



HAL
open science

A collaborative backbone resource for comparative studies of subterranean evolution: The World Asellidae database

Nathanaelle Saclier, Louis Duchemin, Lara Konecny-Dupré, Philippe Grison, David Eme, Chloé Martin, Cécile Callou, Tristan Lefébure, Clémentine François, Colin Issartel, et al.

► **To cite this version:**

Nathanaelle Saclier, Louis Duchemin, Lara Konecny-Dupré, Philippe Grison, David Eme, et al.. A collaborative backbone resource for comparative studies of subterranean evolution: The World Asellidae database. *Molecular Ecology Resources*, 2023, 24 (1), 10.1111/1755-0998.13882 . hal-04313500

HAL Id: hal-04313500

<https://hal.inrae.fr/hal-04313500>

Submitted on 23 Apr 2024

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 Title:

2 **A collaborative backbone resource for comparative studies of subterranean**
3 **evolution: the World Asellidae database**

4

5 Running title: Comparative study in subterranean evolution

6

7 Authors:

8 Nathanaelle Saclier ^{1,2}, Lara Konecny-Dupré ¹, Philippe Grison ³, Louis Duchemin ¹, David
9 Eme ⁴, Chloé Martin ³, Cécile Callou ³, Tristan Lefébure ¹, Clémentine François ¹, Colin Issartel
10 ¹, Julian J. Lewis ^{5,6}, Fabio Stoch ⁷, Boris Sket ⁸, Sanja Gottstein ⁹, Teo Delić ⁸, Maja Zgajmajster
11 ⁸, Michal Grabowski ¹⁰, Dieter Weber ^{11,12}, Ana Sofia P.S. Reboleira ^{13,14}, Dmitry Palatov ¹⁵,
12 Kaloust Paragamian ¹⁶, Lee R.F.D. Knight ¹⁷, Georges Michel ¹⁸, Francois Lefebvre ¹⁹,
13 Mohammad-Javad Malek Hosseini ^{20,21}, Ana I. Camacho ²², Begoña Gartzia De Bikuña ^{23,24},
14 Amina Taleb ²⁵, Nouria Belaidi ²⁵, Raoul P. Tuekam Kayo ²⁶, Diana Maria Paola Galassi ²⁷,
15 Oana Teodora Moldovan ^{28,29}, Christophe J. Douady ^{1,30*}, Florian Malard ^{1,*}

16

17 * Equal contribution

18

19 Correspondence: Florian Malard, Univ Lyon, Université Claude Bernard Lyon 1, CNRS,

20 ENTPE, UMR 5023 LEHNA, F-69622, Villeurbanne, France.

21 Email: Florian.Malard@univ-lyon1.fr

22

23

24

25 Authors' affiliations:

26

27 ¹ Univ Lyon, Université Claude Bernard Lyon 1, CNRS, ENTPE, UMR 5023 LEHNA, Ville-
28 urbanne, France

29 ² ISEM, CNRS, Univ. Montpellier, IRD, EPHE, Montpellier, France

30 ³ BBEES, Unité Bases de données sur la Biodiversité, Ecologie, Environnement et Sociétés,
31 Muséum National d'Histoire Naturelle, CNRS, Paris, France

32 ⁴ INRAE, UR-RiverLY, Lyon, France

33 ⁵ Virginia Museum of Natural History, Martinsville, Virginia, USA

34 ⁶ Lewis and Associates, Cave, Karst and Groundwater Biological Consulting, Borden,
35 Indiana, USA

36 ⁷ Evolutionary Biology & Ecology, Université libre de Bruxelles (ULB), Bruxelles, Belgium

37 ⁸ University of Ljubljana, Biotechnical Faculty, Department of Biology, SubBio Lab,
38 Jamnikarjeva 101, Ljubljana, Slovenia

39 ⁹ University of Zagreb, Faculty of Science, Department of Biology, Horvatovac 102A, 10000
40 Zagreb, Croatia

41 ¹⁰ Department of Invertebrate Zoology & Hydrobiology, Faculty of Biology & Environmental
42 Protection, University of Lodz, 90-237 Lodz, Poland

43 ¹¹ Musée National d'Histoire Naturelle de Luxembourg, 25 Rue Münster, L-2160
44 Luxembourg, Luxembourg

45 ¹² Senckenberg Deutsches Entomologisches Institut, Eberswalder Straße 90, 15374
46 Müncheberg, Germany

47 ¹³ Departamento de Biologia Animal, and Centre for Ecology, Evolution and Environmental
48 Changes & CHANGE – Global Change and Sustainability Institute, Faculdade de Ciências,
49 Universidade de Lisboa, Campo Grande 1749-016, Lisbon, Portugal

50 ¹⁴ Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15,
51 2100 Copenhagen, Denmark

52 ¹⁵ A.N. Severtsov Institute of Ecology and Evolution of the Russian Academy of Sciences,
53 Moscow, Russia.

54 ¹⁶ Hellenic Institute of Speleological Research, Megalou Alexandrou 179, 71306 IRAKLEIO,
55 Crete, Greece

- 56 ¹⁷ Hypogean Crustacea Recording Scheme, No.1, The Linhay, North Kenwood Farm, Oxton,
57 Nr Kenton, Devon, EX6 8EX, United Kingdom
- 58 ¹⁸ CWEPS, Commission Wallonne d'Etude et de Protection des Sites Souterrains, avenue G.
59 Gilbert 20, 1050 Bruxelles, Belgium
- 60 ¹⁹ SEPANSO Aquitaine, 1 rue de Tauzia, 33800 Bordeaux, France
- 61 ²⁰ Jovan Hadži Institute of Biology, Research Centre of the Slovenian Academy of Sciences
62 and Arts (ZRC-SAZU), Ljubljana, Slovenia
- 63 ²¹ Department of Organisms and Ecosystems Research, National Institute of Biology (NIB),
64 Ljubljana, Slovenia
- 65 ²² Museo Nacional de Ciencias Naturales (CSIC). Dpto. Biodiversidad y Biología Evolutiva.
66 C/ José Gutiérrez Abascal 2., 28006-Madrid, Spain
- 67 ²³ Anbiotek, Investigación científica y técnica del medio ambiente, Limnology Area
68 Directora, Ribera de Axpe 11. B201, 48950 Erandio, Bizkaia, Spain
- 69 ²⁴ Anbiolab, BIC Bizkaia Astondo bidea, edificio 612, Parque Científico y Tecnológico de
70 Bizkaia, 48160 Derio, Spain
- 71 ²⁵ Laboratoire d'Écologie et Gestion des Ecosystèmes Naturels, University of Tlemcen, Alge-
72 ria
- 73 ²⁶ University of Bamenda, Faculty of Science, Department of Zoology, POBox 39, Bambili,
74 Cameroon
- 75 ²⁷ Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila,
76 Italy
- 77 ²⁸ Emil Racovita Institute of Speleology, Clinicilor 5, 400006 Cluj-Napoca, Romania
- 78 ²⁹ Centro Nacional de Investigación sobre la Evolución Humana, Paseo Sierra de Atapuerca 3,
79 09002 Burgos, Spain
- 80 ³⁰ Institut Universitaire de France, Paris, France
- 81
- 82

83 **Abstract**

84 Transition to novel environments, such as groundwater colonization by surface organisms,
85 provides an excellent research ground to study phenotypic evolution. However, interspecific
86 comparative studies on evolution to groundwater life are few because of the challenge in
87 assembling large ecological and molecular resources for species-rich taxa comprised of surface
88 and subterranean species. Here, we make available to the scientific community an operational
89 set of working tools and resources for the Asellidae, a family of freshwater isopods containing
90 hundreds of surface and subterranean species. First, we release the World Asellidae database
91 (WAD) and its web application, a sustainable and FAIR solution to producing and sharing data
92 and biological material. WAD provides access to thousands of species occurrences, specimens,
93 DNA extracts and DNA sequences with rich metadata ensuring full scientific traceability.
94 Second, we perform a large-scale dated phylogenetic reconstruction of Asellidae to support
95 phylogenetic comparative analyses. Of 424 terminal branches, we identify 34 pairs of surface
96 and subterranean species representing independent replicates of the transition from surface
97 water to groundwater. Third, we exemplify the usefulness of WAD for documenting phenotypic
98 shifts associated with colonization of subterranean habitats. We provide the first
99 phylogenetically controlled evidence that body size of males decreases relative to that of
100 females upon groundwater colonization, suggesting competition for rare receptive females
101 selects for smaller, more agile males in groundwater. By making these tools and resources
102 widely accessible, we open up new opportunities for exploring how phenotypic traits evolve in
103 response to changes in selective pressures and trade-offs during groundwater colonization.

104

105 **KEYWORDS** collaborative database, phylogeny, comparative analysis, phenotypic
106 evolution, molecular resources, subterranean biodiversity

107 **1 | INTRODUCTION**

108 *Homo sapiens* has been fascinated by the subterranean world throughout its history (Mammola
109 and Martinez, 2020) and the peculiar features of subterranean organisms have attracted
110 scientists since their first discovery over the 16th to 17th centuries (Malard, 2022). However, it
111 was not until the mid-twentieth century that the idea that the subterranean world provides an
112 excellent research ground for addressing general scientific questions in ecology and evolution
113 gained momentum (Poulson and White, 1969; Mammola et al., 2020). A long-standing
114 perspective of subterranean life evolution is that of convergence whereby phylogenetically
115 distant organisms acquire similar phenotypes because of a convergent selective environment
116 that includes no light, environmental stability and energy limitation (Christiansen, 1961; Pipan
117 and Culver, 2012). Since the 2000's, a broader evolutionary perspective of subterranean life
118 has emerged, one that has also incorporated the role of non-adaptive processes (Lefébure et al.,
119 2017; Wilkens and Strecker, 2017; Policarpo et al., 2021) and divergent selection (Trontelj et
120 al., 2012; Fišer et al., 2023) in shaping the phenotype of organisms.

121

122 Phylogenetically controlled and replicated comparisons between closely related surface and
123 subterranean organisms provide ideal models to study evolution during colonization of a novel
124 environment (Protas and Jeffery, 2012; Saclier et al., 2018; Rétaux and Jeffery, 2023). Indeed,
125 surface organisms that colonize subterranean habitats experience dramatic environmental
126 changes (e.g. darkness, food limitation) and evolve characteristic regressive (e.g. reduced eyes
127 and pigment) and constructive (e.g. increased extra-optic sensory structures) traits (Culver and
128 Pipan, 2019; Hose et al. 2022). Subterranean colonization is considered an irreversible habitat
129 transition because it leads to eye degeneration (Niemiller et al., 2013; Langille et al., 2022).

130 Only in very rare cases, blind and depigmented animals can re-colonize surface habitats that
131 are characterized by low competitive pressure (Copilas-Ciocianu et al., 2018).

132

133 The scientific scope of surface-subterranean comparative studies ultimately depends on the
134 acquisition of ecological and molecular resources - from biological trait data to phylogenetic
135 and genomic resources - in model organisms. These resources are increasingly becoming
136 available at intraspecific level in species comprised of surface and subterranean populations
137 such as the teleost *Astyanax mexicanus* (Kowalko et al., 2020; Gross et al., 2023), the isopod
138 *Asellus aquaticus* (Konec et al., 2015; Protas et al., 2023), the amphipod *Gammarus minus*
139 (Fong et al., 2023) and the urodele amphibian *Proteus anguinus* (Kostanjšek et al., 2023).

140 However, comparative studies at the interspecific level remain scarce essentially because of the
141 difficulty in assembling large-scale phylogenetic and species trait data sets in clades comprised
142 of multiple surface and subterranean species (Stern et al., 2017; Lefébure et al., 2017; Saclier
143 et al., 2018; Mammola et al., 2019; Langille et al., 2022). Although intraspecific studies often
144 provide deeper insights into the genetic and developmental basis of phenotypic traits, only
145 interspecific studies can document evolutionary changes taking place over time periods longer
146 than the lifespan of natural populations.

147

148 Performing phylogenetic comparative analyses of clades comprised of surface and subterranean
149 species faces several challenges. First, only a few clades of metazoans contain both a high
150 number of surface species and subterranean species because the surface ancestors of many
151 subterranean taxa went extinct (Humphreys, 2000). Candidate clades often have a wide
152 geographic distribution, sometimes spanning several continents, which makes it particularly
153 difficult to obtain biological material (Mammola and Isaia, 2017; Faille, 2019; Fišer, 2019a;

154 [Lukić, 2019](#)). Second, the taxonomic units to be used in comparative analyses are not firmly
155 established. Molecular species delimitation methods often reveal highly divergent operational
156 taxonomic units within subterranean described species that have been historically delimited
157 based on morphological criteria ([Fišer et al., 2018](#); [Eme et al., 2018](#)). Third, we lack large dated
158 phylogenies of clades with multiple independent subterranean colonization events ([but see](#)
159 [Ledford et al., 2011](#); [Morvan et al., 2013](#); [Stern et al., 2017](#)). Last, when phylogenetic inferences
160 are available, biological traits for the taxonomic units of interest are often not available in the
161 literature and voucher specimens for measuring those traits are difficult to locate ([but see](#)
162 [Mammola et al., 2022](#)).

163

164 Here, we address the aforementioned challenges by releasing the World Asellidae database
165 (WAD) and phylogeny, a backbone resource to support comparative studies on life evolution
166 in subterranean habitats. The Asellidae (Isopoda, Pancrustacea) is one of the few families of
167 aquatic metazoans containing both surface and subterranean species, thereby potentially
168 providing multiple independent replicates of the transition from surface water to groundwater.
169 First, we describe the guiding principles and content of WAD, a collaborative database
170 specifically designed to promote the joint production and sharing of primary ecological and
171 molecular data and metadata by multiple research laboratories. Second, we take advantage of
172 new sequence data available in WAD for two mitochondrial genes and two nuclear genes to
173 perform a large-scale dated phylogenetic reconstruction of the Asellidae family that can be used
174 more widely in future comparative studies. Third, we exemplify the usefulness of WAD for
175 documenting phenotypic changes associated with colonization of subterranean habitats. We use
176 the Asellidae phylogeny and body size (BS) data from literature articles and morphological
177 measurements made on specimen lots referenced in WAD to test for differences in male and

178 female BS between surface and groundwater habitats. We predict no difference in female BS
179 between habitats because fecundity selection probably favors large-bodied females with large
180 brood sizes in both habitats. In contrast, we predict smaller-bodied males in groundwater than
181 in surface water due to a shift in male mating strategy. In surface water, we hypothesize that
182 competition for synchronously receptive females selects for large males that are more likely to
183 win mating contests (Bertin and Cezilly, 2003). In groundwater, competition for rare, highly
184 asynchronous, receptive females potentially favors smaller, more agile males that are more
185 likely to be successful in finding mates (Andersson, 1994; Blanckenhorn, 2000).

186

187 2 | MATERIALS AND METHODS

188 2.1 | The World Asellidae Database (WAD)

189 We use the free and open-source application GOTIT (<https://github.com/gotit-dev/gotit>; Malard
190 et al., 2020) to input, manage and share ecological and molecular data and metadata in WAD.
191 The application manages every step of an every-day laboratory workflow process leading to the
192 production of species occurrence data and DNA sequences. A demo version of GOTIT
193 application is available at <https://gotit.cnrs.fr>. WAD hosts all species occurrence data, sampling
194 and sequencing metadata and biological vouchers (specimens, microscopic slides and DNA
195 extracts) generated over the workflow (Table 1). The database also manages species occurrence
196 data and DNA sequence metadata from the literature and biodiversity facilities, the
197 bibliographic referencing of information and the assignment of DNA sequences to molecular
198 operational taxonomic units (MOTUs). We provide in supplemental figures 1 and 2 (SI Figures
199 1 and 2) the simplified and full logical models of the database, respectively. User access to
200 WAD, either as a data end-user or contributor, is at <https://gotit.univ-lyon1.fr> upon request from
201 the corresponding author.

202

203 2.2 | Species delimitation and dated phylogeny

204 2.2.1 | Taxon sampling and molecular data

205 To build the phylogeny, we extracted from WAD an initial molecular data set representing 299
206 described and undescribed morphospecies of Aselloidea (278 Asellidae and 21 Stenasellidae
207 used as outgroup). Specimens were collected at 943 localities in Europe, North America, North
208 Africa and Asia (SI Table 1). Localities spanned a wide range of surface and subsurface fresh-
209 water habitats including lotic and lentic surface water bodies, cave streams and pools, the
210 hyporheic zone of surface streams and groundwater in unconsolidated sediments. Throughout
211 this paper, we used the term morphospecies to refer to species, either formally described or
212 undescribed (i.e. waiting a formal description), that were identified based on morphological
213 criteria. Species names of North American asellids follow the latest taxonomic revision to be
214 published by Lewis and coauthors (2023). For morphological identification of specimens to
215 species level, we relied on the shape of male copulatory organs (pleopods 2), plus secondary
216 characteristics including the morphology of the male pereopods 1 and 4, pleopods 3-5, and
217 uropods (Lewis et al., 2023). We dissected copulatory pleopods 1 and 2 of male specimens and
218 mounted them on slide for examination using a compound microscope.

219

220 We used the Chelex protocol of Casquet and coauthors (2012) to extract DNA from specimen.
221 We incubated three pereopods of each specimen in a solution of 150 µl of 7% chelex and 10 µl
222 of proteinase K at 15 mg / ml for 90 minutes at 56 °C, and then 15 minutes at 90 °C. We
223 amplified DNA using primers targeting the mitochondrial cytochrome oxidase subunit I (COI)
224 gene, the 16S mitochondrial rDNA gene, the FASTKD4 nuclear gene and the 28S nuclear
225 rDNA gene. We provide in supplemental tables 2 and 3 (SI Tables 2 and 3) the list of all PCR

226 primers, among which 66 were specifically designed as part of this study. For the two
227 mitochondrial genes, we applied several methods to prevent misleading inclusion of nuclear
228 mitochondrial DNA segments (NUMT) in the data set, including different primer pairs, long-
229 range amplification and pre-PCR dilution of genomic DNA (Calvignac et al., 2011).

230

231 We amplified 16S fragments with two independent pairs of primers (SI Tables 2 & 3). PCR
232 settings were as follows: one step of 3 min at 95 °C; 35 cycles of 20 s at 95 °C, 30 s at 53 °C,
233 30 s at 72 °C; and one step of 5 min at 72 °C. We performed PCRs for COI fragments using a
234 previously optimized protocol (Calvignac et al., 2011), but with a Taq polymerase (Eurobiotaq)
235 amount of 0.05 U instead of 0.15 U and a PCR volume of 25 µl instead of 35 µl. We used the
236 following PCR settings: one step of 3 min at 95 °C, 37 cycles of 20 s at 95 °C, 30 s at 51 °C,
237 45 s at 72 °C, and one step of 5 min at 72 °C. A semi-nested PCR was performed whenever the
238 first amplification failed. Using the first PCR product as DNA template, we performed a second
239 PCR using one of the two primers used in the first PCR and another, different primer. The
240 second round PCR was run on 1 µl of the first round PCR product, using the same settings as
241 above but 35 cycles. We amplified FASTKD4 fragments using several pairs of primers (SI
242 Tables 2 & 3) with the following PCR settings: one step of 5 min at 95 °C, 38 cycles of 30 s at
243 95 °C, 45 s at 54 °C, 45 s at 72 °C, and one step of 5 min at 72 °C. As for the COI gene, we
244 performed a semi-nested PCR whenever the first amplification failed. We completed PCRs for
245 28S fragments with two independent pairs of primers in order to detect divergent copies. We
246 used the following PCR settings: one step of 3 min at 95 °C; 37 cycles of 30 s at 95 °C, 30 s at
247 62 °C, 30 s at 72 °C; and one step of 5 min at 72 °C.

248

249 Microsynth France SAS (Vaulx-en-Velin, France) performed Sanger sequencing for the four
250 genes. Chromatograms were visualized with FinchTV (Geospiza, Seattle, WA, USA). All
251 sequences were aligned with Muscle as implemented in Seaview (Gouy et al., 2010) and
252 checked visually for the presence of anomalies, including stop codons and frameshifts for
253 protein coding genes.

254

255 2.2.2 | Molecular operational taxonomic units (MOTUs)

256 We delimited MOTUs based on a COI alignment of 1385 haplotypes, which were defined from
257 the sequences obtained from 2093 specimens belonging to 299 morphospecies of Aselloidea
258 (SI Tables 1 & 4). We used the following procedure to select specimens for which we obtained
259 COI sequences. Whenever possible, we first obtained 16S sequences from three specimens of
260 each morphospecies present at a site. Second, we obtained a COI sequence for each specimen
261 whose 16S sequence differed by more than 5 nucleotides with any 16S sequence of the two
262 other specimens.

263

264 We used the fixed COI threshold method (TH) implemented by Lefébure and coauthors (2006)
265 for crustaceans, and the Poisson tree processes (PTP) proposed by Zhang and coauthors (2013),
266 to delimit MOTUs. The TH method was previously used in several studies for delimiting
267 species of asellids (Morvan et al., 2013; Eme et al., 2013, 2018). It is based on the observation
268 made from hundreds COI sequences of crustaceans that two clades diverging by more than 0.16
269 substitution per site, as measured by patristic distances, have a strong probability (ca. 0.99%)
270 of belonging to different described morphospecies. It is a conservative method insofar as it
271 identifies both fewer MOTUs and MOTUs that are more divergent than tree-based methods
272 such as the PTP method (Eme et al., 2018). We applied the TH and PTP methods on a COI

273 haplotype alignment in which the longest sequence with the fewest ambiguities was retained as
274 the best representative sequence for any given haplotype. We constructed a COI haplotype
275 phylogeny in maximum likelihood with PhyML (Guindon et al., 2010) using a GTR + G + I
276 model of evolution and Stenasellidae species as outgroup. We computed patristic distances
277 from this phylogeny with the R package “ape” (Paradis et al., 2004) and delimited MOTUs
278 according to the TH method with the “cluster” package (Maechler et al., 2002). To delimit
279 MOTUs with the PTP method, we ran mPTP v0.2.2 (<https://github.com/Pas-Kapli/mptp>) using
280 400 000 MCMC generations, with a thinning of 400 and 0.1 (10%) burn-in.

281

282 We performed pairwise taxonomic comparisons between the three different sets of species
283 hypotheses delimited using morphology, the TH method, and the PTP method. For each
284 pairwise comparison, we provided the number of species delimited by each of the two methods
285 as well as the number of matches, splits, lumps and reshuffling (see Eme et al. (2018) for a
286 definition of these four categories).

287

288 2.2.3 | Four-gene alignment and phylogeny

289 We produced alignments of the COI, 16S, FASTKD4 and 28S genes for 424 MOTUs of
290 aselloids delimited with the TH method (SI Tables 5 to 9). Using MOTUs delimited with the
291 TH method rather than the PTP method limited the risk of considering two populations of the
292 same species as belonging to two distinct MOTUs. In each alignment, we retained the longest
293 sequence with the fewest ambiguities to represent each MOTU, using the chimera assembler
294 script (https://github.com/TristanLefebure/chimera_assembler.pl). We aligned the 28S and 16S
295 genes with MAFFT Q-INS-i using default parameters (Katoh and Standley, 2013) and the COI
296 and FASTKD4 genes with PRANK codon (Löytynoja and Goldman, 2008). Sites ambiguously

297 aligned were removed with Gblocks (Castresana 2000). We used the four genes for 424 MOTUs
298 to build a phylogeny with PhyloBayes (Lartillot et al., 2009) under a CAT-GTR model of
299 evolution. To guarantee the absence of polytomy, a threshold of 10% was set to obtain the
300 majority consensus tree, meaning that each clade must be found in at least 10% of the trees of
301 the Markov process after burn-in. We computed *posterior probabilities* to estimate the support
302 of tree topologies and rooted the tree using species of Stenasellidae as outgroup.

303

304 Using the phylogeny, we identified pairs of surface and groundwater asellid species that
305 provided independent replicates of the ecological transition from surface water to groundwater.
306 We ensured independence among pairs by selecting them so that the tree paths from one species
307 to the other within a pair did not contain any branches in common with any other pairs
308 (Felsenstein, 1988). For comparison with intraspecific studies, using independent species pairs
309 is statistically more robust than using replicate pairs of surface and groundwater populations
310 within a single species (Rétaux and Jeffery, 2023). Indeed, replicate populations pairs within
311 species can be statistically dependent if gene flow still occurs among surface populations.

312

313 2.2.4 | Time-scale phylogeny

314 In the absence of fossil records for the Aselloidea, we used well-identified paleobiogeographic
315 events to constrain the age of 17 nodes in the phylogeny (see SI Table 10 for a description of
316 these events). Paleobiogeographic calibration points spanned a period ranging from 300 to 2
317 Myr before present. We estimated divergence times among aselloids with PhyloBayes using a
318 CAT-GTR + G + I model, the 17 calibration points as soft bounds, a birth-death prior on
319 divergence time and a log-normal auto-correlation of the substitution rates among branches

320 (Lepage et al., 2007). The effect of any given calibration point on divergence time estimates
321 was assessed by removing that given calibration point during time tree reconstruction.

322

323 2.3 | Comparative phylogenetic analyses of body size

324 2.3.1 | Body size and sexual body size dimorphism

325 Here, we provide a case study of body size and sexual body size dimorphism to show how
326 WAD resources and the World Asellidae phylogeny allow exploring how phenotypic traits
327 evolve upon groundwater colonization. We completed literature data with laboratory
328 measurements made on specimen lots contained in WAD to quantify the maximum body size
329 of adult males and females of 162 asellid MOTUs included in the World Asellidae phylogeny
330 (SI Table 11). We defined body size as the distance between the anterior margin of the cephalon
331 and the posterior margin of the pleotelson (Prevorčnik et al. 2004). Maximum body size
332 (subsequently abbreviated to BS) provides an estimator of the size of full-grown specimens in
333 a species: it avoids including immature specimens and is often the only measurement provided
334 in publications. For each MOTU, we provide in SI Table 11 our best estimate of the number of
335 specimens used for quantifying BS, as the exact number is not always reported in the source
336 articles. For measurements made on specimen lots contained in WAD, we took photos of
337 specimens with a DP25 Olympus camera connected to a dissecting microscope (SZX16
338 Olympus) and measured BS using ImageJ (Schneider et al., 2012).

339

340 To quantify sexual body size dimorphism (SBSD), we used the size dimorphism index (SDI)
341 as follows (Lovich and Gibbons, 1992; Fairbairn, 2007):

$$342 \text{SDI} = \frac{\text{Body size of largest sex}}{\text{Body size of smallest sex}} - 1$$

343 SDI equals zero for monomorphic species in which the two sexes have the same body size and
344 is arbitrarily given a negative sign when males are larger than females.

345

346 2.3.2 | Habitat specialization and habitat size

347 We used presence and absence of eyes and body pigment as evidence of specialization to
348 surface water and groundwater habitats, respectively. Hence, in the ensuing text, groundwater
349 species designate eyeless and depigmented species whereas surface water species designate
350 ocellated and/or pigmented species. Of the 162 MOTUs included in the phylogenetic
351 comparative analyses (see below), 61 were surface water species and 101 were groundwater
352 species.

353

354 We assessed the size of habitat or pore volume available to species because it is potentially a
355 major determinant of maximum BS (Pipan and Culver, 2017). We used a fuzzy coding approach
356 (Chevenet et al., 1994; Degen et al., 2018) to assess habitat size because most groundwater
357 ecological studies do not provide any quantitative estimates of pore volume available to species.
358 For the 162 asellid MOTUs incorporated in the comparative analyses, we attributed positive
359 scores (from 0=no affinity to 3=strong affinity) to three categories of habitat size (large,
360 medium and small pore volumes). We attributed habitat size scores independently of the eye
361 and pigmentation status of species. Hence, we assigned a high affinity for large size habitats to
362 species living in the benthic layer of surface streams as well as to those living in the benthic
363 layer of cave streams. Scores were attributed to all species separately by two of us (F.M. and
364 J.J.L.) using species occurrence data per habitat category as guideline data (data extracted from
365 WAD). Then, we corrected for inconsistencies between the two sets of scores to produce a
366 single “habitat trait categories per species” matrix. We provide the species habitat scores and

367 the scoring procedure in supplemental table 11 (SI Table 11) to ensure data traceability and
368 reproducibility, and potential revision of scores in the event of new habitat data of species.
369 Then, we performed a fuzzy correspondence analysis (COA) of the “habitat trait categories per
370 species” matrix (Chevenet et al., 1994) and used the coordinates of species along the first axis
371 of the COA, representing 85% of total variability, as quantitative surrogates of their habitat size.
372 The COA was performed using the R package ”ade4” (Thioulouse et al., 2018).

373

374 2.3.3 | Data analysis

375 We performed phylogenetic generalized least-squares (PGLS) regression models (Martins and
376 Hansen, 1997) to test for the effect of habitat specialization and habitat size and its interaction
377 on BS of females and males and SDI. To account for phylogenetic non-independence among
378 species, we used the Asellidae time-scale phylogeny, pruned to the 162 MOTUs for which BS
379 data were available for the two sexes. We selected the best model of trait evolution and its
380 associated covariance structure - in this study, the Brownian motion model - according to
381 minimum Akaike information criterion. We tested the significance of each predictor (i.e. habitat
382 specialization and habitat size) in the regression by comparing with a likelihood ratio test (LRT)
383 a model without the predictor to a model with the predictor. We assessed the proportion of
384 variance explained by phylogenetic regressions using Cox-Snell pseudo-R². PGLS were
385 performed in R using the “APE” (Paradis et al., 2004) and “nlme” (Pintero et al., 2022)
386 packages.

387

388 3 | RESULTS

389 3.1 | The World Asellidae Database (WAD)

390 The database contains 9438 distributional records for 163 surface water species and 285
391 groundwater species of Asellidae belonging to 23 genera (Tables 1 and 2). Asellids are widely
392 distributed in the Northern Hemisphere with species belonging to four formerly recognized
393 groups of morphospecies, which occupy distinct but partially overlapping distribution ranges
394 (Figure 1). All four groups include both surface and groundwater species, although in different
395 proportions. (Table 2). The first group is the "Asellus pattern", so named by Henry and Magniez
396 (1995) in reference to the specific shape of copulatory organs shared by several genera of
397 Asellidae. It has nine genera (61 species); all distributed in Asia and North Western America,
398 except the genus *Asellus*, which is also represented in Europe by the widespread *Asellus*
399 *aquaticus* species complex (Verovnik et al., 2005). The second group to which we refer as the
400 North American asellids include seven genera (152 described species), all located in North
401 America, except *Gallasellus* and *Baicalasellus*, which are endemic to western France and Lake
402 Baikal (Russia), respectively. The third group containing the two genera *Bragasellus* and
403 *Synasellus* (56 species) is endemic to the Iberian Peninsula. The fourth group corresponding to
404 the genus *Proasellus* (174 species) extends from southern Scandinavia to northern Africa and
405 from Portugal to Iran.

406

407 WAD describes the content of 1943 specimen lots, which were sampled by 324 collectors in 38
408 countries. Lot description comprises the number of male and female mature specimens,
409 juveniles and ovigerous females. Collection material referenced in WAD also includes 4362
410 specimen DNA extracts and 1584 specimen microscopic slides. Specimen lots and DNA
411 extracts are preserved at -20°C in the zoology collection at University Claude Bernard of Lyon:
412 they are available for subsequent collaborative morphological and molecular analyses upon
413 request from the corresponding author.

414

415 WAD provides metadata – from sampling of specimens to PCR and chromatogram settings –
416 for 8914 validated sequences of Asellidae belonging to two mitochondrial and three nuclear
417 genes (Table 1). Of these, 3692 sequences were submitted to NCBI, essentially from the present
418 article authors, as part of previous studies, and 4082 sequences were submitted as part of the
419 present study (SI Table 12). In WAD, COI sequences are assigned to MOTUs using different
420 molecular species delimitation methods. The geographic distribution of MOTUs within
421 morphospecies can be visualized using ready-to-use queries implemented in GOTIT application
422 (SI Figure 3).

423

424 3.2 | The Asellidae timetree

425 3.2.1 | Molecular operational taxonomic units (MOTUs)

426 The TH and PTP molecular species delimitation methods provided respectively, 1.6 and 1.9
427 more MOTUs than morphospecies (Figure 2). The two molecular methods essentially split
428 morphospecies into smaller clusters of individuals. Reshuffling cases were rare: of the 466 and
429 557 MOTUs respectively delimited by TH and PTP, only 10 (2.1%) and 12 (2.2%) fell in that
430 category. PTP split morphospecies into smaller clusters than TH, thereby generating 1.2 more
431 species hypotheses than TH.

432

433 3.2.2 | Time-scale phylogeny

434 The phylogeny included 384 MOTUs of asellids delimited with the TH method. They
435 collectively represented 268 morphospecies, among which 195 were formally described (Figure
436 3, Table 2). The tree topology recovered the monophyly of the four main species groups
437 described above and that of all asellid genera, with posterior probabilities > 0.9 , except the

438 genus *Conasellus* ($PP=0.43$) (Figure 3, SI Figure 4). The *Asellus* pattern (group 1 in Figures 1
439 and 3) formed a sister clade to the rest of a larger clade comprised of the North American asellid
440 clade (group 2), the *Bragasellus* + *Synasellus* clade (group 3), and the *Proasellus* clade (group
441 4). However, relationships among the later three clades were not resolved. The tree topology
442 for *Proasellus* was also consistent with earlier subdivisions of this species-rich genus into four
443 clades (Morvan et al., 2013). Within that genus, the slavus clade was sister to a larger clade
444 comprised of the ibero-aquitanian, anophthalmus-coxalis, and Alpine clades, but the
445 relationships among the later three clades were not resolved (Figure 3).

446

447 Divergence time estimates were robust to the removal of any single paleobiogeographic
448 calibration point, except the deepest one that constrained the divergence between Stenasellidae
449 and Asellidae to be more recent than 300 Myr (SI Figure 5). Removing this point yielded older
450 divergence times, notably pushing back the divergence between the Stenasellidae and Asellidae
451 to 300 Myr (95% Credibility Interval [CI]: 415-222 Myr) instead of 139 Myr (CI: 174-106
452 Myr), when including it (Figure 3, SI Figure 5). The diversification of Asellidae might have
453 started in early Cretaceous (132 Myr, CI: 168-102 Myr) and that of *Proasellus* at the end of
454 Cretaceous or beginning of the Paleogene (72 Myr, CI: 88-58 Myr).

455

456 We identified up to 34 independent pairs of surface and groundwater asellid species in the
457 phylogeny (Figure 3, SI Figure 4). Species pairs were present in all four major groups of asellid
458 species, although several species-rich clades were almost exclusively comprised of
459 groundwater species, including the Alpine *Proasellus* clade, *Synasellus* and *Caecidotea*. The
460 uneven distribution of species pairs among the *Proasellus* (21 pairs), North American asellids
461 (10 pairs) and the *Asellus* pattern (one pair) essentially reflected a too low sampling in the latter

462 two groups (Table 2). Transitions to groundwater have probably occurred throughout the
463 evolutionary history of the Asellidae (Figure 3, SI Figure 4). Using the speciation event leading
464 to a congeneric species pair as a surrogate of the transition time to groundwater (but see
465 discussion), some transitions occurred less than 10 million years ago (6.9, CI: 11.4 -4.2 Myr for
466 the *Asellus aquaticus* – *A. kosswigi* species pair), whereas others potentially occurred much
467 longer ago (at most 40.4, CI: 50.9 - 30.4 Myr for the *Conasellus burkensis* - *Conasellus reddelli*
468 species pair) (SI Figure 4).

469

470 3.3 | Comparative phylogenetic analyses of body size

471 Female and male BS ranged from 2.3 to 18 mm (mean=6.3±2.7 mm, n=162 MOTUs) and from
472 2.1 to 25 mm mean=7.3±4.1 mm, n=162 MOTUs), respectively (SI Table 11, SI Figure 6). BS
473 increased significantly with habitat size, both for females and males (Table 3, Figure 4, SI Table
474 13). Species colonizing open habitats, both above (e.g. surface lakes and streams) and below
475 ground (e.g. cave streams), had larger BS than species colonizing interstitial habitats (e.g.
476 groundwater in unconsolidated sediment). The effect of habitat specialization on BS was gender
477 dependent (Table 3, Figure 4, SI Table 13). Male BS was significantly smaller in groundwater
478 species than in surface water species, whereas we found no significant differences in female BS
479 between surface water and groundwater species. However, habitat specialization accounted for
480 a smaller proportion of variance in male body size (Cox-Snell $R^2 = 0.074$) than habitat size (R^2
481 = 0.161) (SI Table 13). We found no interactions between the effects of habitat size and habitat
482 specialization on male and female body size, indicating that constraints imposed by the size of
483 habitats on body size applied similarly to eyeless and depigmented species and oculated and/or
484 pigmented species.

485

486 Asellids showed substantial variation in sexual body size dimorphism (SBSD) among species.
487 Of the 162 species examined in this study, 94 (58 %) exhibited male-biased dimorphism, 62
488 (38.3 %) exhibited female-biased dimorphism and 6 (3.7 %) were monomorphic for BS. We
489 found a significant effect of habitat specialization on SBSD: mean SDI was -0.33 ± 0.22 ($n=61$)
490 and 0.01 ± 0.28 ($n=101$) for surface water species and groundwater species, respectively (Table
491 3, Figure 4, SI Table 13). Males were larger than females in 57 of 61 (i.e. 93.4 %) surface water
492 species examined in this study, whereas they were larger than females in only 37 of 101 (i.e.
493 36.6 %) groundwater species. The size dimorphism index (SDI) decreased significantly with
494 increasing habitat size (Table 3, Figure 4). Male-biased SBSD ($SDI < 0$ in Figure 4)
495 predominated in open habitats whereas female-biased SBSD (i.e. $SDI > 0$) predominated in
496 interstitial habitats. However, habitat size accounted for a smaller proportion of variance in
497 SBSD (Cox-Snell $R^2 = 0.069$) than habitat specialization ($R^2 = 0.110$). We found no interactions
498 between the effect of habitat size and habitat specialization on SBSD.

499

500 4 | DISCUSSION

501 4.1 | The World Asellidae Database (WAD)

502 Collaborative databasing has become essential to biodiversity sciences because the amount of
503 data and biological material needed to address broad-in-scope questions exceeds the production
504 capabilities of even the most performing laboratories (Nelson et al., 2011; Hobern et al., 2012;
505 Fišer, 2019b). The structure of the database used in the present study and its web application
506 GOTIT have been conceived to provide scientists with an efficient tool to jointly produce
507 multiple-species ecological and molecular resources to study life evolution in groundwater. The
508 tool has been used with success since 2017 to amass worldwide data at an unprecedented rate
509 for the Asellidae. WAD provides to date one of the most important resource of species

510 occurrence, DNA sequences and biological material for testing eco-evolutionary hypotheses
511 pertaining to groundwater colonization using comparative phylogenetic methods and
512 evolutionary model fitting (Stern et al., 2017; Lefébure et al., 2017; Saclier et al., 2018; Langille
513 et al., 2022).

514

515 Beyond Asellidae, the tool offers several desirable features when collaboratively producing
516 species occurrence and sequencing data (Malard et al., 2020). First, the database structure
517 portrays a standard workflow - from field sampling to DNA sequencing - that is common to
518 many laboratories. Second, a user-friendly web application allows implementing that laboratory
519 workflow on a day-by-day basis while simultaneously feeding a centralized database. Third,
520 the database guarantees scientific repeatability by offering a full traceability of field and
521 laboratory protocols and biological vouchers. Fourth, intellectual property rights and citation
522 issues are resolved in a way to encourage information sharing before publishing. Sequence
523 metadata are available to all as DNA sequence production flows, hence well before publicly
524 releasing the latter. Sharing metadata before publishing data is key to minimize duplication of
525 work among producers, thereby promoting sustainable data production. Fifth, the database
526 structure and its web application are free and open-source, so that the developer community can
527 modify the source code to address new user requirements. Four updates of the tool have been
528 released since its publication in 2019, with the last update containing a user-friendly query
529 builder for non-SQL experts to extract large data sets (<https://github.com/gotit-dev/gotit>).

530

531 Current development efforts are following two main directions. The first direction is widening
532 the database structure for housing biological species trait data, including but not limited to
533 morphological traits, which are measured on referenced specimens (see for example Lefébure

534 *et al. (2017)* and *Saclier et al. (2018)* for data on genome size and rate of molecular evolution,
535 respectively). The second direction consists in providing user-friendly tools to promote
536 expertise sharing among users. One example expert tool could guide sequence producers in
537 selecting the most appropriate primers for sequencing any given species from the hundred
538 primers available in the database (see [SI Tables 2 and 3](#)).

539

540 **4.2 | Species delimitation and dated phylogeny**

541 The presence of highly genetically divergent units (i.e. MOTUs) – often referred to as cryptic
542 species - within morphospecies is a common phenomenon across most animal taxa (*Bickford*
543 *et al., 2007*; *Pfenninger and Schwenk, 2007*), and asellids do not escape the rule (*Eme et al.,*
544 *2013*; *Morvan et al., 2013*). Hence, species molecular delimitation methods based on the COI
545 gene typically provide more species hypotheses – in the present study 1.6 to 1.9 times as many
546 – than morphological delimitation. Although molecular methods typically split asellid
547 morphospecies into several MOTUs, they very rarely reshuffle MOTUs among morphospecies.
548 Using different elementary species units in biodiversity research can provide novel insights into
549 the mechanisms underlying biodiversity patterns (*Fišer et al., 2018*). In their analysis of the
550 range size pattern of groundwater Asellidae and Niphargidae (Amphipoda) in Europe, *Eme and*
551 *coauthors (2018)* showed that using MOTUs instead of morphospecies reinforced the Rapoport
552 effect of increasing range size at higher latitudes and increased the proportion of variance in
553 range size explained by historical climates. In WAD, we are continuously updating the
554 geographic coverages of COI sequences and MOTUs within morphospecies (see [SI Figure 3](#)),
555 thereby accumulating data for rigorously testing the hypothesis that groundwater species have
556 a reduced range size compared to their surface counterparts. Despite being a long-standing
557 hypothesis (*Malard et al., 2023*), the crayfish study by *Stern and coauthors (2017)* remains the

558 only phylogenetically controlled test to date, even though the authors used morphospecies
559 rather than MOTU-level data. WAD also provides one of the most comprehensive reference
560 barcode libraries of groundwater taxa for accurately assigning to existing known species the
561 COI sequences that arise from a growing number of DNA-based biodiversity studies
562 (Zagmajster et al., 2022). Such a WAD reference barcode library offers great opportunities to
563 combine environmental DNA sampling, metabarcoding, DNA taxonomy and traditional
564 taxonomy to speed up the acquisition of species occurrence data in difficult-to-access
565 groundwater habitats (Fontaneto, et al., 2015; Saccò et al., 2022; Verdier et al., 2022).

566

567 The World Asellidae phylogeny provides one of the most comprehensive phylogenetic
568 frameworks available to date for undertaking comparative studies on evolution to groundwater
569 life (but see also Stern et al., 2017). Here, we highlight key improvements to the phylogeny
570 since a previous version published by Morvan and coauthors (2013). First, the present version
571 of the phylogeny contains 2.5 and 2.4 times more MOTUs and morphological species of asellids
572 respectively, than its previous version. Its geographical coverage is also considerably wider, as
573 it includes not only European species but also many North American and eastern Mediterranean
574 species. Yet, the phylogeny is far from being complete since it presently contains 60 % of
575 described species of asellids, the most species-deficient group being the *Asellus* pattern in Asia
576 with only 13 % of described species included in the phylogeny. Second, we improved dating of
577 divergence times in the phylogeny by adding 14 paleobiogeographic calibration points to the
578 three points originally used by Morvan and coauthors (2013). This addition resulted in overall
579 younger divergence times. Thus, in the present phylogeny, the early diversification of the four
580 *Proasellus* clades is dated to the Paleogene and not to the Upper Jurassic, as estimated by
581 Morvan and coauthors (2013). However, paleobiogeographic calibration points are still

582 relatively unevenly distributed across the phylogeny, with only a single point for the North
583 American, albeit species-rich, clade. Adding new calibration points to this clade would require
584 sampling US regions where species-rich clades might have diversified "on place" following
585 emergence of lands from the sea (e.g., eastern Texas, Florida and Chesapeake Bay). Third, still
586 in comparison with [Morvan and coauthors \(2013\)](#)'s phylogeny, we more than doubled the
587 number of replicates of groundwater evolution by identifying 34 independent pairs of surface
588 and groundwater asellid species, among which 21 within the genus *Proasellus*. Further
589 sampling will likely provide additional species pairs within the *Asellus* pattern and North
590 American asellids, thereby providing a more even distribution of groundwater transitions
591 among three of the four major groups of asellids. Obtaining many replicate species pairs is
592 crucial to robust testing of common principles of groundwater evolution while accounting for
593 the effects of local contingencies. Up to now, comparative studies have relied on few replicates
594 of evolution to groundwater life - i.e., on 3 to 13 independent species pairs - for assessing
595 changes in the evolution of genome size and rate of molecular evolution in asellids ([Lefébure](#)
596 [et al., 2017](#); [Saclier et al., 2018](#)), vision genes in beetles, crayfishes, and fishes ([Stern and](#)
597 [Crandall, 2018](#); [Policarpo et al., 2021](#); [Langille et al., 2022](#)), and gene repertoires in beetles
598 ([Balart-García et al., 2023](#)).

599

600 Another desirable attribute of a biological study system for understanding trait evolution in
601 groundwater is to have species that have colonized groundwater for different lengths of time.
602 Time is undoubtedly an important factor controlling the evolution of traits, at least those that
603 evolve under relaxed selection, such as the regression of eyes in subterranean animals ([Wilkens](#)
604 [and Strecker, 2017](#); [Policarpo et al., 2021](#); [Langille et al., 2022](#)). Among the asellids,
605 depigmented and reduced-eye subterranean populations of the surface species *Asellus aquaticus*

606 colonized groundwater less than one hundred thousand years ago (Protas and Jeffery, 2012;
607 Protas et al., 2023), whereas some eyeless and depigmented species of *Proasellus* have resided
608 in groundwater for over 10 million years (Lefébure et al., 2017). However, in a phylogeny, it is
609 usually unclear at which point along a terminal branch leading to a groundwater species
610 colonization of groundwater exactly occurred. Specifically, groundwater colonization may be
611 much more recent than the speciation event leading to a pair of surface and groundwater species
612 if now-extinct surface species have persisted long after that speciation event. In asellids, a
613 promising approach is to use the pseudogenization of genes coding for opsin light-sensitive
614 proteins to estimate the groundwater colonization time, assuming that loss-of-function
615 mutations accumulate early in the process of groundwater colonization. In a study by Lefébure
616 and coauthors (2017), colonization time was measured for 19 asellid species as a function of
617 the speciation time and an estimate of the pseudogenization of the opsin genes on branches
618 leading to subterranean species. Increasingly sequencing the opsin genes across asellid species
619 (see Table 2) paves the way for accounting for the effect of colonization time on the evolution
620 of phenotype in comparative studies. Of note, however, the pseudogenization approach to
621 dating groundwater colonization times reaches its limits when the gene fails to be amplified,
622 presumably due to a too long period of time a species spent underground (Lefébure et al., 2017;
623 Langille et al., 2022).

624

625 **4.3 | Comparative phylogenetic analyses of body size and sexual body size dimorphism**

626

627 Our phylogenetic comparative study of BS and SBSD between surface- and groundwater
628 species illustrates the usefulness of WAD for documenting evolutionary changes during
629 transition to novel habitats. We found that BS in asellids was constrained by the size of habitat
630 in both sexes. This corroborates Pipan and Culver (2017)'s hypothesis that BS within clades

631 containing subterranean species is in part controlled by habitat volume because pore size
632 between rocks can set an upper limit to maximum BS.

633

634 We provide the first, phylogenetically-controlled evidence that the difference in BS between
635 surface- and groundwater species is sex-dependent. Body size of males was significantly larger
636 in surface- than in groundwater species. We propose that competition for synchronously
637 receptive females selects for large males in surface species, while competition for rare, highly
638 asynchronous, receptive females favors small males in groundwater species (Andersson, 1994;
639 Blanckenhorn, 2000; Kelly et al., 2008; Balázs et al., 2021). In precopulatory mate guarding
640 crustaceans, among which many surface asellid species are known (Jormalainen, 2007), large
641 males have a mating advantage because they can more easily displace small guarding males
642 from their guarded females (Ridley and Thompson, 1979). In groundwater asellids, males no
643 longer guard females prior to copulation (Henry, 1976) and selection probably favors small
644 males that are more agile and can attain receptive females more rapidly. In addition, small males
645 can use energy that they do not invest in growth for searching for mates.

646

647 In contrast to male BS, we found that female BS did not differ between surface water and
648 groundwater asellids. Whatever the habitat, strong fecundity selection probably favors large
649 female size because brood size increases with increasing BS (Ridley and Thompson, 1979;
650 Pincheira-Donoso and Hunt, 2017). However, groundwater females take longer to grow than
651 surface water females (Henry, 1976). Life history studies of asellids also showed that
652 groundwater species were long-lived ($\gg 2$ yr) and iteroparous, whereas surface water species
653 had short lifespan (ca. 1 year) and were semelparous (Steel, 1961; Henry, 1976).

654

655 We found that habitat specialization significantly influenced SBSD. A predominant pattern of
656 male-biased SBSD occurred in surface species, whereas groundwater species were in average
657 monomorphic in BS but exhibited much larger variation in SBSD (Figure 4). We provide two
658 non-mutually exclusive explanations for difference in SBSD between habitats, in addition to
659 selective factors influencing male and female BS discussed above. First, the degree of SBSD
660 decreases in groundwater species because females mate with multiple males and produce
661 multiple clutches of offspring during their life. Both aspects diminish the sex difference in the
662 opportunity for selection and hence the potential for SBSD (Shuster and Wade, 2003; Shuster
663 et al., 2013). Second, in the absence of precopulatory mate guarding, groundwater males may
664 still prefer larger females that produce more eggs, but they no longer have to be bigger than
665 females to carry them prior to copulation (Adams et al., 1985).

666

667 A recent morphological study by Balázs and coauthors (2021) investigated sexual dimorphism
668 in 17 morphological traits, including body size, using nine surface and six cave groundwater
669 populations of *Asellus aquaticus* showing various degrees of reduction of eyes and body
670 pigments (see also Biró et al., 2022). The authors showed that several morphological traits were
671 significantly less male-biased in cave than in surface populations (for example the shape of
672 pereopods I). However, contrary to the present study, they found no significant reduction in
673 male-biased dimorphism in body size upon cave groundwater colonization. A potential
674 explanation is that the intraspecific comparative study by Balázs and coauthors (2021) may
675 have been unable to detect a reduction in male-biased SBSD in cave populations of *A. aquaticus*
676 due to insufficient time for BS to evolve. Of note, males were reported to be smaller than
677 females in several depigmented and eyeless subterranean *Asellus* species including *A.*
678 *amamiensis*, *A. hyugaensis*, *A. primoryensis* and *A. tamaensis* (Matsumoto, 1960, 1961 1963;

679 [Henry and Magniez, 1993](#)). A potentially important proportion of the variance in SBS
680 exhibited by groundwater species might be due to differences in groundwater colonization time
681 among species. If so, using colonization time as a predictor instead of a qualitative present-day
682 biological status (i.e., eyeless and depigmented vs. ocellated and pigmented) would contribute
683 to a better understanding of trait changes associated with groundwater transitions. This may
684 become possible in a near future as sequences of genes accumulating loss-of-function mutations
685 during colonization (e.g. opsin gene, see [Lefebvre et al. 2017](#)) become available for a large
686 number of species.

687

688 Dimorphism also significantly depended on habitat size. Groundwater species exhibiting male-
689 biased dimorphism occurred in habitats of larger size than groundwater species exhibiting
690 female-biased dimorphism or monomorphism. A potential explanation is that the mating
691 selective pressure for more agile and hence smaller males is less in cave habitats than in
692 interstitial habitats. Another non-mutually exclusive hypothesis is that even with equivalent
693 mating selective pressures for BS in both habitats, only the smallest specimens of a surface
694 population can colonize interstitial habitats. Hence, even those populations that have recently
695 colonized interstitial habitats would exhibit a weak sexual dimorphism in body size. Yet,
696 populations that have recently colonized cave habitats would exhibit male-biased dimorphism
697 until sexual selection has had time to act.

698

699 Beyond BS, WAD provides many of the necessary resources for testing predictions on how
700 phenotypic traits linked to mating success, fecundity, and survival evolve in response to
701 changes in selective pressures and trade-offs during groundwater colonization. We provide
702 below three example predictions. First, if searching for rather than fighting for mates is key to

703 determining mating success of male groundwater species, then, selection is likely to target
704 sensory organs that improve the ability of males to find females. More specifically, the hundreds
705 of specimens referenced in WAD can be used to test whether males of groundwater species
706 have longer antennae, relative to BS, than surface males and groundwater females, because long
707 antennae are advantageous for detecting receptive females (Bertin and Cézilly, 2003; Balázs et
708 al., 2021). Second, life history theory predicts that relative to their BS, groundwater, iteroparous
709 species should produce fewer but larger eggs per reproductive event than surface, semelparous
710 species (Fišer, 2019a; Venarsky et al., 2023). WAD keeps full record of the number of
711 ovigerous females contained in hundreds of specimen lots for testing this hypothesis. Third,
712 WAD resources can be used to test for the occurrence of a trade-off between transient fecundity
713 (i.e. the number of offspring produced per brood per single reproductive event) and adult
714 survival in long-lived, iteroparous groundwater species. Fecundity selection favors increase in
715 BS, whereas selection for survival may favor narrow and elongated body shapes that allow
716 individuals to withdraw into tiny hiding places to escape predators (Miller, 1933; Fišer et al.,
717 2013; Fišer Ž. et al., 2019). A trade-off may arise because an elongated brood pouch prevents
718 good ventilation of eggs beyond a certain BS. If such a trade-off exists, we predict variation in
719 BS to be more evolutionarily constrained in groundwater females than in surface females. This
720 prediction can be tested by comparing best-fit evolution models of BS and shape between
721 habitats and sexes.

722

723 5 | CONCLUSION

724 The asellids fulfill many of the desirable attributes of a model animal system for studying
725 evolution during colonization of a new environment, in particular here groundwater. Recently,
726 Protas and coauthors (2023) synthesized the ecological and molecular resources available for

727 studying microevolutionary dynamics of groundwater colonization from multiple cave and
728 surface populations of *Asellus aquaticus*. Here, we make available to the scientific community
729 a comprehensive set of taxonomic, distributional and molecular resources and biological
730 material that have been acquired for studying macroevolutionary dynamics of groundwater
731 colonization from multiple-species data. Looking at trait variation among multiple independent
732 colonization events across a wide range of times since colonization can provide better
733 understanding into the temporal dynamics of phenotypic evolution.

734

735 **ACKNOWLEDGEMENTS**

736 We thank all collectors, including the many speleologists and naturalists, who kindly provided
737 specimens of asellids: their donation together with their names are gratefully acknowledged in
738 the database. The World Asellidae database GOTIT is hosted by the CNRS/IN2P3 Computing
739 Center (Villeurbanne, France). Sampling, taxonomic identification, DNA sequencing, database
740 management, body size measurements and data analysis were supported financially by: French
741 National Research Agency projects CONVERGENOMICS (ANR-15-CE32-0005), EUR
742 H20'Lyon project (ANR-17-EURE-0018), and Biodiversa+ Project DarCo (F.M., N.S., C.J.D.,
743 L.K.D., T.L., C.F., C.I.); the Slovenian Research Agency through the Research Core
744 Programme Funding P1-0184 (T.D., B.S. and M.Z.); Spanish project PID2019-110243GB-100
745 of MICINN/FEDER (A.I.C.); Romania Ministry of Research, Innovation and Digitization
746 grant, CNCS/CCCDI – UEFISCDI, project 2/2019 (DARKFOOD), within PNCDI III
747 (O.T.M.); the Polish Ministry of Science and Higher Education through projects no. N N303
748 579439 and 5818/ B/P01/2010/39 (M.G.); the VILLUM FONDEN (research grant 15471) and
749 Portuguese National Funds through "Fundação para a Ciência e a Tecnologia" (FCT) within the

750 cE3c Unit funding UIDB/00329/2020 (A.S.P.S.R.); Belgium exceptional grant ARES-CCD
751 and French SCAC project (R.P.T.K.).

752

753 **CONFLICT OF INTEREST**

754 The authors declare that they have no competing interests.

755

756 **AUTHOR CONTRIBUTIONS**

757 All authors revised the manuscript and approved the publication. Manuscript conception,
758 writing, editing, illustrations: N.S., F.M., C.J.D, D.E. Database structure conception: leaders:
759 F.M., P.G.; contributors: L.K.D., T.L., C.J.D., C.M., C.C., D.E. Writing of code for the database
760 and web application: P.G., L.D. Database management: F.M., L.K.D. Molecular data
761 acquisition and management: L.K.D., C.J.D., T.L., C.F. Phylogenetic and comparative
762 analyses: N.S., C.J.D., F.M. Acquisition of body size data: C.I., F.M., N.S. The authors
763 hereinafter largely contributed to sampling and identification of biological material for the
764 following groups and/or regions. North American asellids / North America: J.J.L.; multiple sites
765 in the European Union: N.S., F.M., C.J.D., C.F., D.E., T.L.; coxalis group (*Proasellus*): F.S.;
766 The Balkan Peninsula: B.S., S.G., T.D., M.Z., M.G.; Germany / Luxemburg: D.W.; Portugal:
767 A.S.P.S.R.; Russia: D.P.; Crete: K.P.; United Kingdom: L.R.F.D.K.; Belgium: G.M.; Nouvelle
768 Aquitaine (France): F.L.; Iran: M.J.M.H.; Spain: B.G.D.B., A.I.C.; Cameroon / Stenasellidae:
769 R.P.T.K; Algeria: A.T., N.B.; Italy: D.M.P.G.; Romania: O.T.M.

770

771 **DATA AVAILABILITY AND BENEFIT SHARING**

772 The sequence data generated as part of this study have been deposited in NCBI. Sequence
773 metadata, sequence alignments, the World Asellidae phylogeny, and data used in comparative

774 analyses are included as Supporting Information at the publisher's website or archived in
775 Zenodo. Access to metadata and data stored in WAD is at <https://gotit.univ-lyon1.fr> upon
776 request from the corresponding author. Temporary logins in the read mode for the reviewers
777 are as follows: User name: REVIEWER; Password: MOLECOLRES123456. The web
778 application GOTIT and structure of the World Asellidae Database are distributed with full
779 documentation at <https://github.com/gotit-dev/gotit> under the terms of GNU General Public
780 License. A demo version of GOTIT application is available at <https://gotit.cnrs.fr>.
781 The work presented herein is from a collaborative group of researchers who is committed to
782 international scientific partnerships, as well as institutional capacity building. Scientific
783 collaborators are included as co-authors and the results of research are shared with the sample
784 provider communities and the broader scientific community via the World Asellidae Database.

785

786 **ORCID**

787 *Nathanaelle Saclier* <https://orcid.org/0000-0003-1522-9644>

788 *Lara Konecny-Dupré* <https://orcid.org/0009-0004-4604-687X>

789 Philippe Grison

790 *Louis Duchemin* <https://orcid.org/0000-0002-7984-9554>

791 *David Eme* <https://orcid.org/0000-0001-8790-0412>

792 Chloé Martin

793 *Cécile Callou* <https://orcid.org/0000-0002-8540-8114>

794 *Tristan Lefébure* <https://orcid.org/0000-0003-3923-8166>

795 *Clémentine François* <https://orcid.org/0000-0001-7781-8781>

796 Colin Issartel

797 Julian J. Lewis

- 798 *Fabio Stoch* <https://orcid.org/0000-0003-4535-3769>
- 799 *Boris Sket* <https://orcid.org/0000-0002-7398-5483>
- 800 *Sanja Gottstein* <https://orcid.org/0000-0002-1424-2911>
- 801 *Teo Delić* <https://orcid.org/0000-0003-4378-5269>
- 802 *Maja Zagmajster* <https://orcid.org/0000-0003-1323-9937>
- 803 *Michal Grabowski* <https://orcid.org/0000-0002-4551-3454>
- 804 *Dieter Weber* <https://orcid.org/0000-0001-7813-842X>
- 805 *Ana Sofia P.S. Reboleira* <https://orcid.org/0000-0002-4756-7034>
- 806 *Dmitry Palatov* <https://orcid.org/0000-0002-8826-9316>
- 807 *Kaloust Paragamian* <https://orcid.org/0000-0001-7372-733X>
- 808 Lee R.F.D. Knight
- 809 Georges Michel
- 810 Francois Lefebvre
- 811 *Mohammad-Javad Malek Hosseini* <https://orcid.org/0000-0001-7411-2150>
- 812 *Ana I. Camacho* <https://orcid.org/0000-0003-0596-7678>
- 813 *Gartzia De Bikuña Begoña* <https://orcid.org/0000-0001-6089-9458>
- 814 *Amina Taleb* <https://orcid.org/0000-0003-4045-590X>
- 815 *Nouria Belaidi* <https://orcid.org/0009-0004-2780-8031>
- 816 *Raoul P. Tuekam Kayo* <https://orcid.org/0000-0002-2578-4898>
- 817 *Diana Maria Paola Galassi* <https://orcid.org/0000-0002-6448-2710>
- 818 *Oana Teodora Moldovan* <https://orcid.org/0000-0002-1262-0675>
- 819 *Christophe J. Douady* <https://orcid.org/0000-0002-4503-8040>
- 820 *Florian Malard* <https://orcid.org/0000-0001-8037-4464>
- 821

822 SUPPORTING INFORMATION

823 The following additional supporting information may be found in the online version of the
824 article at the publisher's website or at <https://doi.org/10.5281/zenodo.6474972> (i.e. the World
825 Asellidae phylogeny with credibility Intervals for the age of the nodes).
826 SI Table 1. Metadata for the 2093 COI sequences used in the study.
827 SI Table 2. List of primers.
828 SI Table 3. Selection of most often used primers.
829 SI Table 4. Alignment of the 2093 COI sequences used for the delimitation of MOTUs.
830 SI Table 5. Alignment of the 424 COI sequences used for the four-gene dated phylogeny.
831 SI Table 6. Alignment of the 424 16S sequences used for the four-gene dated phylogeny.
832 SI Table 7. Alignment of the 424 FASTKD4 sequences used for the four-gene dated phylogeny.
833 SI Table 8. Alignment of the 424 28S sequences used for the four-gene dated phylogeny.
834 SI Table 9: Metadata for the DNA sequences used for the 4-gene dated phylogeny.
835 SI Table 10. Paleobiogeographic events used to constrain species divergence times.
836 SI Table 11. Data on body size, sexual body size dimorphism, habitat specialization and habitat
837 size used in comparative analyses.
838 SI Table 12. Metadata for the DNA sequences deposited in NCBI as part of this study.
839 SI Table 13. Results of likelihood ratio tests for testing the effects of habitat specialization and
840 habitat size on body size and sexual body size dimorphism.
841 SI Figure 1. Simplified logical model of the World Asellidae Database.
842 SI Figure 2. Full logical model of the World Asellidae Database.
843 SI Figure 3. Mapping of occurrence data, specimen lots, COI sequences and MOTUs within
844 morphospecies with GOTIT.
845 SI Figure 4. Timetree of Asellidae (Isopoda, Pancrustacea).
846 SI Figure 5. Lineage through time plots of Aselloidea.
847 SI Figure 6. Phylogeny and comparative data set for the analysis of body size and sexual body
848 size dimorphism.
849

850 REFERENCES

851 Adams, J., Greenwood, P., Pollitt, R., & Yonow, T. (1985). Loading constraints and sexual
852 size dimorphism in *Asellus aquaticus*. *Behaviour*, 92(3/4), 277–287.
853 <http://www.jstor.org/stable/4534415>
854 Andersson, M. (1994). *Sexual selection*. Princeton University Press.
855 Balart-García, P., Aristide, L., Bradford, T., Beasley-Hall, P., Polak, S., Ribera, I., Cooper, S.,
856 & Fernandez, R. (2023). Genomic exaptation and convergent evolution paved the way to
857 independent subterranean colonization across beetle lineages. Research Square Preprint
858 under Review at Nature Portfolio. <https://doi.org/10.21203/rs.3.rs-2254102/v1>
859 Balázs, G., Biró, A., Fišer, Ž., Fišer, C., & Herczeg, G. (2021). Parallel morphological evolution
860 and habitat-dependent sexual dimorphism in cave- vs. surface populations of the *Asellus*
861 *aquaticus* (Crustacea: isopoda: Asellidae) species complex. *Ecology and Evolution*, 11(21),
862 15389e15403. <https://doi.org/10.1002/ece3.8233>
863 Bertin, A., & Cezilly, F. (2003). Sexual selection, antennae length and the mating advantage of
864 large males in *Asellus aquaticus*. *Journal of Evolutionary Biology*, 16(4), 698–707.
865 <https://doi.org/10.1046/j.1420-9101.2003.00565.x>

- 866 Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K.,
867 & Das, L. (2007). Cryptic species as a window on diversity and conservation. *Trends in*
868 *Ecology & Evolution*, 22 (3), 148e155. <https://doi.org/10.1016/j.tree.2006.11.004>
- 869 Biró, A., Balázs, G., Fišer, Ž., & Herczeg, G. (2022). Gender inequality in the dark: are
870 adaptations to the cave environment sex-specific? *Karstologia Memoires*, 21(1), 229–232.
- 871 Blanckenhorn, W. U. (2000). The evolution of body size: what keeps organisms small? *The*
872 *Quarterly Review of Biology*, 75(4), 385–407. <http://www.jstor.org/stable/2664968>
- 873 Calvignac, S., Konecny, L., Malard, F., & Douady, C. J. (2011). Preventing the pollution of
874 mitochondrial data sets with nuclear mitochondrial paralogs (numts). *Mitochondrion*, 11,
875 246–254. <https://doi.org/10.1016/j.mito.2010.10.004>
- 876 Casquet, J., Thebaud, C., & Gillespie, R. G. (2012). Chelex without boiling, a rapid and easy
877 technique to obtain stable amplifiable DNA from small amounts of ethanol stored
878 spiders. *Molecular Ecology Resources*, 12, 136–141. <https://doi.org/10.1111/j.1755-0998.2011.03073.x>
- 879
880 Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in
881 phylogenetic analysis. *Molecular Biology and Evolution*, 17, 540–552.
882 <https://doi.org/10.1093/oxfordjournals.molbev.a026334>
- 883 Chevenet, F., Dolédec, C., & Chessel, D. (1994). A fuzzy coding approach for the analysis of
884 long-term ecological data. *Freshwater Biology*, 31, 295–309. <https://doi.org/10.1111/j.1365-2427.1994.tb01742.x>
- 885
886 Christiansen, K. A. (1961). Convergence and parallelism in cave Entomobryinae. *Evolution*,
887 15, 288e301. <https://doi.org/10.2307/2406229>
- 888 Copilas-Ciocianu, D., Fišer, C., Borza, P., & Petrusek, A. (2018). Is subterranean lifestyle
889 reversible? Independent and recent large-scale dispersal into surface waters by two species
890 of the groundwater amphipod genus *Niphargus*. *Molecular Phylogenetics and Evolution*,
891 119, 37e49. <https://doi.org/10.1016/j.ympev.2017.10.023>
- 892 Culver, D. C., & Pipan, T. (2019). *The Biology of Caves and Other Subterranean Habitats*.
893 Oxford University Press.
- 894 Degen, R., Aune, M., Bluhm, B. A., Cassidy, C., Kędra, M., Kraan, C., Vandepitte, L.,
895 Włodarska-Kowalczyk, M., Zhulay, I., Albano, P. G., Bremner, J., Grebmeier, J. M., Link,
896 H., Morata, N., Nordström, M. C., Shojaei, M. G., Sutton, L., & Zuschin, M. (2018). Trait-
897 based approaches in rapidly changing ecosystems: A roadmap to the future polar oceans.
898 *Ecological Indicators*, 91, 722–736. <https://doi.org/10.1016/j.ecolind.2018.04.050>
- 899 Eme, D., Malard, F., Konecny-Dupré, L., Lefébure, T., & Douady, C. J. (2013). Bayesian
900 phylogeographic inferences reveal contrasting colonization dynamics among European
901 groundwater isopods. *Molecular Ecology*, 22, 5685e5699.
902 <https://doi.org/10.1111/mec.12520>
- 903 Eme, D., Zgamaister, M., Delić, T., Fišer, C., Flot, J.-F., Konecny-Dupré, L., Pálsson, S., Stoch,
904 F., Zakšek, V., Douady, C. J., & Malard, F. (2018). Do cryptic species matter in
905 macroecology? Sequencing European groundwater crustaceans yields smaller ranges but
906 does not challenge biodiversity determinants. *Ecography*, 41, 424–436.
907 <https://doi.org/10.1111/ecog.02683>
- 908 Faille, A. (2019). Beetles. In W. B. White, D. C. Culver, & T. Pipan (Eds.), *Encyclopedia of*
909 *Caves* (pp. 102–108). Academic Press. <https://doi.org/10.1016/B978-0-12-814124-3.00014-5>
- 910
911 Fairbairn, D. J. (2007). Introduction: the enigma of sexual size dimorphism. In D. J. Fairbairn,
912 W. U. Blanckenhorn, & T. Székely (Eds.), *Sex, Size, and Gender Roles: Evolutionary Studies*

913 of *Sexual Size Dimorphism* (pp. 1–10). Oxford University Press.
 914 <https://doi.org/10.1093/acprof:oso/9780199208784.003.0001>

915 Felsenstein, J. (1988). Phylogenies and Quantitative Characters. *Annual Review of Ecology and*
 916 *Systematics*, 19, 445–471. <http://www.jstor.org/stable/2097162>

917 Fišer, C., Zagamajster, M., & Zakšek, V. (2013). Coevolution of life history traits and
 918 morphology in female subterranean amphipods. *Oikos*, 122, 770–778.
 919 <http://www.jstor.org/stable/41937725>

920 Fišer, C., Robinson, C. T., & Malard, F. (2018). Cryptic species as a window into the paradigm
 921 shift of the species concept. *Molecular Ecology*, 27, 613–635.
 922 <https://doi.org/10.1111/mec.14486>

923 Fišer, C., Brancelj, A., Yoshizawa, M., Mammola, S., & Fišer, Ž. (2023). Dissolving
 924 morphological and behavioral traits of groundwater animals into a functional phenotype. In
 925 F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater Ecology and Evolution* (pp. 415–
 926 438). Academic Press. <https://doi.org/10.1016/B978-0-12-819119-4.00012-3>

927 Fišer, Ž., Prevorčnik, S., Lozej, N., & Trontelj, P. (2019). No need to hide in caves: shelter-
 928 seeking behavior of surface and cave ecomorphs of *Asellus aquaticus* (Isopoda: Crustacea).
 929 *Zoology*, 134, 58e65. <https://doi.org/10.1016/j.zool.2019.03.001>

930 Fišer, C. (2019a). Adaptation: Morphological. In W. B. White, D. C. Culver, & T. Pipan (Eds.),
 931 *Encyclopedia of Caves* (pp. 33–39). Academic Press. <https://doi.org/10.1016/B978-0-12-814124-3.00005-4>

932

933 Fišer, C. (2019b). Collaborative databasing should be encouraged. *Trends in Ecology and*
 934 *Evolution*, 34, 184–185. <https://doi.org/10.1016/j.tree.2018.12.001>

935 Fong, D. W., & Carlini, D. B. (2023). Ecological and evolutionary perspectives on groundwater
 936 colonization by the amphipod crustacean *Gammarus minus*. In F. Malard, C. Griebler, & S.
 937 Rétaux (Eds.), *Groundwater Ecology and Evolution* (pp. 373–392). Academic Press.
 938 <https://doi.org/10.1016/B978-0-12-819119-4.15007-3>

939 Fontaneto, D., Flot, J.-F., & Tang, C. Q. (2015). Guidelines for DNA taxonomy, with a focus
 940 on the meiofauna. *Marine Biodiversity*, 45, 433–451. <https://doi.org/10.1007/s12526-015-0319-7>

941

942 Gouy, M., Guindon, S., & Gascuel, O. (2010). Seaview version 4: a multiplatform graphical
 943 user interface for sequence alignment and phylogenetic tree building. *Molecular biology and*
 944 *evolution*, 27, 221–224. <https://doi.org/10.1093/molbev/msp259>

945 Gross, J. B., Boggs, T. E., Rétaux, S., & Torres-Paz, J. (2023). Developmental and genetic basis
 946 of troglomorphic traits in the teleost fish *Astyanax mexicanus*. In F. Malard, C. Griebler, &
 947 S. Rétaux (Eds.), *Groundwater Ecology and Evolution* (pp. 351–371). Academic Press.
 948 <https://doi.org/10.1016/B978-0-12-819119-4.00004-4>

949 Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010).
 950 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
 951 performance of phylml 3.0. *Systematic Biology*, 59, 307–321.
 952 <https://doi.org/10.1093/sysbio/syq010>

953 Henry, J.-P., & Magniez, G. (1995). Nouvelles données sur les Asellidae épigés d Extrême-
 954 Orient (Crustacea, Isopoda, Asellota). *Contributions to Zoology*, 65(2), 101–122.
 955 <https://doi.org/10.1163/26660644-06502003>

956 Henry, J.-P., & Magniez, G. (1993). Présence d'Asellides stygobies (Crustacea, Isopoda,
 957 Aselloidea) dans la région du Primorje, Sibérie sud-orientale. *Contributions to Zoology*,
 958 62(3), 179–191. <https://doi.org/10.1163/26660644-06203003>

959 Henry, J.-P. (1976). *Recherches sur les Asellidae hypogés de la lignee cavaticus* (Crustacea,
 960 *Isopoda, Asellota*). PhD thesis, University of Dijon, Dijon, France.

- 961 Hidding, B., Michel, E., Natyaganova, A. V., & Sherbakov, D. Y. (2003). Molecular evidence
962 reveals a polyphyletic origin and chromosomal speciation of Lake Baikal's endemic asellid
963 isopods. *Molecular Ecology*, 12(6), 1509–1514. [https://doi.org/10.1046/j.1365-](https://doi.org/10.1046/j.1365-294X.2003.01821.x)
964 [294X.2003.01821.x](https://doi.org/10.1046/j.1365-294X.2003.01821.x)
- 965 Hobern, D., Apostolico, A., Arnaud, E., Bello, J. C., Canhos, D., Dubois, G., ... Willoughby,
966 S. (2012). *Global biodiversity informatics outlook: delivering biodiversity knowledge in the*
967 *information age*. Global Biodiversity Information Facility. [https://doi.org/10.15468/6jxa-](https://doi.org/10.15468/6jxa-yb44)
968 [yb44](https://doi.org/10.15468/6jxa-yb44)
- 969 Hose, G., Chariton, A., Daam, M., Di Lorenzo, T., Galassi, D., Halse, S., Reboleira, A.S.P.S.,
970 Robertson, A., Schmidt, S., & Korbel, K. (2022). Invertebrate traits, diversity and the
971 vulnerability of groundwater ecosystems. *Functional Ecology*, 36(9), 2200–2214.
972 <https://doi.org/10.1111/1365-2435.14125>
- 973 Humphreys, W.F. (2000). Relict faunas and their derivation. In H. Wilkens, D. C. Culver, &
974 W. F. Humphreys (Eds.), *Subterranean Ecosystems* (pp. 417–432). Elsevier.
- 975 Jormalainen, V., 2007. Mating strategies in isopods: from mate monopolization to conflicts. In
976 J. E. Duffy, & M. Thiel (Eds.), *Evolutionary Ecology of Social and Sexual Systems:*
977 *Crustaceans as Model Organisms* (pp. 167–190). Oxford University Press.
978 <https://doi.org/10.1093/acprof:oso/9780195179927.003.0008>
- 979 Katoh, K., & Standley, D. M. (2013). MAFFT Multiple sequence alignment software version
980 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4),
981 772–780. <https://doi.org/10.1093/molbev/mst010>
- 982 Kelly, C. D., Bussière, L. F., & Gwynne, D. T. (2008). Sexual selection for male mobility in a
983 giant insect with female-biased size dimorphism. *The American Naturalist*, 172(3), 417–
984 423. <https://www.journals.uchicago.edu/doi/10.1086/589894>
- 985 Konec, M., Prevorčnik, S., Sarbu, S. M., Verovnik, R., & Trontelj, P. (2015). Parallels between
986 two geographically and ecologically disparate cave invasions by the same species, *Asellus*
987 *aquaticus* (Isopoda, Crustacea). *Journal of Evolutionary Biology*, 28(4), 864e875.
988 <https://doi.org/10.1111/jeb.12610>
- 989 Kostanjsek, R., Zaksek, V., Bizjak-Mali, L., & Trontelj, P. (2023). The olm (*Proteus anguinus*),
990 a flagship groundwater species. . In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater*
991 *Ecology and Evolution* (pp. 305–327). Academic Press. [https://doi.org/10.1016/B978-0-12-](https://doi.org/10.1016/B978-0-12-819119-4.15006-1)
992 [819119-4.15006-1](https://doi.org/10.1016/B978-0-12-819119-4.15006-1)
- 993 Kowalko, J., Dickinson, M. H., Voshall, L. B., & Dow, J. A. T. (2020). Utilizing the blind
994 cavefish *Astyanax mexicanus* to understand the genetic basis of behavioral evolution.
995 *Journal of Experimental Biology*, 223 (Suppl. 1), jeb208835.
996 <https://doi.org/10.1242/jeb.208835>
- 997 Langille, B. L., Tierney, S. M., Bertozzi, T., Beasley-Hall, P. G., Bradford, T. M., Fagan-
998 Jeffries, E. P., Hyde, J., Leijts, R., Richardson, M., Saint, K. M., Stringer, D. N., Villastrigo,
999 A., Humphreys, W. F., Austin, A. D., & Cooper, S. J. B. (2022). Parallel decay of vision
1000 genes in subterranean water beetles. *Molecular Phylogenetics and Evolution*, 173, 107522.
1001 <https://doi.org/10.1016/j.ympev.2022.107522>
- 1002 Lartillot, N., Lepage, T., & Blanquart, S. (2009). PhyloBayes 3: a bayesian software package
1003 for phylogenetic reconstruction and molecular dating. *Bioinformatics*, 25(17), 2286–2288.
1004 <https://doi.org/10.1093/bioinformatics/btp368>
- 1005 Ledford, J., Paquin, P., Cokendolpher, J., Campbell, J., & Griswold, C. (2011). Systematics of
1006 the spider genus *Neoleptoneta* Brignoli, 1972 (Araneae : Leptonetidae) with a discussion of
1007 the morphology and relationships for the North American Leptonetidae. *Invertebrate*
1008 *Systematics*, 25, 334–388. <https://www.publish.csiro.au/is/IS11014>

- 1009 Lefebure, T., Douady, C. J., Gouy, M., & Gibert, J. (2006). Relationship between
1010 morphological taxonomy and molecular divergence within Crustacea: proposal of a
1011 molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution*,
1012 40, 435–447. <https://doi.org/10.1016/j.ympev.2006.03.014>
- 1013 Lefébure, T., Morvan, C., Malard, F., François, C., Konecny-Dupré, L., Guéguen, L., Weiss-
1014 Gayet, M., Seguin-Orlando, A., Ermini, L., Der Sarkissian, C., Charrier, N. P., Eme, D.
1015 Mermillod-Blondin, F., Duret, L., Vieira, C., Orlando, L., & Douady, C. J., (2017). Less
1016 effective selection leads to larger genomes. *Genome Research*, 27, 1016–1028.
1017 <https://doi.org/10.1101/gr.212589.116>
- 1018 Lepage, T., Bryant, D., Philippe, H., & Lartillot, N. (2007). A general comparison of relaxed
1019 molecular clock models. *Molecular Biology and Evolution*, 24(12), 2669–2680.
1020 <https://doi.org/10.1093/molbev/msm193>
- 1021 Lewis, J. J., Lewis, S. L., Orndorff, W., Orndorff, Z., Malard, F., Konecny-Dupré, L., Saclier,
1022 N., & Douady, C. J. (2023). *The groundwater isopods of Virginia (Isopoda : Asellidae and*
1023 *Cirolanidae)*. Virginia Museum of Natural History, Martinsville, VA, Special Publication
1024 19. In press.
- 1025 Lovich, J. E., & Gibbons, J. W. (1992). A review of techniques for quantifying sexual size
1026 dimorphism. *Growth Development and Aging*, 56, 269–281.
- 1027 Löytynoja, A., & Goldman, N. (2008). Phylogeny-aware gap placement prevents errors in
1028 sequence alignment and evolutionary analysis. *Science*, 320(5883), 1632–1635.
1029 <https://www.science.org/doi/10.1126/science.1158395>
- 1030 Lukić, M. (2019). Collembola. In W. B. White, D. C. Culver, & T. Pipan (Eds.), *Encyclopedia*
1031 *of Caves* (pp. 308–319). Academic Press. [https://doi.org/10.1016/B978-0-12-814124-](https://doi.org/10.1016/B978-0-12-814124-3.00034-0)
1032 [3.00034-0](https://doi.org/10.1016/B978-0-12-814124-3.00034-0)
- 1033 Maechler, M., Rousseeuw, P., Struyf, A., & Hubert, M. (2022). Package ‘cluster’: cluster
1034 analysis extended. R Foundation for Statistical Computing, Vienna, 2002. [https://svn.r-](https://svn.r-project.org/R/packages/trunk/cluster/)
1035 [project.org/R/packages/trunk/cluster/](https://svn.r-project.org/R/packages/trunk/cluster/)
- 1036 Malard, F., Grison, P., Duchemin, L., Konecny-Dupré, L., Lefébure, T., Saclier, N., Eme, D.,
1037 Martin, C., Callou, C., & Douady, C. J. (2020). GOTIT: A laboratory application software
1038 for optimizing multi-criteria species-based research. *Methods in Ecology and Evolution*, 11,
1039 159–167. <https://doi.org/10.1111/2041-210X.13307>
- 1040 Malard, F., Machado, E. G., Casane, D., Cooper, S., Fišer, C., & Eme, D. (2023). Dispersal and
1041 geographic range size in groundwater. In F. Malard, C. Griebler, & S. Rétaux (Eds.),
1042 *Groundwater Ecology and Evolution* (pp. 185–207). Academic Press.
1043 <https://doi.org/10.1016/B978-0-12-819119-4.15003-6>
- 1044 Malard, F. (2022). Groundwater Metazoans. In T. Mehner, & K. Tockner (Eds.), *Encyclopedia*
1045 *of Inland Waters* (pp. 474–487). Second Edition, Elsevier. [https://doi.org/10.1016/B978-0-](https://doi.org/10.1016/B978-0-12-819166-8.00176-6)
1046 [12-819166-8.00176-6](https://doi.org/10.1016/B978-0-12-819166-8.00176-6)
- 1047 Mammola, S., & Isaia, M. (2017). Spiders in caves. *Proceedings of the Royal Society B:*
1048 *Biological Sciences*, 284, 20170193. <https://doi.org/10.1098/rspb.2017.0193>
- 1049 Mammola, S., & Martínez, A. (2020). Let research on subterranean habitats resonate!
1050 *Subterranean Biology*, 36: 63–71. <https://doi.org/10.3897/subtbiol.36.59960>
- 1051 Mammola, S., Cardoso, P., Angyal, D., Balázs, G., Blick, T., Brustel, H., Carter, J., Ćurčić, S.,
1052 Danflous, S., Dányi, L., Déjean, S., Deltshv, C., Elverici, M., Fernández, J., Gasparo, F.,
1053 Komnenov, M., Komposch, C., Kováč, L., Kunt, K. B., Mock, A., Moldovan, O., Naumova,
1054 M., Pavlek, M., Prieto, C. E., Ribera, C., Rozwałka, R., Růžička, V., Vargovitsh, R. S.,
1055 Zaenker, S., & Isaia, M. (2019). Continental data on cave-dwelling spider communities

1056 across Europe (Arachnida: Araneae). *Biodiversity Data Journal*, 7, e38492.
1057 <https://doi.org/10.3897/BDJ.7.e38492>

1058 Mammola, S., Amorim, I. R., Bichuette, M. E., Borges, P. A. V., Cheeptham, N., Cooper, S. J.
1059 B., Culver, D. C., Deharveng, L., Eme, D., Ferreira, R. L., Fišer, C., Fišer, Ž., Fong, D. W.,
1060 Griebler, C., Jeffery, W. R., Jugovic, J., Kowalko, J. E., Lilley, T. M., Malard, F., Manenti,
1061 R., Martínez, A., Meierhofer, M. B., Niemiller, M. L., Northup, D. E., Pellegrini, T. G.,
1062 Pipan, T., Protas, M., Reboleira, A. S. P. S., Venarsky, M. P., Wynne, J. J., Zagamajster, M.,
1063 & Cardoso, P. (2020). Fundamental research questions in subterranean biology. *Biological*
1064 *Reviews*, 95, 1855–1872. <https://doi.org/10.1111/brv.12642>

1065 Mammola, S., Pavlek, M., Huber, B. A., Isaia, M., Ballarin, F., Tolve, M., Čupić, I., Hesselberg,
1066 T., Lunghi, E., Mouron, S., Graco-Roza, C., & Cardoso, P. (2022). A trait database and
1067 updated checklist for European subterranean spiders. *Scientific Data*, 9(1), 236.
1068 <https://doi.org/10.1038/s41597-022-01316-3>

1069 Martins, E. P., & Hansen, T. F. (1997). Phylogenies and the comparative method: A general
1070 approach to incorporating phylogenetic information into the analysis of interspecific data.
1071 *The American Naturalist*, 149(4), 646–667. <https://doi.org/10.1086/286013>

1072 Matsumoto, K. (1960). Subterranean isopods of the Kyushu District, with the descriptions of
1073 three new species. *Bulletin of the Biogeographical Society of Japan*, 22(3), 27–44.

1074 Matsumoto, K. (1961). Two subterranean isopods from the Amami Group (Ryukyu Islands),
1075 with a description of a new species. *Annotationes Zoologicae Japonenses*, 34(4), 208–215.

1076 Matsumoto, K. (1963). *Studies on the subterranean Isopoda of Japan with notes on the well-*
1077 *water fauna of Japan*. Part I: Studies on the subterranean Isopoda of Japan (No. 1).
1078 Supplement of the Annual Report, XIII Tokyo-to Laboratories for Medical Sciences, pp. 1–
1079 77.

1080 Miller, M. A. (1933). A new blind isopod, *Asellus californicus*, and a revision of the
1081 subterranean asellids. *University of California Publications in Zoology*, 39(4), 97–110.

1082 Morvan, C., Malard, F., Paradis, E., Lefébure, T., Konecny-Dupré, L., & Douady, C. J. (2013).
1083 Timetree of Aselloidea reveals species diversification dynamics in groundwater. *Systematic*
1084 *Biology*, 62(4), 512–522. <https://doi.org/10.1093/sysbio/syt015>

1085 Nelson, E. K., Piehler, B., Eckels, J., Rauch, A., Bellew, M., Hussey, P., Ramsay, S., Nathe,
1086 C., Lum, K., Krouse, K., Stearns, D., Connolly, B., Skillman, T., & Igra, M. (2011). LabKey
1087 Server: an open source platform for scientific data integration, analysis and collaboration.
1088 *BMC Bioinformatics*, 12, 71. <https://doi.org/10.1186/1471-2105-12-71>

1089 Niemiller, M. L., Fitzpatrick, B. M., Shah, P., Schmitz, L., & Near, T. J. (2013). Evidence for
1090 repeated loss of selective constraint in rhodopsin of amblyopsid cavefishes (Teleostei:
1091 Amblyopsidae). *Evolution*, 67(3), 732–748. <https://doi.org/10.1111/j.1558-5646.2012.01822.x>

1092

1093 Paradis, E., Claude, J., & Strimmer, K. (2004). APE: analyses of phylogenetics and evolution
1094 in R language. *Bioinformatics*, 20, 289–290. <https://doi.org/10.1093/bioinformatics/btg412>

1095 Pfenninger, M., & Schwenk, K. (2007). Cryptic animal species are homogeneously distributed
1096 among taxa and biogeographical regions. *BMC Evolutionary Biology*, 7, 121.
1097 <https://doi.org/10.1186/1471-2148-7-121>

1098 Pincheira-Donoso, D., & Hunt, J. (2017). Fecundity selection theory: concepts and evidence.
1099 *Biological Reviews*, 92(1), 341–356. <https://doi.org/10.1111/brv.12232>

1100 Pinheiro J, Bates D., & R Core Team (2022). nlme: Linear and Nonlinear Mixed Effects
1101 Models. R package version 3.1-160, <https://CRAN.R-project.org/package=nlme>.

- 1102 Pipan, T., & Culver, D. C. (2012). Convergence and divergence in the subterranean realm: a
 1103 reassessment. *Biological Journal of the Linnean Society*, 107, 1e14.
 1104 <https://doi.org/10.1111/j.1095-8312.2012.01964.x>
- 1105 Pipan, T., & Culver, D.C. (2017). The unity and diversity of the subterranean realm with respect
 1106 to invertebrate body size. *Journal of Cave and Karst Studies*, 79 (1), 1e9.
 1107 https://digitalcommons.usf.edu/kip_articles/5638
- 1108 Policarpo, M., Fumey, J., Lafargeas, P., Naquin, D., Thermes, C., Naville, M., Dechaud, C.,
 1109 Volff, J. N., Cabau, C., Klopp, C., Møller, P. R., Bernatchez, L., García-Machado, E.,
 1110 Rétaux, S., & Casane, D. (2021). Contrasting gene decay in subterranean vertebrates:
 1111 insights from cavefishes and fossorial mammals. *Molecular Biology and Evolution*, 38(2),
 1112 589e605. <https://doi.org/10.1093/molbev/msaa249>
- 1113 Poulson, T. L., & White, W. B. (1969). The cave environment. *Science*, 165, 971–981.
 1114 <https://www.science.org/doi.org/10.1126/science.165.3897.971>
- 1115 Prevorčnik, S., Blejec, A., & Sket, B. (2004). Racial differentiation in *Asellus aquaticus* (L.)
 1116 (Crustacea Isopoda Asellidae). *Archiv für Hydrobiologie*, 160(2), 193–214.
 1117 <https://doi.org/10.1127/0003-9136/2004/0160-0193>
- 1118 Protas, M., & Jeffery, W. R. (2012). Evolution and development in cave animals: from fish to
 1119 crustaceans. *Wiley Interdisciplinary Reviews: Developmental Biology*, 1, 823–845.
 1120 <https://doi.org/10.1002/wdev.61>
- 1121 Protas, M., Trontelj, P., Prevorčnik, S., & Fišer, Ž. (2023). The *Asellus aquaticus* species
 1122 complex: an invertebrate model in subterranean evolution. In F. Malard, C. Griebler, & S.
 1123 Rétaux (Eds.), *Groundwater Ecology and Evolution* (pp. 329–350). Academic Press.
 1124 <https://doi.org/10.1016/B978-0-12-819119-4.00016-0>
- 1125 Rétaux, S., & Jeffery, W. R. (2023). Voices from the underground: animal models for the study
 1126 of trait evolution during groundwater colonization and adaptation. In F. Malard, C. Griebler,
 1127 & S. Rétaux (Eds.), *Groundwater Ecology and Evolution* (pp. 285–304). Academic Press.
 1128 <https://doi.org/10.1016/B978-0-12-819119-4.00002-0>
- 1129 Ridley, M., & Thompson, D. J. (1979). Size and mating in *Asellus aquaticus* (Crustacea:
 1130 Isopoda). *Zeitschrift für Tierpsychologie*, 51, 380–397. <https://doi.org/10.1111/j.1439-0310.1979.tb00697.x>
- 1132 Saccò, M., Guzik, M. T., van der Heyde, M., Nevill, P., Cooper, S. J. B., Austin, A. D., Coates,
 1133 P. J., Allentoft, M. E., & White, N. E. (2022). eDNA in subterranean ecosystems:
 1134 applications, technical aspects, and future prospects. *Science of the Total Environment*, 820,
 1135 153223. <https://doi.org/10.1016/j.scitotenv.2022.153223>
- 1136 Saclier, N., François, C. M., Konecny-Dupré, L., Lartillot, N., Guéguen, L., Duret, L., Malard,
 1137 F., Douady, C. J., & Lefébure, T. (2018). Life history traits impact the nuclear rate of
 1138 substitution but not the mitochondrial rate in isopods. *Molecular Biology and Evolution*, 35,
 1139 2900–2912. <https://doi.org/10.1093/molbev/msy184>
- 1140 Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of
 1141 image analysis. *Nature Methods*, 9, 671–675. <https://doi.org/10.1038/nmeth.2089>
- 1142 Shuster, S. M., & Wade, M. J. (2003). *Mating systems and strategies*. Princeton University
 1143 Press.
- 1144 Shuster, S. M., Briggs, W. R., & Dennis, P. A. (2013). How multiple mating by females affects
 1145 sexual selection. *Philosophical Transactions of the Royal Society B*, 368, 20120046.
 1146 <https://doi.org/10.1098/rstb.2012.0046>
- 1147 Steel, E. A. (1961). Some observations on the life history of *Asellus aquaticus* (L.) and *Asellus*
 1148 *meridianus* Racovitza (Crustacea: Isopoda). *Proceedings of the Zoological Society of*
 1149 *London*, 137, 71–87. <https://doi.org/10.1111/j.1469-7998.1961.tb06162.x>

- 1150 Stern, D. B., & Crandall, K. A. (2018). Phototransduction gene expression and evolution in
1151 cave and surface crayfishes. *Integrative and Comparative Biology*, 58(3), 398–410.
1152 <https://doi.org/10.1093/icb/icy029>
- 1153 Stern, D. B., Breinholt, J., Pedraza-Lara, C., López-Mejía, M., Owen, C. L., Bracken-Grissom,
1154 H., Fetzner, J. W. Jr, & Crandall, K. A. (2017). Phylogenetic evidence from freshwater
1155 crayfishes that cave adaptation is not an evolutionary dead-end. *Evolution*, 71(10), 2522–
1156 2532. <https://doi.org/10.1111/evo.13326>
- 1157 Thioulouse, J., Dray, S., Dufour, A.-B., Siberchicot, A., Jombart, T., & Pavoine, S. (2018).
1158 *Multivariate Analysis of Ecological Data with ade4*. Springer.
1159 <https://link.springer.com/book/10.1007/978-1-4939-8850-1>
- 1160 Trontelj, P., Blejec, A., & Fišer, C. (2012). Ecomorphological convergence of cave
1161 communities. *Evolution*, 66(12), 3852–3865. [https://doi.org/10.1111/j.1558-
1162 5646.2012.01734.x](https://doi.org/10.1111/j.1558-5646.2012.01734.x)
- 1163 Venarsky, M., Niemiller, M. L., Fišer, C., Saclier, N., & Moldovan, O. T. (2023). Life histories
1164 in groundwater organisms. In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater
1165 Ecology and Evolution* (pp. 439–456). Academic Press. [https://doi.org/10.1016/B978-0-12-
1166 819119-4.00013-5](https://doi.org/10.1016/B978-0-12-819119-4.00013-5)
- 1167 Verdier, H., Konecny-Dupré, L., Marquette, C., Reveron, H., Tadier, S., Grémillard, L.,
1168 Barthès, A., Datry, T., Bouchez, A., & Lefébure, T. (2022). Passive sampling of
1169 environmental DNA in aquatic environments using 3D-printed hydroxyapatite samplers.
1170 *Molecular Ecology Resources*, 22(6), 2158–2170. <https://doi.org/10.1111/1755-0998.13604>
- 1171 Verovnik, R., Sket, B., & Trontelj, P. (2005). The colonization of Europe by the freshwater
1172 crustacean *Asellus aquaticus* (Crustacea: Isopoda) proceeded from ancient refugia and was
1173 directed by habitat connectivity. *Molecular Ecology*, 14(14), 4355–4369.
1174 <https://doi.org/10.1111/j.1365-294X.2005.02745.x>
- 1175 Wilkens, H., & Strecker, U. (2017). *Evolution in the Dark. Darwin's Loss Without Selection*.
1176 Springer. <https://doi.org/10.1007/978-3-662-54512-6>
- 1177 Zgmajster, M., Borko, Š., Delić, T., Douady, C. J., Eme, D., Malard, F., Trontelj, P., & Fišer,
1178 C. (2022). Availability of DNA barcodes in subterranean amphipods of Europe. *Karstologia
1179 Memoires*, 21(1), 361–364.
- 1180 Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation
1181 method with applications to phylogenetic placements. *Bioinformatics*, 29, 2869–2876.
1182 <https://doi.org/10.1093/bioinformatics/btt499>
1183
- 1184
- 1185

1186 **TABLES AND FIGURES (WITH CAPTIONS)**

1187 Table 1: Summary content of the World Asellidae Database (WAD). All items, except
 1188 references, are for Asellidae (Isopoda, Pancrustacea). Data extraction on 10 March 2023.

1189

Items	Number
Species Occurrence	
Species and subspecies	448
Record	9438
Country	55
Collection material ¹	
Specimen lot	1943
Specimen microscopic slide	1584
Specimen used for DNA extraction	4901
DNA extract	4362
DNA sequencing metadata	
Primer ²	138
PCR	22743
Chromatogram	12052
Sequence ³	
16S	3562
COI	2866
FASTKD4	922
28S	1202
Opsin	362
Literature reference	641

1190

1191 ¹ All specimens and DNA extracts are stored at -20°C

1192 ² See SI Tables 2 and 3

1193 ³ Validated sequences. Numbers differ from the number of sequences used in this study because
 1194 the database is regularly updated with new data.

1195

1196 Table 2: Numbers of surface water (Surf.) and groundwater (Grou.) described species contained
 1197 in the World Asellidae database (WAD) and numbers of morphospecies and MOTUs included
 1198 in the Asellidae timetree. Numbers in bold are totals.
 1199

Morphospecies groups /genera	WAD		Asellidae timetree				Species pairs ⁶
	Described species		Morphospecies ⁵		MOTUs		
	Surf.	Grou.	Surf.	Grou.	Surf.	Grou.	
1 - Asellus pattern	31	30	7	1	8	1	1
<i>Asellus</i>	28	10	5	1	6	1	1
<i>Calasellus</i>		2					NA
<i>Columbasellus</i>		1					NA
<i>Limnoasellus</i> ¹	1		1		1		0
<i>Mesoasellus</i>	1		1		1		0
<i>Nipponasellus</i>		5					NA
<i>Phreatoasellus</i>	1	9					NA
<i>Sibirasellus</i>		2					NA
<i>Uenasellus</i>		1					NA
2 - North American asellids ²	68	84	51	24	60	48	10
<i>Baicalasellus</i>	4		2		2		1
<i>Caecidotea</i>	9	39		9		28	1
<i>Conasellus</i>	21	24	13	9	17	10	6
<i>Gallasellus</i>		1		1		5	0
<i>Lirceolus</i>		6					NA
<i>Lirceus</i>	34	4	36	2	41	2	2
<i>Pseudobaicalasellus</i>		10		3		3	0
3 - Bragasellus & Synasellus	3	53	2	20	2	32	1
<i>Bragasellus</i>	3	18	2	8	2	19	1
<i>Synasellus</i>		35		12		13	0
4 - Proasellus ³	61	113	54	108	51	181	21
Others ⁴		5		1		1	1
<i>Bowmanasellus</i>		1					NA
<i>Oregonasellus</i>		1					NA
<i>Salmasellus</i>		2		1		1	1
<i>Stygasellus</i>		1					NA
Asellidae	163	285	114	154	121	263	30

1200

1201 ¹ *Nomen nudum* in [Hidding et al. \(2003\)](#)

1202 ² Genera according to recent revision by [Lewis et al. \(2023\)](#)

1203 ³ Including *Chthonasellus bodoni* Argano & Messana, 1991

1204 ⁴ Genera that cannot be assigned to any of the four species groups.

1205 ⁵ Numbers include undescribed morphospecies

1206 ⁶ Number of independent species pairs containing a surface water and a groundwater asellid
 1207 species.

1208 Table 3: Results of phylogenetic generalized least-squares regression models for testing the
 1209 effects of habitat size and specialization (i.e. surface vs groundwater habitats) on body size of
 1210 females and males and sexual dimorphism index. Significant P values are in bold.

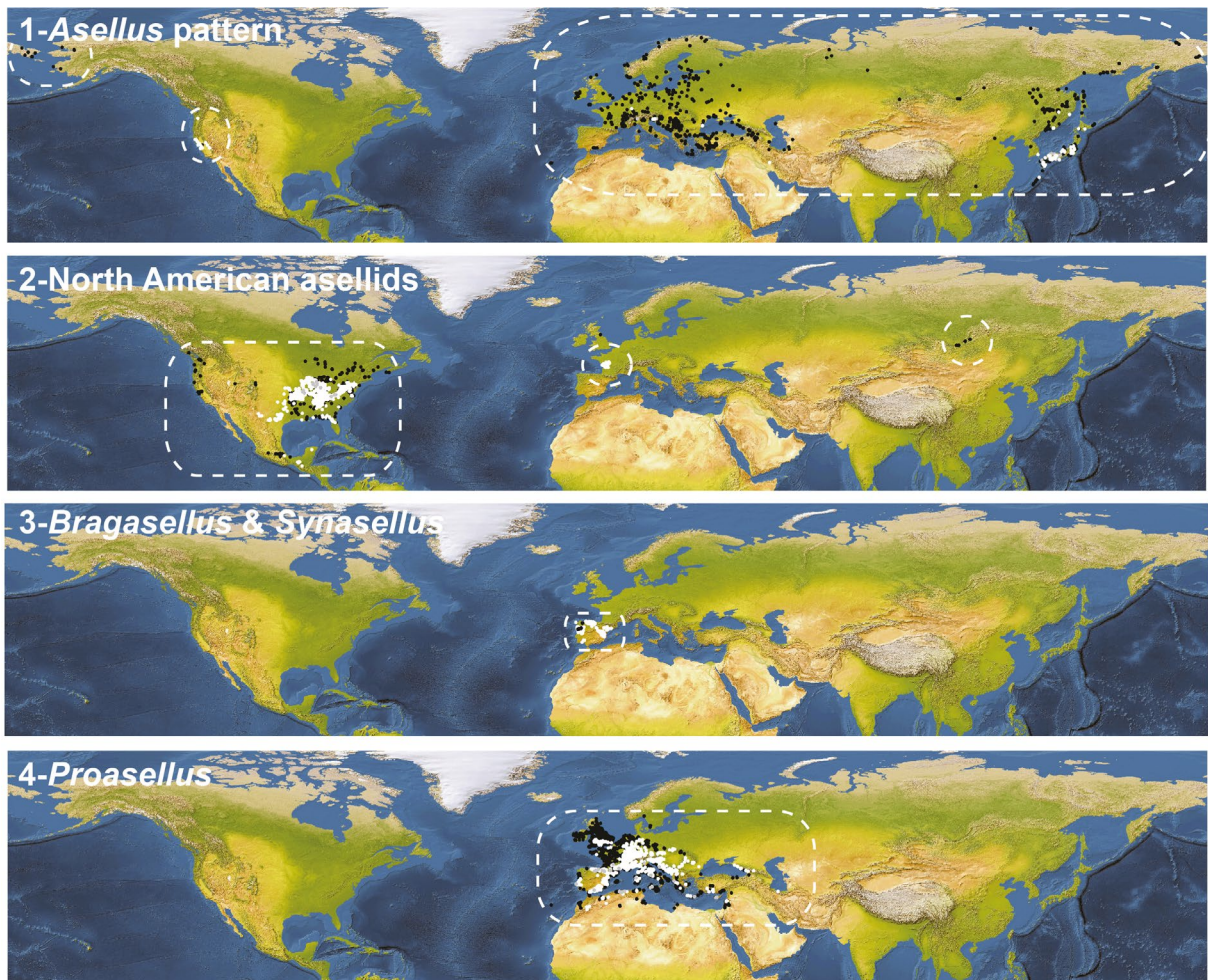
1211

Dependent variable	Explanatory variable	Parameter estimate	Standard error	<i>t</i>	P
Male body size	Intercept (groundwater)	1.871	0.401	4.662	
	Habitat size	0.186	0.044	4.242	<0.001
	Habitat specialization	0.166	0.080	2.084	0.0387
	Habitat Size × habitat specialization	-0.039	0.091	-0.427	0.670
Female body size	Intercept (groundwater)	1.768	0.354	4.994	
	Habitat size	0.129	0.039	3.350	0.001
	Habitat specialization	-0.008	0.070	-0.111	0.912
	Habitat Size × habitat specialization	-0.008	0.080	-0.098	0.922
Sexual dimorphism index	Intercept (groundwater)	-0.134	0.289	-0.472	
	Habitat size	-0.072	0.032	-2.268	0.025
	Habitat specialization	-0.203	0.057	-3.534	0.001
	Habitat Size × habitat specialization	0.044	0.065	0.676	0.500

1212

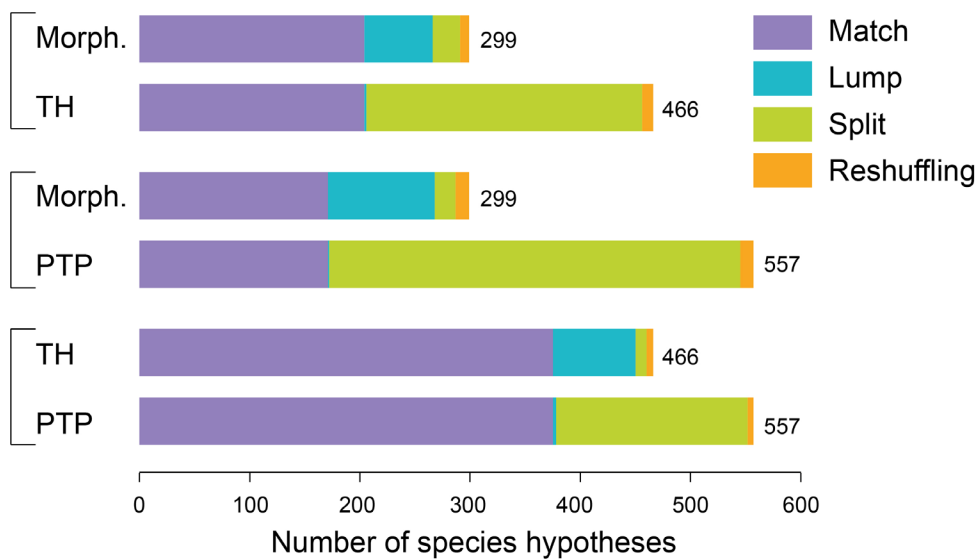
1213

1214 Figure 1: Distribution of four major species groups of Asellidae (Isopoda, Pancrustacea). Dots
1215 are species occurrence data contained in the World Asellidae Database (black dots: surface
1216 water species; white dots: groundwater species).
1217



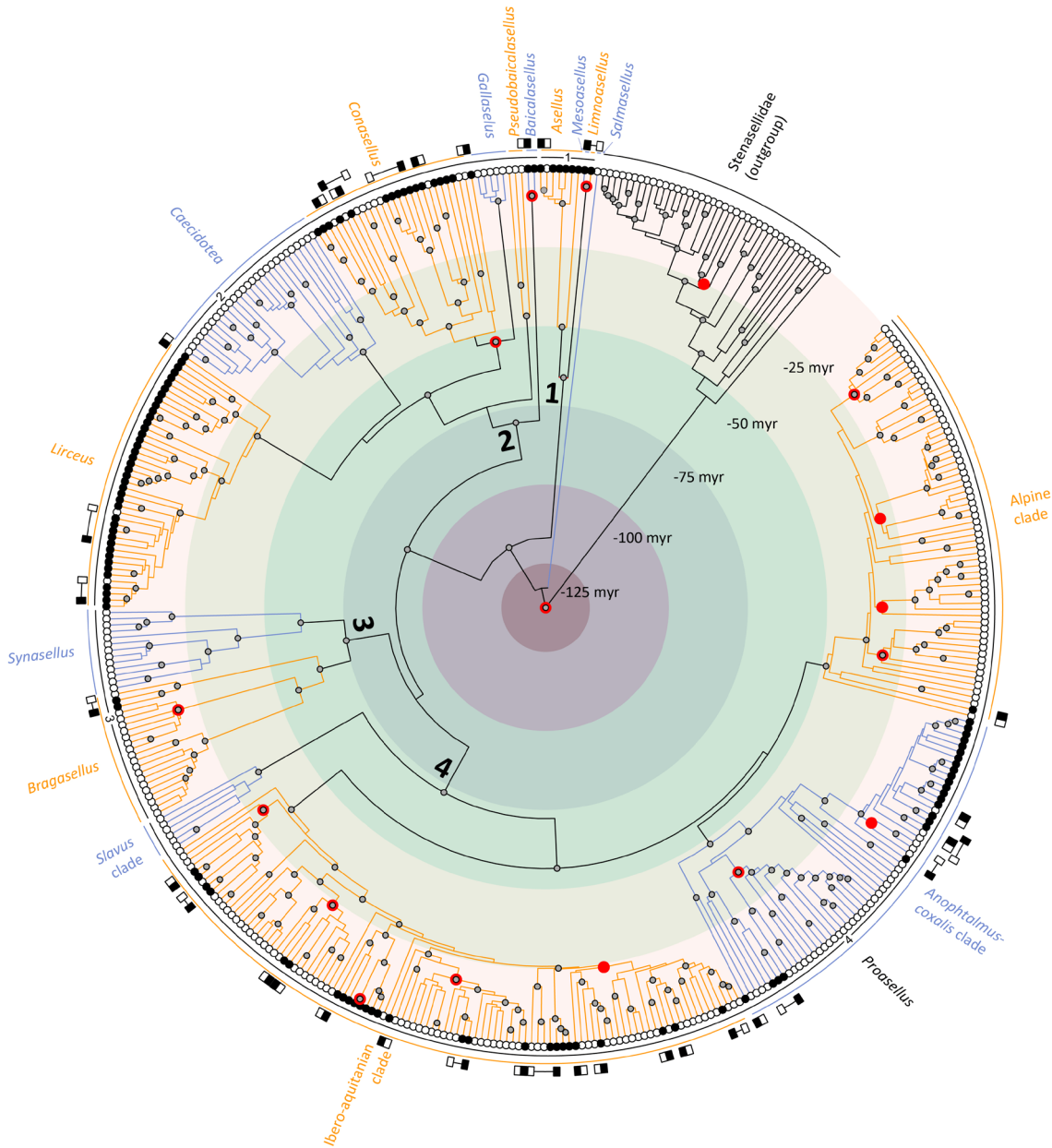
1218

1219 Figure 2: Pairwise taxonomic comparisons between the three different sets of aselloid species
 1220 hypotheses delimited using morphology (Morph.), a COI divergence threshold (TH), and the
 1221 Poisson tree processes model (PTP).
 1222



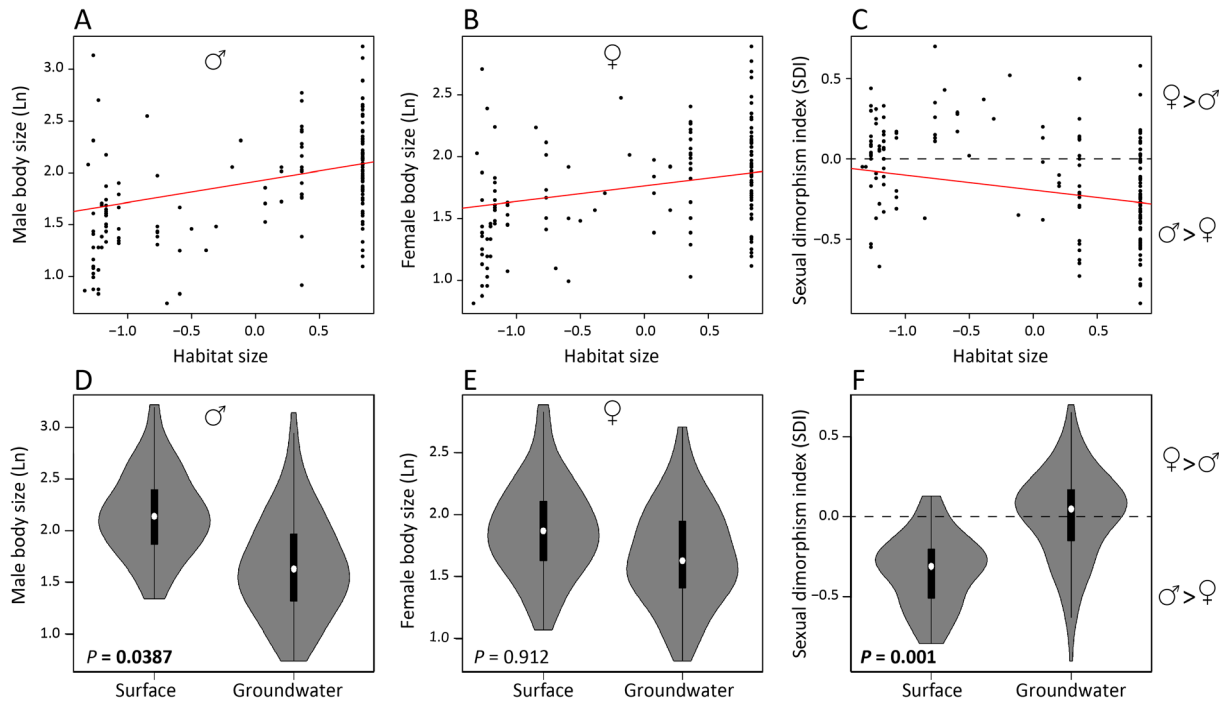
1223

1224 Figure 3: Timetree of Asellidae (Isopoda, Pancrustacea). The tree is rooted using Stenasellidae
 1225 as outgroup. Terminal nodes are molecular operational taxonomic units (MOTUs) as delimited
 1226 with the fixed COI threshold method (TH) implemented by **Lefébure and coauthors (2006)**.
 1227 White terminal nodes are eyeless and depigmented MOTUs; black terminal nodes are ocellated
 1228 and/or pigmented MOTUs. Color rings show time. Red and gray dots show paleobiogeographic
 1229 calibration points and node supports with posterior probabilities > 0.9, respectively. Black and
 1230 white squares on the outer ring show independent pairs of surface (black) and groundwater (white)
 1231 asellid species (see definition of species pairs in materials and methods). Legends show
 1232 genera and main species groups within the Asellidae family and *Proasellus* genus. Groups are
 1233 as follows: for Asellidae: 1 – *Asellus* pattern, 2 – North American asellids, 3 – *Bragasellus* +
 1234 *Synasellus*, 4 – *Proasellus*; for *Proasellus*: slavus – ibero-aquitanian – anophthalmus-coxalis –
 1235 alpine.
 1236



1237
 1238

1239 Figure 4: A-C: Relationships between body size (males and females) and habitat size and
 1240 between sexual dimorphism index (SDI) and habitat size. Data for habitat size correspond to
 1241 the coordinates of species along the first axis of the fuzzy correspondence analysis performed
 1242 on the “habitat trait categories per species” matrix (See SI Table 11). SDI is negative when
 1243 males are larger than females and positive when females are larger than males. The red lines
 1244 represent the phylogenetic generalized least square regressions. All regressions are statistically
 1245 significant. D-F: Violin plots showing the difference in body size (males and females) and
 1246 sexual dimorphism index (SDI) between surface water- and groundwater-habitat specialist
 1247 species. The white dot, thick black bar, and thin black line show the median value, interquartile
 1248 range, and 95% of all data, respectively. Significant *P* values are in bold.



1250
 1251
 1252