 We focused on the cephalic labial gland secretions (CLGS) of males, the most commonly
215 reproductive trait involved in the bumblebee pre-mating recognition (Ayasse & Jarau, 2014).
216 Currently, the composition of the CLGS is known for 15 species of *Thoracobombus* (Ayasse
217 & Jarau, 2014; Brasero et al., 2015; Brasero et al., 2018a,b) and this study will increase the
218 knowledge inside this group to 17 species. These secretions are species-specific (Calam,
219 1969) and provide an efficient diagnostic characters for species delimitation (Brasero et al.,
220 2015). They are complex mixtures of mainly aliphatic compounds synthesized *de novo*
221 (Coppée et al., 2008; Žáček et al., 2013) in the head of bumblebee males. Main components
222 were defined as those with the highest relative amount in at least one individual of the taxon.

223 All specimens were killed by freezing at -20°C and the CLGS were extracted with 400 µl of
224 heptane (method described in De Meulemeester, Gerbaux, Boulvin, Coppée, & Rasmont,
225 2011). Samples were stored at -40°C prior to analysis. We were able the sample 111
226 specimens belonging to the six studied taxa (Tab. 1).

227 We determined the CLGS composition by gas chromatography-mass spectrometry (GC/MS)
228 using a Focus GC (Thermo Scientific) with a non-polar DB-5 ms capillary column [5%
229 phenyl (methyl) polysiloxane stationary phase; column length 30 m; inner diameter 0.25 mm;
230 film thickness 0.25 µm] coupled to DSQ II quadrupole mass analyzer (Thermo Scientific)
231 with 70 eV electron impact ionization. We used a splitless injection mode (220°C) and
232 helium as a carrier gas (one ml/min). The temperature program of the oven was set to 70°C
233 for two minutes and then heated up at a rate of 10°C/min to 320°C. The temperature was then
234 held at 320°C for five minutes. Compounds were identified in Xcalibur™ using the retention
235 times (t_r) and mass spectra of each peak, in comparison to those at National Institute of
236 Standards and Technology library (NIST, U.S.A). Double bond positions (C=C) were
237 determined by dimethyl disulfide (DMDS) derivatization (Cvacka et al., 2008).

238 The CLGS compounds of all samples were quantified by a gas chromatograph-flame
239 ionization detector Shimadzu GC-2010 with a SLB-5ms non-polar capillary column (5%
240 phenyl (methyl) polysiloxane stationary phase; 30-m column length; 0.25-mm inner diameter;
241 0.25-µm film thickness) with the same chromatographic conditions as in GC/MS. Peak areas
242 of compounds were detected in GCsolution Postrun (Shimadzu Corporation) with automatic
243 peak detection and noise measurement. We calculated relative amounts (RA in %) of
244 compounds in each sample by dividing the peak areas of compounds by the total area of
245 compounds. We discarded all compounds for which RA were recorded as less than 0.1% for
246 all specimens (De Meulemeester et al., 2011). The data matrix (supplementary file 2) for each

247 taxon was based on the alignment of each relative proportion of compound between all
248 samples performed with GCAAligner 1.0 (Dellicour & Lecocq, 2013).

249 Before each samples injection, a standard (Kovats) containing a mix of hydrocarbons
250 (alkanes) from C10 (decane) to C40 (tetracontane) was injected in order to facilitate the
251 alignment and the identification of compounds. Kovats indices were calculated with GC
252 Kovats 1.0 (Dellicour & Lecocq, 2013).

253 Clustering method were performed using R 3.3.2 (R Development Core Team, 2016) to
254 detected CLGS differentiations between taxa. A Canberra distance matrix based on CLGS
255 (RA of each compound) was computed. An unweighted pair group method with arithmetic
256 mean (UPGMA) was used as clustering method (R-package ape, Suzuki & Shimodaira,
257 2011). We assessed the uncertainty in hierarchical cluster analysis using p -values calculated
258 via multiscale bootstrap resampling with 50 000 bootstrap replications (significant branch
259 support > 0.85) (R-package pvclust, Suzuki & Shimodaira, 2011). We assessed CLGS
260 differentiations between taxa by performing a permutation multivariate analysis of variance
261 using distance matrix (perMANOVA) (R package vegan; Oksanen et al., 2011). When a
262 significant difference was detected, we performed a pairwise multiple comparison with an
263 adjustment of P -values (Bonferroni correction) to avoid type I errors.

264 To determine the indicator compounds (IC) of each taxon, we used the indicator value
265 (IndVal) method (Claudet, Pelletier, Jouvenel, Bachet & Galzin, 2006; Dufrêne & Legendre,
266 1997). The value given is the product of relative abundance and relative frequency of
267 occurrence of a compound within a group. We evaluated the statistical significance of a
268 compound as an indicator at the 0.01 level with a randomization procedure.

269 *Decision framework of taxonomic status*

270 We used an approach that corresponds to the strictest commonly used framework called
271 "integration by congruence" by Padial et al. (2010). We thus assigned the specific status to a

272 taxon (i) genetically differentiated in all genetic markers (i.e. unique haplotypes), (ii) which
273 constitute a monophyletic group with high branch support, and (iii) with a significant
274 differentiation in CLGS composition (significance differentiation in perMANOVA and high
275 bootstrap values > 0.85). All criteria used must be convergent to assign specific status. This
276 approach can lead to underestimation of species differentiation but reduces the taxonomic
277 inflation. We assigned the subspecies status to phenotypically distinct allopatric population
278 that does not diverge in all lines of evidence. This decision reduces the ‘underestimate’s risk’
279 of our strict integrative approach by naming them as a subspecies (Ennen et al., 2014;
280 Hawlitschek et al., 2012).

281

282 **Results**

283 *Genetic trait analyses*

284 All phylogenetic analyses (ML and MB) showed the expected differentiation between
285 outgroup and ingroup (Fig. 2A, 3A). Seven strongly supported (bootstrap>0.95) monophyletic
286 groups are detected in our COI phylogenetic analyses (Fig. 2A): (i) one group with
287 *inexpectatus* specimens; (ii) one group with *runderarius*, *montanus*, *sartus* and *simulatilis*
288 specimens; (iii) one with *velox* specimens; (iv) one group with *veteranus* specimens; (v) one
289 group including all *sylvarum*, *nigrescens*, *daghestanicus* and *rogenhoferi* specimens; (vi) one
290 group with *mlokosievitzii* specimens and (vii) the outgroup *B. mucidus*. Within "*runderarius*
291 group", two unexpected lineages were revealed (Fig. 2A): one main lineage strongly
292 supported including the widespread *runderarius* taxa, *montanus* and *sartus* and one another
293 lineage not strongly supported (bootstrap<0.95) including *simulatilis* taxon from Turkey and
294 Iran. As expected, we found a less structured tree in the PEPCK gene fragment for all taxa.
295 PEPCK gene analyses showed little phylogenetic information and failed to reconstruct the
296 phylogenetic affinities between all taxa (Fig. 3A). However, we found some strongly

297 supported groups as *inexpectatus* and *mlokosievitzii*. As for COI gene, no geographical
298 entities were underlined by our PEPCK phylogenetic analyses expected for "*runderarius*
299 group" where the unexpected lineage *simulatilis* was revealed and strongly supported in our
300 MB (Fig. 3A) but not in our ML analysis.

301 15 private haplotypes were highlighted for COI gene while 11 and 10 haplotypes were
302 highlighted in PEPCK genes respectively considering or not the gaps sites by our network
303 analysis (Fig. 2B, 3B, C). Within the "*runderarius* and *sylvarum* groups", only *simulatilis*
304 displayed specific haplotypes in COI and PEPCK, while *daghestanicus* had specific COI
305 haplotype (Fig. 2B, 3B, C). Considering all taxa, COI revealed 72/624 nucleotide sites
306 substitution corresponding to 11.54% of phylogenetic information. Compared to the taxa
307 inside their groups, *simulatilis* and *daghestanicus* revealed respectively 0.96% (6/624
308 diagnostic nucleotide sites) and 0.64% (4/624 diagnostic nucleotide sites) of phylogenetic
309 information in COI gene (Fig. 2B). Considering all taxa, PEPCK gene revealed 143/894
310 (considering gaps) and 23/894 (not considering gaps) nucleotide sites respectively
311 corresponding to 16% and 2.6% of phylogenetic information. Compared to the taxa inside
312 their groups, *simulatilis* revealed 1.79% (16/894 diagnostic nucleotide sites) of phylogenetic
313 information when gaps are considered (Fig. 3C) and 0.34% (3/894 diagnostic nucleotide sites)
314 when gaps are not considered (Fig. 3B).

315 The tree generated by bGMYC analysis highlighted several entities with low probabilities
316 (<0.05) to be conspecific with the other ones (Fig. 4). It suggested the delimitation of 7
317 prospective species (P<0.05) (Fig. 4) which match with those obtained in the COI tree (ML
318 and MB, Fig. 2A): (i); one group with *inexpectatus* specimens (P=1); (ii); one group with
319 *velox* specimens (P=1); (iii) one group with *runderarius*, *sartus*, *montanus* and *simulatilis*
320 specimens (P>0.16-1); (iv) one group with *veteranus* specimens (P=1); (v) one group
321 including *sylvarum*, *daghestanicus*, *nigrescens*, *rogenhoferi* specimens (P>0.34-1); (vi) one

322 group with *mlokosievitzii* specimens (P=1) and (vii) one group with the outgroup *B. mucidus*
323 (P=1). The pairwise matrix shows a non significant heterospecificity threshold between
324 *simulatilis* and the other taxa from *runderarius* group (P>0.05).

325 *Eco-chemical trait analyses*

326 The chemical analyses highlighted 7 strongly supported groups (bootstrap > 85%) (Fig. 5). **A:**
327 *velox* (Turkey); **B:** *inexpectatus* (Italy); **C:** *mlokosievitzii* (Turkey), **D:** *sylvarum* +
328 *daghestanicus* + *nigrescens* + *rogenhoferi* (Belgium, France, Germany, Iran, Italy, Poland,
329 Serbia, Slovakia, Spain, Sweden and Turkey); **E:** *veteranus* (Belgium, France and Holland);
330 **F:** *montanus* + *runderarius* (Andorra, Estonia, France, Macedonia, Serbia, Slovakia, Spain,
331 Sweden and Switzerland) and **G:** *simulatilis* (Turkey). 95 compounds were detected
332 (respectively 17 from **A**, 36 from **B**, 34 from **C**, 55 from **D**, 36 from **E**, 49 from **F** and 44
333 from **G**; see supplementary file 2). For each group, we detected main compounds, **A:**
334 octadecen-1-ol (88.2%); **B:** hexadec-7-en-1-ol (17.2-57.8%), octadec-11-en-1-ol (12.2-32.6%)
335 and octadec-11-enyl acetate (10-23.4%); **C:** hexadec-7-en-1-ol (14-46.7%) and hexadec-7-
336 enyl acetate (19.6-45.9%); **D:** hexadec-7-en-1-ol (2.2-57.5%) and hexadec-7-enyl acetate
337 (20.1-66.6%); **E:** hexadec-9-en-1-ol (1.1-57.8%) and ethyl octadec-11-enoate (1.5-31.4%); **F:**
338 hexadec-9-en-1-ol (0-62.3%), hexadec-9-enyl acetate (0-32.7%), tricos-9-ene + tricos-7-ene
339 (2.1-59.2%) and hexadecenyl hexadecenoate2 (0.1-25.1%) and **G:** hexadec-9-en-1-ol (4.5-
340 59.9%), octadecadien-1-ol (0.4-22%), and tricos-9-ene + tricos-7-ene (5.7-54.2%) and
341 tricosane (2-20%). Moreover, the IndVal method highlighted several significantly indicator
342 compounds (IC) (IndVal value>70) for the group **A** (IC=4), **B** (IC=11), **C** (IC=3), **D** (IC=7),
343 **E** (IC= 5), **F** (IC=1) and **G** (IC=0) (supplementary file 2). Despite the separation we observed
344 between *montanus* + *runderarius* and *simulatilis* our IndVal method did not find indicator
345 compounds for *simulatilis*, however, our statistical analysis significantly supports this
346 differentiation (bootstrap > 85%). perMANOVA test confirmed the differentiations between

347 all groups (see supplementary file 3) and particularly between *montanus* + *runderarius* and
348 *simulatilis* (p-value: 0.041; F: 2.9986).

349 *Decision framework of taxonomic status*

350 Within the "*runderarius* group" (*runderarius*, *sartus*, *montanus* and *simulatilis*), operational
351 criteria did not reflect the current subspecific differentiation (Tab. 2). CLGS and haplotype
352 networks analyses detected differentiation between *simulatilis* and the other taxa. We then
353 considered *B. runderarius simulatilis* as a subspecies inside the "*runderarius* group". This
354 subspecies occurs in mountainous regions of Anatolia, the Caucasus region as well in north
355 Iran. We also assigned the subspecific status *B. runderarius runderarius* (Müller, 1776) to the
356 other taxa (Tab. 2).

357 The genetic structure observed in COI, PEPCK and CLGS did not reflect the current
358 subspecific status encountered in the "*sylvarum* group" (*sylvarum*, *daghestanicus*, *nigrescens*
359 and *rogenhoferi*). Based on its unique haplotype in COI we decided to consider *B. sylvarum*
360 *daghestanicus* as a subspecies inside the "*sylvarum* group" as described by Radoszkowski
361 (1877). This subspecies occurred in mountainous regions of Anatolia, the Caucasus region as
362 well in Iran. We then assigned the subspecific status *B. sylvarum sylvarum* (Müller, 1776) to
363 the other taxa inside which we considered the forms *sylvarum*, *nigrescens* and *rogenhoferi*
364 based on their color patterns (Tab. 2). The specific status of *B. inexpectatus*, *B.*
365 *mlokosievitzii*, *B. velox* and *B. veteranus* were confirmed by our results.

366

367 **Discussion**

368 Genetic character along with CLGS can confirm with a high degree of certainty the speciation
369 processes between taxa. We established the species status of our different taxa according to
370 their divergences in genetic markers (mitochondrial and nuclear) and their CLGS
371 differentiation (Tab. 2).

372 Our integrative approach support the specific status of all closely related consubgeneric six
373 species studied here (Cameron et al., 2007; Williams, 1998). Similar interspecific
374 differentiations were recorded independently based on DNA (Cameron et al., 2007, Hines &
375 Cameron, 2010) and CLGS composition (Appelgren, Bergstrom, Svensson & Cederberg,
376 1991; Brasero et al., 2018a; Terzo et al., 2005). However, the analyses in this study were
377 carried out on a larger sampling covering the distribution of the species. Within "*runderarius*
378 group" and "*sylvarum* group" and despite some phenotypic differentiations, all studied taxa
379 are regarded as conspecific (Tab. 2). *runderarius* and *sylvarum* groups appear respectively as
380 single species with high geographic phenotypic differentiations. In the light of our results, we
381 decide for the "*sylvarum* group" to use the oldest available nominal species name as the valid
382 name for all of them: *B. sylvarum sylvarum* except for *daghestanicus* where we maintain the
383 subspecific status *B. sylvarum daghestanicus* from east of Turkey and Iran because of its
384 private haplotype in COI. Inside the "*runderarius* group" we decide to conserve the subspecific
385 status *B. runderarius simulatilis* for the population in the east of Turkey and Iran based on their
386 CLGS differentiation and private haplotype (COI and PEPCK) and raised the other
387 populations to the oldest available nominal species name *B. runderarius runderarius* (see
388 decision framework Tab. 2). For the comparison groups *B. inexpectatus*, *B. mlokosievitzii*, *B.*
389 *veteranus* and *B. velox*, our analyses (Fig 2A, 3A, 4 & 5) confirm clear differentiation
390 between these species (Cameron et al., 2007; Hines & Cameron, 2010). Inside *B.*
391 *mlokosievitzii*, we have not been able to analyze the subspecies *convergens* and *vogtiellus*.

392 *Intraspecific differentiation of B. sylvarum and B. runderarius*

393 *B. sylvarum* is a widespread species and can be found in Central Spain, Sicily, southern Italy,
394 Greece and Southern Turkey of its distribution. It reaches Ireland and northern Portugal to the
395 West and Mongolia to the East. To the north, this species nearly reaches the Arctic Circle
396 (Rasmont et al., 2015a). Divergence of color patterns (sometimes very pronounced: see the

397 melanic form *nigrescens*), seems to be a simple intraspecific variation as encountered in other
398 bumblebees species such as *B. pascuorum* (Lecocq, Brasero, Martinet, Valterova & Rasmont,
399 2015c). However, even if unique haplotypes (Fig. 2B) are found in *daghestanicus* (COI gene),
400 there were not strongly supported as reciprocally monophyletic group in our phylogeny (Fig.
401 2A). No private haplotype was found in PEPCK gene for this taxon (Fig. 3B, C). Moreover,
402 inside the "*sylvarum* group", high probabilities (>0.05) to be conspecific is showed by the
403 bGMYC analyses (Fig. 4). On the other side, the CLGS did not highlight any divergence
404 between the different populations which would imply that all of them could recognize each
405 other as sexual partners.

406 *B. ruderarius* is also a widespread species and its distribution is almost identical to that of *B.*
407 *sylvarum*. However, *B. ruderarius* is absent from Sicily and the southern Italy but the
408 subspecies *B. ruderarius tunensis* occurs in the east of Algeria and Tunisia (Rasmont et al.,
409 2015a). The same configuration as in *B. sylvarum* has been highlighted. Indeed, the
410 divergence in color patterns seems to be a simple intraspecific variation and no diagnostic
411 character was found to discriminate significantly the different population inside *B. ruderarius*.
412 The unique haplotypes found in COI and PEPCK (Fig. 2B, 3B, C) for *simulatilis* were not
413 strongly supported as reciprocally monophyletic group in our phylogenies (Fig. 2A, 3A) as
414 well in our bGMYC analyses (Fig. 4). However, the CLGS of this Turkish population differ
415 from other populations and it is well supported in our cluster analysis (Fig. 5) but share main
416 compounds (hexadec-9-en-1-ol, tricos-9-ene + tricos-7-ene) with the other *B. ruderarius* taxa
417 (supplementary file 2). We can therefore hypothesize that this divergence involved in the
418 recognition system is not sufficient to induce a pre-mating reproductive barrier. This
419 geographic differentiation sharing main compounds could suggest only a dialect between
420 populations as already described in moths (Tòth et al., 1992) and solitary bees (Vereecken,
421 Mant & Schiestl, 2007). Although the taxon *simulatilis* seems more advanced on the path of

422 speciation than *daghestanicus*, complementary tests based on ethological behavior are needed.
423 We then hypothesize that these results could reflect the recent divergence of these taxa.
424 Unfortunately, we have not been able to analyze the subspecies *citrinofasciatus* and *tunensis*
425 respectively inside the "*sylvarum* group" and the "*runderarius* group".

426 Turkey is divided by a diagonal belt which is a massive structure formed by continuous
427 mountains (Miocene-Pliocene) from the north-east towards the south-west. These mountains
428 isolated populations from east and west in many species of insects or plants (Çiplak,
429 Demisroy & Bozcuk, 1993; Çiplak, 2003; Ekim & Güner, 1986). Moreover, the distribution
430 of taxa from the West of this barrier generally spread through Europe while the taxa from the
431 east spread through Iran and Asia. In addition with this geographical particularities, climatic
432 change during the ice ages pushed many northern populations especially those found in the
433 Balkans and Caucasus to the south and led to their isolation and speciation at higher altitudes
434 (Hewitt, 1996, 2000). These two taxa (*B. ruderarius simulatilis* and *B. sylvarum*
435 *daghestanicus*) are known to occur in the east part of the diagonal belt and may have
436 colonized Turkey from Caucasus and Asia while the taxa to the west of this barrier may have
437 colonized Turkey from Balkans. It is also interesting to note that *B. sylvarum daghestanicus*
438 and *B. ruderarius simulatilis* are characterized by the same color pattern (white bands and red
439 tail) (Fig. 1D, E) which could be explained by a local Müllerian mimicry processes (Hines &
440 Williams, 2012). Other species from other subgenus also posses the same color pattern in this
441 region and could confirm this process: *B. cullumanus apollineus* (*Cullumanobombus*), *B.*
442 *sicheli cazuroi* and *B. incertus* (*Melanobombus*), *B. handlirschianus* (*Mendacibombus*), *B.*
443 *brodmannicus* (*Pyrobombus*) and *B. niveatus* (*Sibiricobombus*) (Rasmont & Iserbyt, 2014).

444 *Consider GAPS or not as phylogenetic information?*

445 Our PEPCCK gene fragment contains gaps sites (representing an insertion or deletion and also
446 called "indels") and our statistical method for phylogenetic estimation (ML and MB) consider

447 the gap as missing data. Indeed, in most of taxonomic studies, the indels are treated as missing
448 data as this is the case for birds (Alström et al., 2015; McKay et al., 2014), frogs (Glaw,
449 Kohler, De la Riva, Vieites & Vences, 2010), mammals (Ndiaye et al., 2014), spiders (Rezac,
450 Gasparo, Kral & Heneberg, 2014). Considered indels will increase the phylogenetic
451 information even if they are located in introns (in our case, all indels are located in the first
452 introns of the PEPCK gene fragment). These gaps reflect the evolutionary history of the taxa
453 and thus contain phylogenetic information which could be very informative particularly for
454 *simulatilis* where the divergence estimation with all other *runderarius* populations remains
455 unclear. It's seems therefore that the treatment of indels has not been fully investigated in
456 taxonomic studies and delimitaion of species. Moreover, in our study, haplotype network
457 analysis considering or not the indels do not bring us any additional information except for *B.*
458 *velox* which not differs from *B. ruderarius* if we not considered indels in our analysis (Fig.
459 3B). Indeed, only gap sites allow to separate *B. velox* from *B. ruderarius*. However, both
460 species are respectively strongly differenciated by the genitalia and other morphological
461 characters.

462 *Conservation implication*

463 Our integrative approach brings to attention the relevance of geographically isolated groups
464 differentiated in genetic and reproductive traits. This strict approach allows defining
465 taxonomic status (species or subspecies) and providing a decision framework for policy-
466 makers and conservation organizations. Except for *B. inexpectatus* (Rasmont et al., 2015b)
467 considered as endangered species in the International Union for Conservation of Nature
468 (IUCN) Red List, *B. mlokosievitzii* as data deficient (Nieto et al., 2014) and *B. velox* who is
469 not yet assessed, the other taxa (*B. ruderarius*, *B. sylvarum* and *B. veteranus*) are all
470 considered as least concern by Nieto et al. (2014). Our new taxonomical hypotheses taking
471 into account two subspecies (*B. ruderarius simulatilis* and *B. sylvarum daghestanicus*) have

472 no implications for the red list assessments. However, the conservation of subspecies with
473 unique haplotypes as well as CLGS differentiation should be seen as source of diversity
474 within respective species and then should be considered and integrated in future management
475 plan.

476

477 **Acknowledgments**

478 The authors thank the Parco Nazionale dei Monti Sibillini for granting permission to collect in
479 their respective territories to Paolo Biella. Special thanks go to A. Cetkovic and A. Popovic
480 (University of Belgrade), R. De Jonghe, P. Sagot and V. Cyriaque (University of Mons) for
481 their help in the sampling but also to Simon Dellicour (University of Leuven) for his advices
482 in phylogenetic analyses. Computational resources have been provided by the Consortium des
483 Equipements de Calcul Intensif (CECI), funded by the Belgian FRS-FNRS (Fonds de la
484 Recherche Scientifique). BM contributes as a PhD student granted by the Research Council of
485 University of Mons and by the FRS-FNRS. Part of this work (Eco-chemical trait
486 differentiation) was supported by the Institute of Organic Chemistry and Biochemistry of the
487 Academy of Sciences of the Czech Republic (#61388963). The research has received funding
488 from the FRS-FNRS (Fonds de la Recherche Scientifique) (2013/V 3/5/307-IB/JN-11888)
489 and the European Community's Seventh Framework Program, STEP Project (Status and
490 Trends of European Pollinators, [www.step-project .net](http://www.step-project.net), grant agreement no 244090,
491 FP7/2007-2013).

492

493 **References**

494 Alström, P., Xia, C., Rasmussen, P.C., Olsson, U., Dai, B., Zhao, J., Leader, P.J., Carey, G.J.,
495 Dong, L., Cai, T., Holt, P.I., Le Manh, H., Song, G., Liu, Y., Zhang, Y. & Lei, F.

496 (2015). Integrative taxonomy of the Russet Bush Warbler *Locustella mandelli* complex
497 reveals a new species from central China. *Avian Research*, 6, 9.

498 Amiet, F. (1996). Hymenoptera Apidae, 1. Teil: Allgemeiner Teil, Gattungsschlüssel, die
499 Gattungen *Apis*, *Bombus* und *Psithyrus*. *Insecta Helvetica Fauna*, 12, 1–98.

500 Appelgren, M., Bergstrom, G., Svensson, B.G. & Cederberg, B. (1991). Marking pheromones
501 of *Megabombus* bumble bee males. *Acta Chem. Scand*, 45, 972–974.

502 Ayasse, M. & Jarau, S. (2014). Chemical ecology of bumble bees. *Annual review of*
503 *entomology*, 59, 299–319.

504 Bertsch, A., Schweer, H., Titze, A. & Tanaka, H. (2005). Male labial gland secretions and
505 mitochondrial DNA markers support species status of *Bombus cryptarum* and *B. magnus*
506 (Hymenoptera, Apidae). *Insectes Sociaux*, 52, 45–54.

507 Brasero, N., Martinet, B., Urbanová, K., Valterová, I., Torres, A., Hoffmann, W., Rasmont, P.
508 & Lecocq, T. (2015). First Chemical Analysis and Characterization of the Male Species-
509 Specific Cephalic Labial-Gland Secretions of South American Bumblebees. *Chemistry &*
510 *Biodiversity*, 12, 1535–1546.

511 Brasero, N., Martinet, B., Lecocq, T., Lhomme, P., Biella, P., Valterova, I., Urbanova, K.,
512 Cornalba, M., Hines, H. & Rasmont, P. (2018a). The cephalic labial gland secretions of
513 two socially parasitic bumblebees *Bombus hyperboreus* (*Alpinobombus*) and *Bombus*
514 *inexpectatus* (*Thoracobombus*) question their inquiline strategy. *Insect Sciences*, 25(1),
515 75-86.

516 Brasero, N., Lecocq, T., Martinet, B., Valterova, I., Urbanova, K., de Jonghe, R. & Rasmont,
517 P. (2018b). Variability in sexual pheromones questions their role in bumblebee pre-
518 mating recognition system. *Journal of Chemical Ecology*, 44(1): 9-17.

519 Calam, D.H. (1969). Species and Sex-specific Compounds from the Heads of Male
520 Bumblebees (*Bombus* spp.). *Nature*, 221, 856–857.

521 Cameron, S.A., Hines, H.M. & Williams, P.H. (2007). A comprehensive phylogeny of the
522 bumble bees (*Bombus*). *Biological Journal of the Linnean Society*, 91, 161–188.

523 Çiplak, B. (2003). Distribution of Tettigoniinae (Orthoptera, Tettigoniidae) bush-crickets in
524 Turkey: the importance of the Anatolian Taurus Mountains in biodiversity and
525 implications for conservation. *Biodiversity and Conservation*, 12, 47–64.

526 Çiplak, B., Demisroy, A. & Bozcuk, N. (1993). Distribution of Orthoptera in Relation to the
527 Anatolian Diagonal in Turkey. *Articulata*, 8, 1–20.

528 Claudet, J., Pelletier, D., Jouvenel, J.Y., Bachet, F. & Galzin, R. (2006). Assessing the effects
529 of marine protected area (MPA) on a reef fish assemblage in a northwestern
530 Mediterranean marine reserve: Identifying community-based indicators. *Biological
531 Conservation*, 130, 349–369.

532 Coppée, A., Terzo, M., Valterová, I. & Rasmont, P. (2008). Intraspecific variation of the
533 cephalic labial gland secretions in *Bombus terrestris* (L.) (Hymenoptera: Apidae).
534 *Chemistry & Biodiversity*, 5, 2654–2661.

535 Cvacka, J., Kofronová, E., Vasícková, S., Stránský, K., Jiros, P., Hovorka, O., Kindl, J. &
536 Valterová, I. (2008) Unusual fatty acids in the fat body of the early nesting bumblebee,
537 *Bombus pratorum*. *Lipids*, 43, 441–450.

538 Dalla Torre, K.W. (1882). Bemerkungen zur Gattung *Bombus* Latr. II. *Bericht des
539 naturwissenschaftlich-medizinischen Vereins in Innsbruck*, 12, 14–31.

540 Dalla Torre, K.W. (1880). Unsere Hummel-(*Bombus*) Arten. *Die Naturhistoriker*, 30, 40–41.

541 Dellicour, S. & Lecocq, T. (2013). *GALIGNER 1.0 and GCKOVATS 1.0 - Manual of a
542 software suite to compute a multiple sample comparison data matrix from eco-chemical
543 datasets obtained by gas chromatography*. University of Mons, Mons (Belgium).

544 De Meulemeester, T., Gerbaux, P., Boulvin, M., Coppée, A., & Rasmont, P. (2011). A
545 simplified protocol for bumble bee species identification by cephalic secretion analysis.

546 *Insectes Sociaux*, 52(2), 227-236.

547 De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, 56,
548 879–886.

549 Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012). Bayesian phylogenetics with
550 BEAUTI and BEAST 1.7. *Molecular biology and evolution*, 29(8), 1969-1973.

551 Dufrière, M. & Legendre, P. (1997). Species assemblages and indicator species: The need for
552 a flexible asymmetrical approach. *Ecological Monographs*, 67, 345–366.

553 Ekim, T. & Güner, A. (1986). The Anatolian Diagonal: fact or fiction? *Proceedings of the*
554 *Royal Society of Edinburgh, Section B: Biological Sciences*, 89, 69–77.

555 Ennen, J.R., Kalis, M.E., Patterson, A.L., Kreiser, B.R., Lovich, J.E., Godwin, J. & Qualls,
556 C.P. (2014). Clinal variation or validation of a subspecies? A case study of the
557 *Graptemys nigrinoda* complex (Testudines: Emydidae). *Biological Journal of the*
558 *Linnean Society*, 111, 810–822.

559 Fabricius, J.C. (1793). *Entomologica systematica emendata at aucta. Secundum classes,*
560 *ordines, genera, species adjectis synonymis, locis observationibus. Vol. 2. Hafniae.*

561 Françoso, E., Freitas de Oliveira, F. & Arias, M.C. (2016). An integrative approach identifies
562 a new species of bumblebee (Hymenoptera: Apidae: Bombini) from northeastern Brazil.
563 *Apidologie*, 47, 171–185.

564 Gibbs, J. (2009). Integrative taxonomy identifies new (and old) species in the *Lasioglossum*
565 (*Dialictus*) *tegulare* (Robertson) species group (Hymenoptera, Halictidae). *Zootaxa*,
566 2032, 1–38.

567 Glaw, F., Kohler, J., De la Riva, I., Vieites, D.R. & Vences, M. (2010). Integrative taxonomy
568 of Malagasy treefrogs: combination of molecular genetics, bioacoustics and comparative
569 morphology reveals twelve additional species of *Boophis*. *Zootaxa*, 2383, 1–82.

570 Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis

571 program for Windows 95/98/NT. *Nucleic Acids Symp. Research*, 41, 95–98.

572 Hawlitschek, O., Nagy, Z.T. & Glaw, F. (2012). Island evolution and systematic revision of
573 comoran snakes: Why and when subspecies still make sense. *PLoS ONE*, 7(8): e42970.

574 Hewitt, G.M. (1996). Some genetic consequences of ice ages, and their role in divergence and
575 speciation. *Biological Journal of the Linnean Society*, 58, 247–276.

576 Hewitt, G.M. (2000). The genetic legacy of the quaternary ice ages. *Nature*, 405, 907–913.

577 Hillis, D.M. & Bull, J.J. (1993). An empirical test of bootstrapping as a method for assessing
578 confidence in phylogenetic analysis. *Systematic Biology*, 42, 182–192.

579 Hines, H.M. & Williams, P.H. (2012). Mimetic colour pattern evolution in the highly
580 polymorphic *Bombus trifasciatus* (Hymenoptera: Apidae) species complex and its
581 comimics. *Zoological Journal of the Linnean Society*, 166, 805–826.

582 Hines, H.M. & Cameron, S.A. (2010). The phylogenetic position of the bumble bee inquiline
583 *Bombus inexpectatus* and implications for the evolution of social parasitism. *Insectes*
584 *sociaux*, 57, 379–383.

585 Hurvich, C.M. & Tsai, C.L. (1989). Regression and time series model selection in small
586 samples. *Biometrika*, 76, 297–307.

587 Lecocq, T., Lhomme, P., Michez, D., Dellicour, S., Valterová, I. & Rasmont, P. (2011).
588 Molecular and chemical characters to evaluate species status of two cuckoo bumblebees:
589 *Bombus barbutellus* and *Bombus maxillosus* (Hymenoptera, Apidae, Bombini).
590 *Systematic Entomology*, 36, 453–469.

591 Lecocq, T., Dellicour, S., Michez, D., Lhomme, P., Vanderplanck, M., Valterová, I., Rasplus,
592 J.Y. & Rasmont, P. (2013). Scent of a break-up: phylogeography and reproductive trait
593 divergences in the red-tailed bumblebee (*Bombus lapidarius*). *BMC evolutionary*
594 *biology*, 13, 263.

595 Lecocq, T., Brasero, N., De Meulemeester, T., Michez, D., Dellicour, S., Lhomme, P., De

596 Jonghe, R., Valterová, I., Urbanová, K. & Rasmont, P. (2015a). An integrative
597 taxonomic approach to assess the status of Corsican bumblebees: implications for
598 conservation. *Animal conservation*, 18, 236–248.

599 Lecocq, T., Dellicour, S., Michez, D., Dehon, M., Dewulf, A., De Meulemeester, T., Brasero,
600 N., Valterová, I., Rasplus, J.Y. & Rasmont, P. (2015b). Methods for species delimitation
601 in bumblebees (Hymenoptera, Apidae, *Bombus*): towards an integrative approach.
602 *Zoologica Scripta*, 44(3), 281-297.

603 Lecocq, T., Brasero, N., Martinet, B., Valterová, I. & Rasmont, P. (2015c). Highly polytypic
604 taxon complex: interspecific and intraspecific integrative taxonomic assessment of the
605 widespread pollinator *Bombus pascuorum* Scopoli 1763 (Hymenoptera: Apidae).
606 *Systematic Entomology*, 40, 881–890.

607 Linnaeus, C. (1761). *Fauna Suecica sistens animalia Suecica regni: Mammalia, Aves,*
608 *Amphibia, Pisces, Insecta, Vermes. Distributa per classes & ordines, genera & species,*
609 *cum differentiis specierum, synonymis auctorum, nominibus incolarum, locis natalium,*
610 *descriptionibus insec.*

611 Martin, S.J., Helantera, H. & Drijfhout, F.P. (2008). Evolution of species-specific cuticular
612 hydrocarbon patterns in Formica ants. *Biological Journal of the Linnean Society*, 95,
613 131–140.

614 Martinet, B., Lecocq, T., Brasero, N., Biella, P., Urbanová, K., Valterová, I., Cornalba, M.,
615 Gjershaug, J.O., Michez, D. & Rasmont, P. (2018). Following the cold: geographical
616 differentiation between interglacial refugia and speciation in the arcto-alpine species
617 complex *Bombus monticola* (hymenoptera: Apidae). *Systematic Entomology*, 43(1), 200-
618 217.

619 Mayr, E. (1942). *Systematics and the Origin of Species*. New York.

620 McKay, B.D., Mays, H.L., Yao, C.T., Wan, D., Higuchi, H. and Nishiumi, I. (2014).

621 Oncorporating color into integrative taxonomy: analysis of the varied tit (*Sittiparus*
622 *varius*) complex in East Asia. *Systematic Entomology*, 63(4), 505-517.

623 Moraes, G.J. (1987). Importance of taxonomy in biological control. *Insect sciences*
624 *Applications*, 8, 841–844.

625 Moure, J.S. & Sakagami, S.F. (1962). As Mamangabas Sociais do Brasil (*Bombus* Latreille)
626 (Hymenoptera, Apoidea). *Studia Entomologica*. 5, 65–194.

627 Müller, A. (2006). A scientific note on *Bombus inexpectatus* (Tkalcu, 1963): evidence for a
628 social parasitic mode of life. *Apidologie*, 37, 408–409.

629 Müller, O.F. (1776). *Zoologiae Danicae prodromus, seu animalium Danicae et Norvegiae*
630 *indigenarum characteres, nomina, et synonyma imprimis popularium*. Hallager (Ed).
631 Haniae (Köbenhavn).

632 Murray, T.E., Fitzpatrick, U., Brown, M.J.F. & Paxton, R.J. (2008). Cryptic species diversity
633 in a widespread bumble bee complex revealed using mitochondrial DNA RFLPs.
634 *Conservation Genetics*, 9, 653–666.

635 Ndiaye, A., Hima, K., Dobigny, G., Sow, A., Dalecky, A., Bâ, K., Thiam, M. and Granjon, L.
636 (2014). Integrative taxonomy of a poorly known Sahelian rodent, *Gerbillus nancillus*
637 (Muridae, Gerbillinae). *Zoologischer Anzeiger*, 253, 430-439.

638 Nieto, A., Roberts, S.P.M., Kemp, J., Rasmont, P., Kuhlmann, M., García Criado, M.,
639 Biesmeijer, J.C., Bogusch, P., Dathe, H.H., De la Rúa, P., De Meulemeester, T., Dehon,
640 M., Dewulf, A., Ortiz-Sánchez, F.J., Lhomme, P., Pauly, A., Potts, S.G., Praz, C.,
641 Quaranta, M., Radchenko, V.G., Scheuchl, E., Smit, J., Straka, J., Terzo, M., Tomozii,
642 B., Window, J. & Michez, D. (2014). European Red List of bees. Luxembourg:
643 Publication Office of the European Union.

644 Oksanen, F.J., Blanchet, G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson,
645 G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2011). Vegan: Community Ecology
646 Package. Tertiary Vegan: Community Ecology Package.

647 Oleksa, A. & Tofilski, A. (2015). Wing geometric morphometrics and microsatellite analysis
648 provide similar discrimination of honey bee subspecies. *Apidologie*, 46, 49–60.

649 Padial, J.M., Miralles, A., De la Riva, I. & Vences, M. (2010). The integrative future of
650 taxonomy. *Frontiers in zoology*, 7, 16.

651 Paquin, P., Dupérré, N., Cokendolpher, J.C., White, K. & Hedin, M. (2008). The fundamental
652 importance of taxonomy in conservation biology: the case of the eyeless *Cicurina*
653 *bandida* (Araneae: Dictynidae) of central Texas, including new synonyms and the
654 description of the male of the species. *Invertebrate Systematics*, 22, 139–149.

655 Paterson, H.E.H. (1985). The Recognition Concept of Species. *Transvaal Museum*
656 *Monograph*, 4, 21–29.

657 Pérez, J. (1879). Contribution à la faune des apiaires de France. 1re partie. *Actes de la Société*
658 *linnéenne de Bordeaux*, 33, 119–229.

659 Phillimore, A.B. & Owens, I.P.F. (2006). Are subspecies useful in evolutionary and
660 conservation biology? *Proceedings of the Royal Society B: Biological Sciences*, 273,
661 1049–1053.

662 Posada, D. (2008). jModelTest: Phylogenetic model averaging. *Molecular Biology and*
663 *Evolution*, 25, 1253–1256.

664 R Development Core Team (2016). R: A language and environment for statistical computing.

665 Radoszkowski, O. (1877). 3 Maya. 1) *Bombus daghestanicus* 2) *Bombus mlokosievitzii*. *Trudy*
666 *russskogo entomologicheskogo Obshchestva*, 13, VII–VIII.

667 Radoszkowski, O. (1888). Etudes hyménoptérologiques. I. Revisions des armures copulatrices
668 des mâles. II: Description de nouvelles espèces russes. *Trudy russskogo*

669 *entomologicheskogo Obshchestva*, 22, 315–337.

670 Rasmont, P. (1983). Catalogue commenté des Bourdons de la région ouest-paléarctique
671 (Hymenoptera, Apoïdea, Apidae). *Notes fauniques de Gembloux*, 7, 1-72.

672 Rasmont, P., Coppée, A., Michez, D. & De Meulemeester, T. (2008). An overview of the
673 *Bombus terrestris* (L. 1758) subspecies (Hymenoptera: Apidae). *Annales de la Société*
674 *Entomologique de France*, 44, 243–250.

675 Rasmont, P. & Iserbyt, S. (2014). Atlas of the European Bees: genus *Bombus*. 3d Edition.
676 STEP Project. *Status and Trends of European Pollinators, Atlas Hymenoptera, Mons,*
677 *Gembloux (Belgium)*, 1.

678 Rasmont, P., Franzen, M., Lecocq, T., Harpke, A., Roberts, S., Biesmeijer, K., Castro, L.,
679 Cederberg, B., Dvorak, L., Fitzpatrick, U., Gonseth, Y., Haubruge, E., Mahe, G.,
680 Manino, A., Michez, D., Neumayer, J., Odegaard, F., Paukkunen, J., Pawlikowski, T.,
681 Potts, S., Reemer, M., Settele, J., Straka, J. & Schweiger, O. (2015a). Climatic Risk and
682 Distribution Atlas of European Bumblebees. *BioRisk*, 10, 1–236.

683 Rasmont, P., Roberts, S., Cederberg, B., Radchenko, V. & Michez, D. (2015b). *Bombus*
684 *inexpectatus*. The IUCN Red List of Threatened Species 2015:
685 e.T13340462A57349805.[http://dx.doi.org/10.2305/IUCN.UK.20151.RLTS.T13340462A](http://dx.doi.org/10.2305/IUCN.UK.20151.RLTS.T13340462A57349805.en)
686 [57349805.en](http://dx.doi.org/10.2305/IUCN.UK.20151.RLTS.T13340462A57349805.en). *The IUCN Red List of Threatened Species 2015: e.T13340462A57349805.*
687 <http://dx.doi.org/10.2305/IUCN.UK.2015-1.RLTS.T13340462A57349805.en>.

688 Reid, N.M. & Carstens, B.C. (2012). Phylogenetic estimation error can decrease the accuracy
689 of species delimitation: A Bayesian implementation of the general mixed Yule-
690 coalescent model. *BMC Evolutionary Biology*, 12.

691 Rezac, M., Gasparo, F., Kral, J. & Heneberg, P. (2014). Integrative taxonomy and
692 evolutionary history of a newly revealed spider *Dysdera ninnii* complex (Areneae:
693 Dysderidae). *Zoological Journal of the Linnean Society*, 172(2), 451-474.

694 Ronquist, F. & Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under
695 mixed models. *Bioinformatics*, 19, 1572–1574.

696 Sackett, L.C., Seglund, A., Guralnick, R.P., Mazzella, M.N., Wagner, D.M. & Martin, A.P.
697 (2014). Evidence for two subspecies of Gunnison’s prairie dogs (*Cynomys gunnisoni*),
698 and the general importance of the subspecies concept. *Biological Conservation*, 174, 1–
699 11.

700 Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., Stauffer, C., Christian, E. & Crozier, R.H.
701 (2010). Integrative taxonomy: a multisource approach to exploring biodiversity. *Annual*
702 *Review of Entomology*, 55, 421–438.

703 Skorikov, A.S. (1908). Novyya formy shmeley (Hymenoptera, Bombidae). (Predvaritel’nye
704 diagnozy). I. *Russkoe entomologicheskoe Obozrenie*, 7, 111–113.

705 Skorikov, A.S. (1914). Les formes nouvelles des bourdons (Hymenoptera, Bombidae). V.
706 *Russkoe entomologicheskoe Obozrenie*, 14, 119–129.

707 Skorikov, A.S. (1922). Palaearctic bumble bees. Part I. General biology (including
708 zoogeography) (in russian). *Izvestiya Severnoi oblastnoi stantsii zashchity rastenii ot*
709 *vreditelei*, 4, 1–160.

710 Suzuki, R. & Shimodaira, H. (2011). Pvcust: Hierarchical Clustering with P-Values via
711 Multiscale Bootstrap Resampling.

712 Terzo, M., Urbanová, K., Valterová, I. & Rasmont, P. (2005). Intra and interspecific
713 variability of the cephalic labial glands’ secretions in male bumblebees: the case of
714 *Bombus (Thoracobombus) ruderarius* and *B. (Thoracobombus) sylvarum* Hymenoptera,
715 Apidae. *Apidologie*, 36, 85–96.

716 Tkalců, B. (1963). Eine neue Hummel-Art der Gattung *Agrobombus* Vogt aus dem
717 Alpengebiet. *Cas. Cesk. Spol. Entomo.*, 60, 183–196.

718 Tkalců, B. (1965). Über *Agrobombus inexpectatus* Tkalců. *Reichenbachia*, 5, 225–230.

719 Tkalcù, B. (1974). Beitrag zur Kenntnis der Hummelfauna der französischen Besses-Alpes
720 (Hymenoptera, Apoidea, Bombinae). *Acta Rerum naturalium Musei nationalis Slovenici*
721 *Bratislava*, 20, 167–186.

722 Tòth, M., Löfstedt, C., Blair, B.W., Cabello, T., Farag, A.I., Hansson, B.S., Kovalev, B.G.,
723 Maini, S., Nesterov, E.A., Pajor, I., Sazonov, A.P., Shamshev, I.V., Subchev, M. &
724 Szöcs, G. (1992). Attraction of male turnip moths *Agrotis segetum* (Lepidoptera:
725 Noctuidae) to sex pheromone components and their mixtures at 11 sites in Europe, Asia,
726 and Africa. *Journal of Chemical Ecology*, 18, 1337–1347.

727 Vereecken, N.J., Mant, J. & Schiestl, F.P. (2007). Population differentiation in female sex
728 pheromone and male preferences in a solitary bee. *Behavioral Ecology and Sociobiology*,
729 61, 811–821.

730 Vogt, O. (1909). Studien über das Artproblem. 1. Mitteilung. über das Variieren der
731 Hummeln. I. Teil. *Schriften der berlinischen Gesellschaft Naturforschender, Freunde,*
732 *Berlin*, 1909, 28–84.

733 Wheeler, Q.D., Raven, P.H. & Wilson, E.O. (2004). Taxonomy: Impediment or expedient?
734 *Science*, 303, 285.

735 White, B.P., Pilgrim, E.M., Boykin, L.M., Stein, E.D. & Mazor, R.D. (2014). Comparison of
736 four species-delimitation methods applied to a DNA barcode data set of insect larvae for
737 use in routine bioassessment. *Freshwater Science*, 33, 338–348.

738 Wilcox, T.P., Zwickl, D.J., Heath, T.A. & Hillis, D.M. (2002). Phylogenetic relationships of
739 the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic
740 support. *Molecular Phylogenetics and Evolution*, 25, 361–371.

741 Williams, P.H. (1998). An annotated checklist of bumble bees with an analysis of patterns of
742 description (Hymenoptera: Apidae, Bombini). *Bulletin of the Natural History Museum*
743 *(Entomology)*, 67, 79-152.

744 Williams, P.H., Brown, M.J.F., Carolan, J.C., An, J., Goulson, D., Aytakin, A.M., Best, L.R.,
745 Byvaltsev, A.M., Cederberg, B., Dawson, R., Huang, J., Ito, M., Monfared, A., Raina,
746 R.H., Schmid-Hempel, P., Sheffield, C.S., Šima, P. & Xie, Z. (2012). Unveiling cryptic
747 species of the bumblebee subgenus *Bombus* s. str. worldwide with COI barcodes
748 (Hymenoptera: Apidae). *Systematics and Biodiversity*, 10, 21–56.

749 Yarrow, I.H.H. (1970). Is *Bombus inexpectatus* (Tkalcu) a workless parasite? *Insectes*
750 *Sociaux*, 17, 95–112.

751 Žáček, P., Prchalová-Hornáková, D., Tykva, R., Kindl, J., Vogel, H., Svatoš, A., Pichová, I. &
752 Valterová, I. (2013). De Novo Biosynthesis of Sexual Pheromone in the Labial Gland of
753 Bumblebee Males. *ChemBioChem*, 14, 361–371.

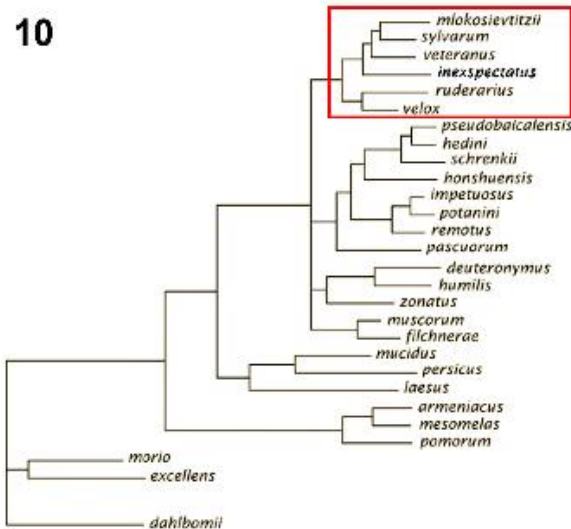
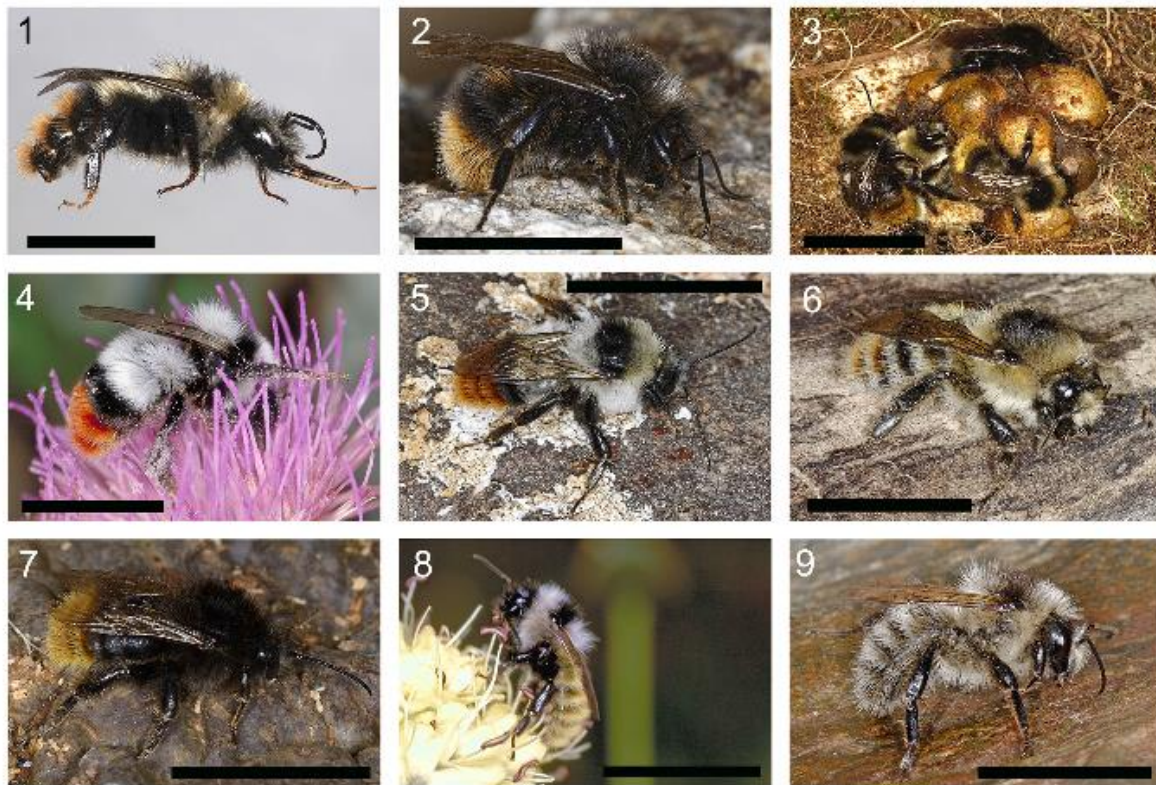
754 Zwickl, D.J. (2006). Genetic algorithm approaches for the phylogenetic analysis of large
755 biological sequence datasets under the maximum likelihood criteria. The University of
756 Texas at Austin Editor.

757

758

759

760 **Figures**

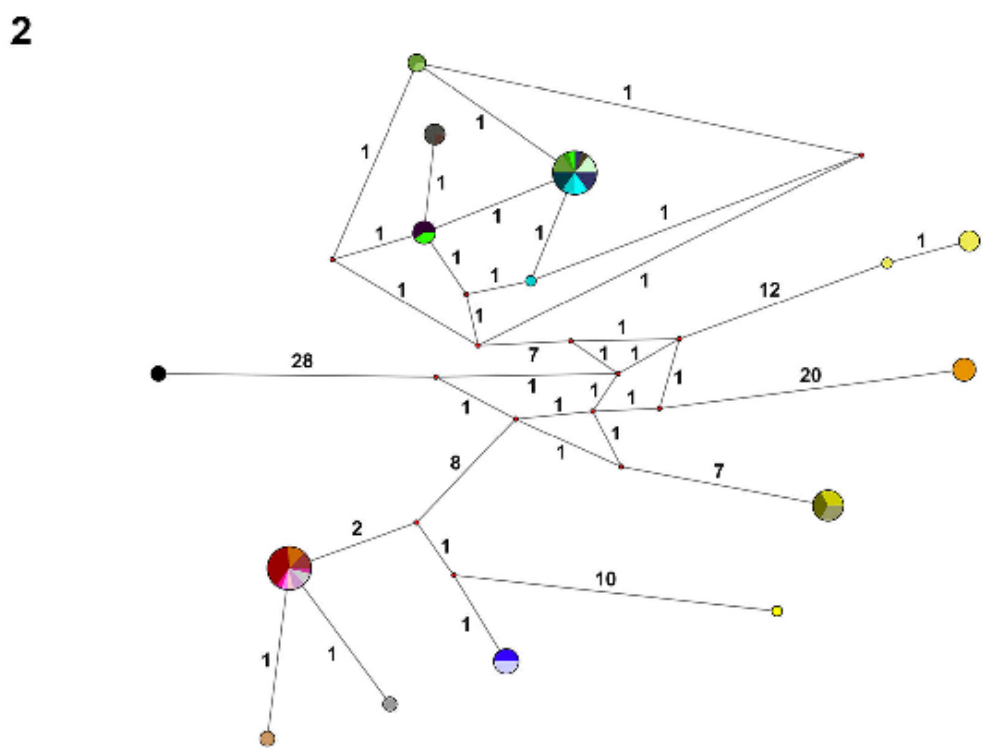
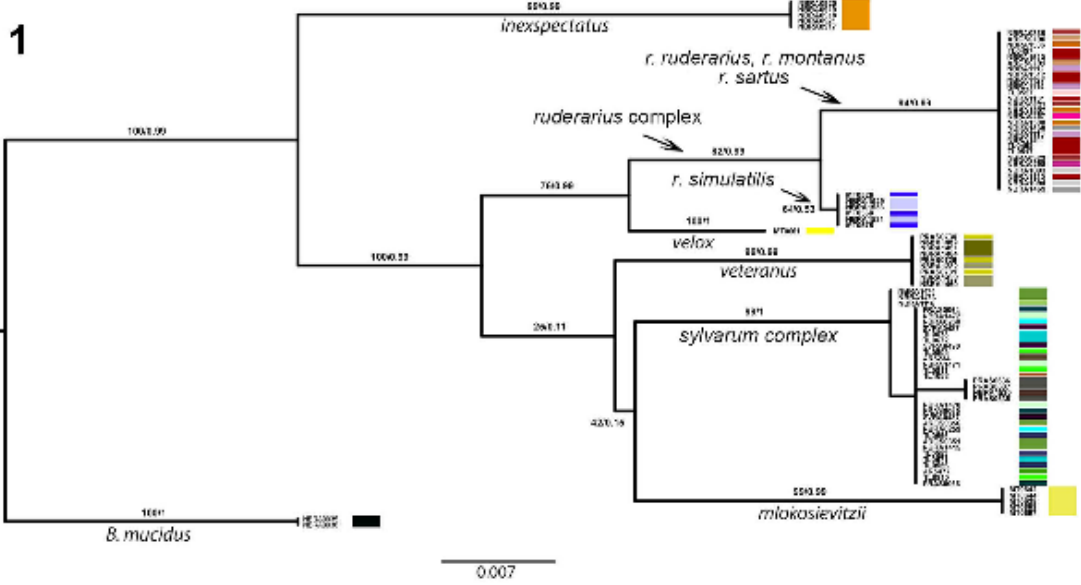


761

762 **Figure 1.** Photos of some studied bumblebee species: **A)** *Bombus inexpectatus* male from
 763 Alps, Italy; **B)** *B. mlokosievtzii vogtiellus* male from Turkey ; **C)** *B. ruderarius ruderarius*
 764 and *B. ruderarius montanus* from Pyrenees, France; **D)** *B. ruderarius simulatilis* female from
 765 Turkey; **E)** *B. sylvarum daghestanicus* male from Turkey; **F)** *B. sylvarum sylvarum* female
 766 from Pyrenees, France; **G)** *B. sylvarum nigrescens* male from Sweden; **H)** *B. velox* male from
 767 Turkey and **I)** *B. veteranus* from the Netherlands. All black lines correspond to a scale of 1

768 cm. All photographs are by P. Rasmont. **J**) Phylogeny of *Thoracobombus sensu stricto*. The
769 red box shows the monophyletic group of all species included in this study (according to
770 Hines & Cameron 2010).

771



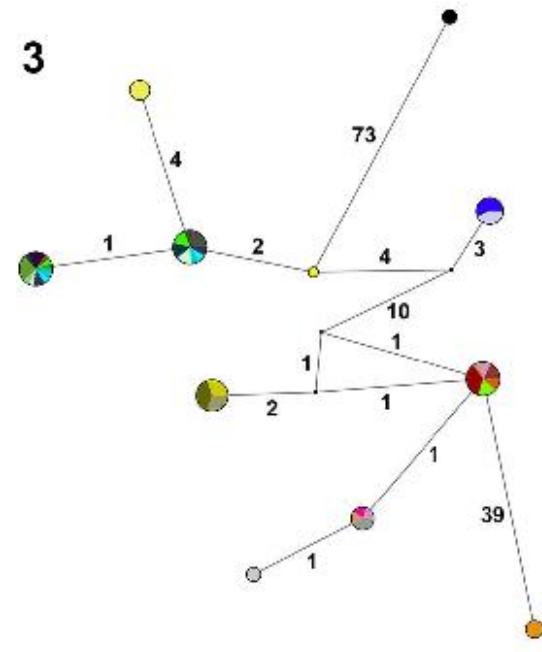
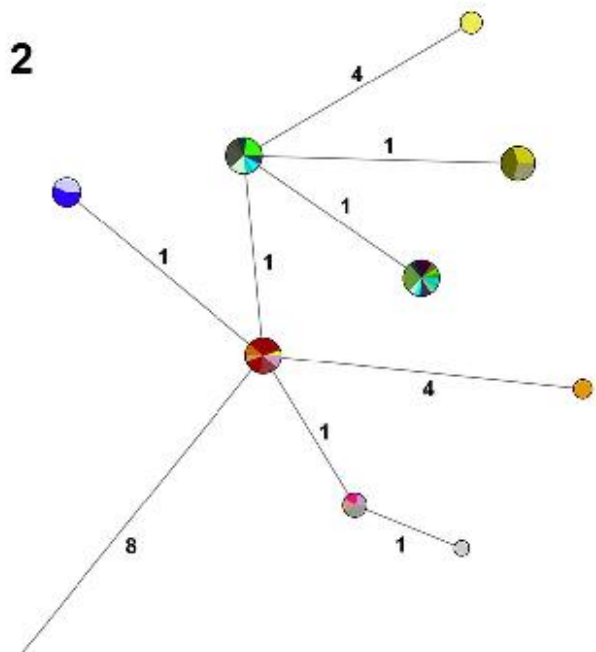
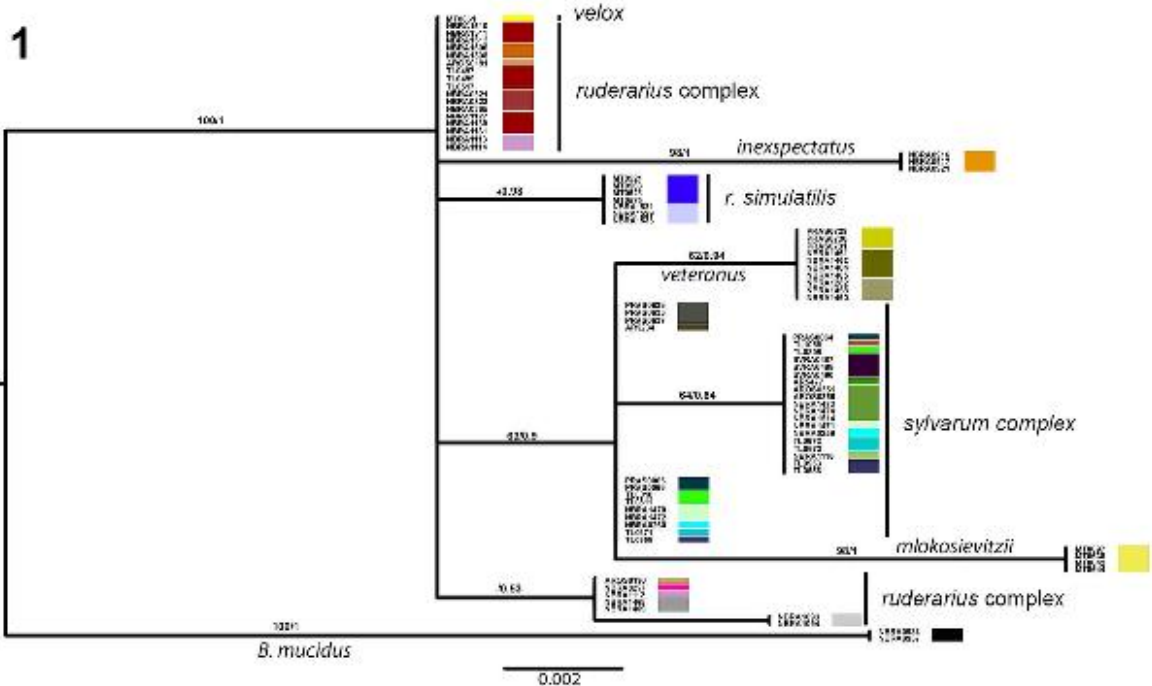
772

773

774

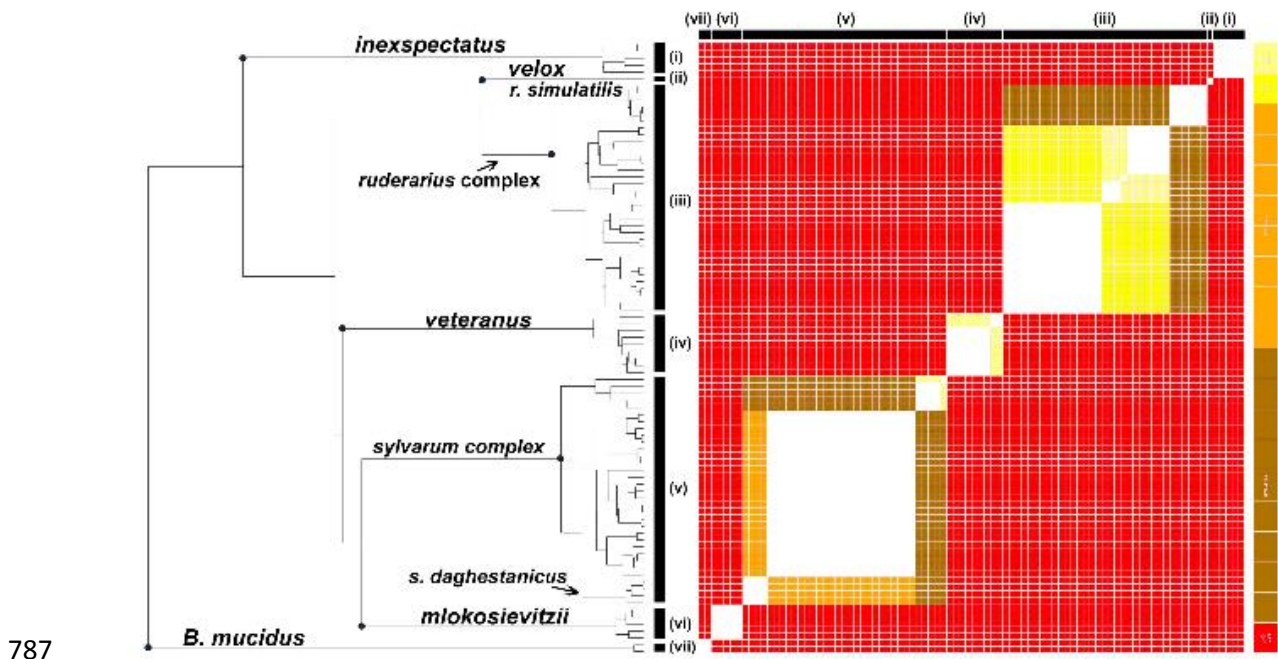
Figure 2. A: Majority rule (50%) consensus tree based on maximum likelihood analyses of COI. Values above tree branches are maximum likelihood bootstrap values/Bayesian posterior

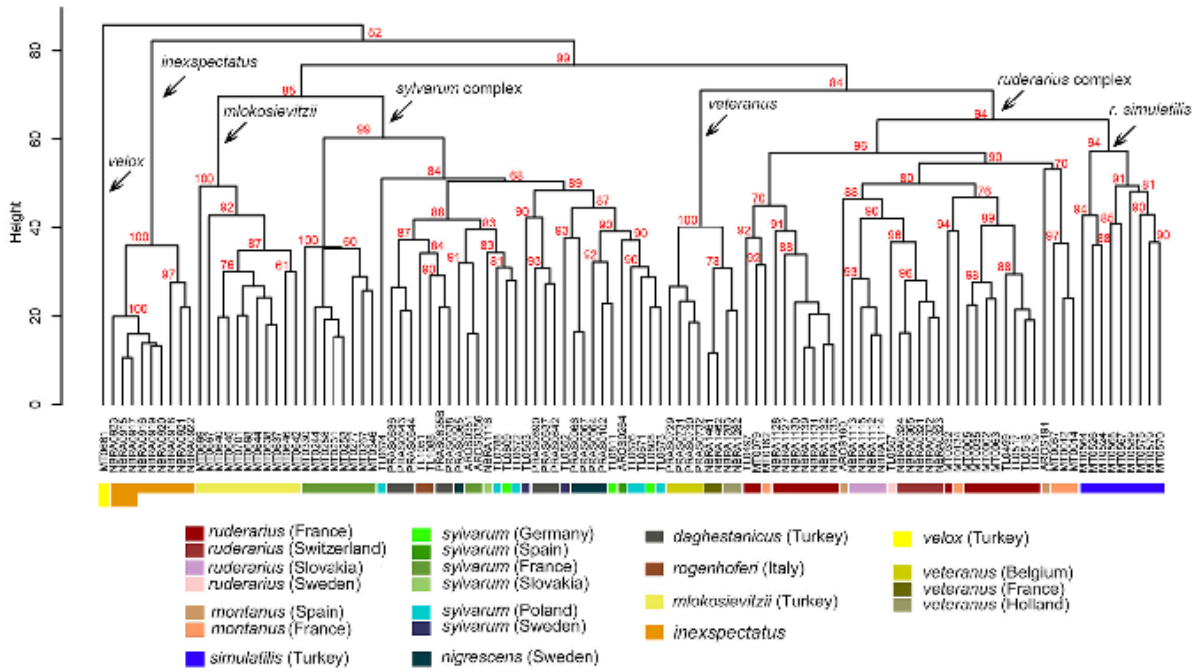
775 probabilities. **B**: Median-joining network of haplotypes based on cytochrome oxidase 1 (COI)
776 sequences. Numbers on the lines correspond to the number of mutation between haplotypes.
777



778
 779 **Figure 3. A:** Majority rule (50%) consensus tree based on maximum likelihood analyses of
 780 PEPCK. Values above tree branches are maximum likelihood bootstrap values/Bayesian

781 posterior probabilities. **B**: Median-joining network of haplotypes based on
782 phosphoenolpyruvate carboxykinase (PEPCK) sequences. This network do not consider the
783 Gap sites. **C**: Median-joining network of haplotypes based on phosphoenolpyruvate
784 carboxykinase (PEPCK) sequences. This network consider the Gap sites. Numbers on the
785 lines correspond to the number of mutation between haplotypes.
786





793

794 **Figure 5.** Unweighted pair group method with arithmetic mean (UPGMA) cluster based on a
 795 Canberra distance matrix calculated from the cephalic labial gland secretions matrix of *velox*,
 796 *veteranus*, *inexpectatus*, *runderarius*, *montanus*, *simulatilis*, *mlokosievitzii*, *sylvanum*,
 797 *daghestanicus*, *nigrescens* and *rogenhoferi* . The values near the node are multiscale bootstrap
 798 resampling.

799