

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

Nathanaelle Saclier^{1,2}  | Louis Duchemin¹  | Lara Konecny-Dupré¹  |
Philippe Grison³ | David Eme⁴  | Chloé Martin³ | Cécile Callou³  |
Tristan Lefébure¹  | Clémentine François¹  | Colin Issartel¹ | Julian J. Lewis^{5,6} |
Fabio Stoch⁷  | Boris Sket⁸  | Sanja Gottstein⁹  | Teo Delić⁸  |
Maja Zagmajster⁸  | Michal Grabowski¹⁰  | Dieter Weber^{11,12}  |
Ana Sofia P. S. Reboleira^{13,14}  | Dmitry Palatov¹⁵  | Kaloust Paragamian¹⁶  |
Lee R. F. D. Knight¹⁷ | Georges Michel¹⁸ | Francois Lefebvre¹⁹ |
Mohammad-Javad Malek Hosseini^{20,21}  | Ana I. Camacho²²  |
Begoña Gartzia De Bikuña^{23,24}  | Amina Taleb²⁵  | Nouria Belaidi²⁵  |
Raoul P. Tuekam Kayo²⁶  | Diana Maria Paola Galassi²⁷  |
Oana Teodora Moldovan^{28,29}  | Christophe J. Douady^{1,30}  | Florian Malard¹ 

Correspondence

Florian Malard, Université Claude Bernard
Lyon 1, CNRS, ENTPE, UMR 5023 LEHNA,
Univ Lyon, Villeurbanne F-69622, France.
Email: florian.malard@univ-lyon1.fr

Funding information

Académie de recherche et d'enseignement
supérieur, Belgique; Agence Nationale
de la Recherche, France, Grant/Award
Number: CONVERGENOMICS ANR-15-
CE32-0005 and EUR H2O'Lyon project
ANR-17-EURE-0018; Biodiversa+ DarCo,
Grant/Award Number: GAN°101052342;
Campus France; Consiliul National al
Cercetării Științifice, Grant/Award
Number: 2/2019 DARKFOOD; Fundação
para a Ciência e a Tecnologia, Grant/
Award Number: UIDB/00329/2020;
Javna Agencija za Raziskovalno Dejavnost
RS, Grant/Award Number: P1-0184;
Ministerio de Ciencia e Innovación, Grant/
Award Number: PID2019-110243GB-I00;
Polish Ministry of Science and Higher
Education, Grant/Award Number: N N303
579439 and 5818/B/P01/2010/39; Villum
Fonden, Grant/Award Number: 15471

Handling Editor: Paula Arribas

Abstract

Transition to novel environments, such as groundwater colonization by surface organisms, provides an excellent research ground to study phenotypic evolution. However, interspecific comparative studies on evolution to groundwater life are few because of the challenge in assembling large ecological and molecular resources for species-rich taxa comprised of surface and subterranean species. Here, we make available to the scientific community an operational set of working tools and resources for the Asellidae, a family of freshwater isopods containing hundreds of surface and subterranean species. First, we release the World Asellidae database (WAD) and its web application, a sustainable and FAIR solution to producing and sharing data and biological material. WAD provides access to thousands of species occurrences, specimens, DNA extracts and DNA sequences with rich metadata ensuring full scientific traceability. Second, we perform a large-scale dated phylogenetic reconstruction of Asellidae to support phylogenetic comparative analyses. Of 424 terminal branches, we identify 34 pairs of surface and subterranean species representing independent replicates of the transition from surface water to groundwater. Third, we exemplify the usefulness of WAD for documenting phenotypic shifts associated with colonization of subterranean habitats. We provide the first phylogenetically controlled evidence that body size of males decreases relative to that of females upon groundwater colonization, suggesting competition for rare receptive females selects for smaller, more agile

Christophe J. Douady and Florian Malard contributed equally.

For Affiliation refer page on 14

males in groundwater. By making these tools and resources widely accessible, we open up new opportunities for exploring how phenotypic traits evolve in response to changes in selective pressures and trade-offs during groundwater colonization.

KEYWORDS

collaborative database, comparative analysis, molecular resources, phenotypic evolution, phylogeny, subterranean biodiversity

1 | INTRODUCTION

Homo sapiens has been fascinated by the subterranean world throughout its history (Mammola & Martínez, 2020) and the peculiar features of subterranean organisms have attracted scientists since their first discovery over the 16th–17th centuries (Malard, 2022). However, it was not until the mid-20th century that the idea that the subterranean world provides an excellent research ground for addressing general scientific questions in ecology and evolution gained momentum (Mammola et al., 2020; Poulson & White, 1969). A long-standing perspective of subterranean life evolution is that of convergence whereby phylogenetically distant organisms acquire similar phenotypes because of a convergent selective environment that includes no light, environmental stability and energy limitation (Christiansen, 1961; Pipan & Culver, 2012). Since the 2000s, a broader evolutionary perspective of subterranean life has emerged, one that has also incorporated the role of non-adaptive processes (Lefébure et al., 2017; Policarpo et al., 2021; Wilkens & Strecker, 2017) and divergent selection (Fišer et al., 2023; Trontelj et al., 2012) in shaping the phenotype of organisms.

Phylogenetically controlled and replicated comparisons between closely related surface and subterranean organisms provide ideal models to study evolution during colonization of a novel environment (Protas & Jeffery, 2012; Rétaux & Jeffery, 2023; Saclier et al., 2018). Among subterranean species, some have evolved following independent colonization of subterranean habitats by surface ancestors, whereas others have evolved from ancestors within the subterranean environment (Cooper et al., 2023). Surface organisms that colonize subterranean habitats experience dramatic environmental changes (e.g. darkness, food limitation) and evolve characteristic regressive (e.g. reduced eyes and pigment) and constructive (e.g. increased extra-optic sensory structures) traits (Culver & Pipan, 2019; Hose et al., 2022). Subterranean colonization is considered an irreversible habitat transition because it leads to eye degeneration (Langille et al., 2022; Niemiller et al., 2013). However, recolonization of surface habitats by blind and depigmented animals is possible, especially when surface habitats are characterized by low competitive pressure (Copilas-Ciocianu et al., 2018; Friedrich, 2019; Prendini et al., 2010).

The scientific scope of surface-subterranean comparative studies ultimately depends on the acquisition of ecological and molecular resources—from biological trait data to phylogenetic and genomic resources—in model organisms. These resources are increasingly becoming available at the intraspecific level in species comprised of

surface and subterranean populations such as the teleost *Astyanax mexicanus* (Gross et al., 2023; Kowalko et al., 2020), the isopod *Asellus aquaticus* (Konec et al., 2015; Protas et al., 2023), the amphipod *Gammarus minus* (Fong & Carlini, 2023) and the urodele amphibian *Proteus anguinus* (Kostanjsek et al., 2023). However, comparative studies at the interspecific level remain scarce essentially because of the difficulty in assembling large-scale phylogenetic and species trait data sets in clades comprised of multiple surface and subterranean species (Langille et al., 2022; Lefébure et al., 2017; Mammola et al., 2019; Saclier et al., 2018; Stern et al., 2017). Although intra-specific studies often provide deeper insights into the genetic and developmental basis of phenotypic traits, only interspecific studies can document evolutionary changes taking place over time periods longer than the lifespan of natural populations.

Performing phylogenetic comparative analyses of clades comprised of surface and subterranean species faces several challenges. First, only a few clades of metazoans contain both a high number of surface species and subterranean species because the surface ancestors of many subterranean taxa went extinct (Humphreys, 2000). Candidate clades often have a wide geographic distribution, sometimes spanning several continents, which makes it particularly difficult to obtain biological material (Faille, 2019; Fišer, 2019a; Lukić, 2019; Mammola & Isaia, 2017). Second, the taxonomic units to be used in comparative analyses are not firmly established. Molecular species delimitation methods often reveal highly divergent operational taxonomic units within subterranean described species that have been historically delimited based on morphological criteria (Eme et al., 2018; Fišer et al., 2018). Third, we lack large dated phylogenies of clades with multiple independent subterranean colonization events (but see Ledford et al., 2011; Morvan et al., 2013; Stern et al., 2017). Last, when phylogenetic inferences are available, biological traits for the taxonomic units of interest are often not available in the literature and voucher specimens for measuring those traits are difficult to locate (but see Mammola et al., 2022).

Here, we address the aforementioned challenges by releasing the World Asellidae database (WAD) and phylogeny, a backbone resource to support comparative studies on life evolution in subterranean habitats. The Asellidae (Isopoda, Pancrustacea) is one of the few families of aquatic metazoans containing both surface and subterranean species, thereby potentially providing multiple independent replicates of the transition from surface water to groundwater. First, we describe the guiding principles and content of WAD, a collaborative database specifically designed to promote the joint production and sharing of primary ecological and molecular data and metadata

by multiple research laboratories. Second, we take advantage of new sequence data available in WAD for two mitochondrial genes and two nuclear genes to perform a large-scale dated phylogenetic reconstruction of the Asellidae family that can be used more widely in future comparative studies. Third, we exemplify the usefulness of WAD for documenting phenotypic changes associated with colonization of subterranean habitats. We use the Asellidae phylogeny and body size (BS) data from published articles and morphological measurements made on specimen lots referenced in WAD to test for differences in male and female BS between surface and groundwater habitats. We predict no difference in female BS between habitats because fecundity selection probably favours large-bodied females with large brood sizes in both habitats. In contrast, we predict smaller-bodied males in groundwater than in surface water due to a shift in male mating strategy. In surface water, we hypothesize that competition for synchronously receptive females selects for large males that are more likely to win mating contests (Bertin & Cezilly, 2003). In groundwater, competition for rare, highly asynchronous, receptive females potentially favours smaller, more agile males that are more likely to be successful in finding mates (Andersson, 1994; Blanckenhorn, 2000).

2 | MATERIALS AND METHODS

2.1 | The World Asellidae Database

We use the free and open-source application GOTIT (<https://github.com/gotit-dev/gotit>; Malard et al., 2020) to input, manage and share ecological and molecular data and metadata in WAD. The application manages every step of an everyday laboratory workflow process leading to the production of species occurrence data and DNA sequences. A demo version of GOTIT application is available at <https://gotit.cnrs.fr>. WAD hosts all species occurrence data, sampling and sequencing metadata and biological vouchers (specimens, microscopic slides and DNA extracts) generated over the workflow (Table 1). The database also manages species occurrence data and DNA sequence metadata from the literature and biodiversity facilities, the bibliographic referencing of information and the assignment of DNA sequences to molecular operational taxonomic units (MOTUs). We provide in Figures S1 and S2 the simplified and full logical models of the database respectively. Free and open access to WAD for data end users is available at <https://gotit.univ-lyon1.fr>. Login details are delivered to data contributors upon request from the corresponding author.

2.2 | Species delimitation and dated phylogeny

2.2.1 | Taxon sampling and molecular data

To build the phylogeny, we extracted from WAD an initial molecular data set representing 299 described and undescribed

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

Items	Number
Species occurrence	
Species and subspecies	448
Record	9438
Country	55
Collection material ^a	
Specimen lot	1943
Specimen microscopic slide	1584
Specimen used for DNA extraction	4901
DNA extract	4362
DNA sequencing metadata	
Primer ^b	138
PCR	22,743
Chromatogram	12,052
Sequence ^c	
16S	3562
COI	2866
FASTKD4	922
28S	1202
Opsin	362
Literature reference	641

^aAll specimens and DNA extracts are stored at -20°C .

^bSee Tables S2 and S3.

^cValidated sequences. Numbers differ from the number of sequences used in this study because the database is regularly updated with new data.

morphospecies of Aselloidea containing 278 species of Asellidae and 21 species of Stenasellidae used as outgroup. Together with the Asellidae, the Stenasellidae forms the super family Aselloidea, within the Asellota suborder, which includes three other superfamilies: the Gnathostenetroidoidea, Stenetrioidea and Janiroidea. Our best evidence of a sister relationship between Asellidae and Stenasellidae is from a phylogeny of 64 species of Aselloidea and one species of Janiroidea (i.e. *Jaera praeheirsuta*) reconstructed by Saclier (2019) using 346 nuclear genes and 13 mitochondrial genes. Specimens were collected at 943 localities in Europe, North America, North Africa and Asia (Table S1). Localities spanned a wide range of surface and subsurface freshwater habitats including lotic and lentic surface water bodies, cave streams and pools, the hyporheic zone of surface streams and groundwater in unconsolidated sediments. Throughout this paper, we used the term morphospecies to refer to species, either formally described or undescribed (i.e. waiting a formal description), that were identified based on morphological criteria. Species names of North American asellids follow the latest taxonomic revision to be published by Lewis et al. (2023). For morphological identification of specimens to species level, we relied on the shape of male copulatory organs

(pleopods 2), plus secondary characteristics including the morphology of the male pereopods 1 and 4, pleopods 3–5 and uropods (Lewis et al., 2023). We dissected copulatory pleopods 1 and 2 of male specimens and mounted them on slide for examination using a compound microscope.

We used the Chelex protocol of Casquet et al. (2012) to extract DNA from specimen. We incubated three pereopods of each specimen in a solution of 150 µL of 7% chelex and 10 µL of proteinase K at 15 mg/mL for 90 min at 56°C, and then 15 min at 90°C. We amplified DNA using primers targeting the mitochondrial cytochrome oxidase subunit I (COI) gene, the 16S mitochondrial rDNA gene, the *FASTKD4* nuclear gene and the 28S nuclear rDNA gene. We provide in Tables S2 and S3 the list of all PCR primers, among which 66 were specifically designed as part of this study. For the two mitochondrial genes, we applied several methods to prevent misleading inclusion of nuclear mitochondrial DNA segments (NUMT) in the data set, including different primer pairs, long-range amplification and pre-PCR dilution of genomic DNA (Calvignac et al., 2011).

We amplified 16S fragments with two independent pairs of primers (Tables S2 and S3). PCR 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, 30 s at 72°C and one step of 5 min at 72°C. We performed PCRs for COI fragments using a previously optimized protocol (Calvignac et al., 2011), but with a Taq polymerase (Eurobiotaq) amount of 0.05 U instead of 0.15 U and a PCR volume of 25 µL instead of 35 µL. We used the following PCR settings: one step of 3 min at 95°C, 37 cycles of 20 s at 95°C, 30 s at 51°C, 45 s at 72°C and one step of 5 min at 72°C. A semi-nested PCR was performed whenever the first amplification failed. Using the first PCR product as DNA template, we performed a second PCR using one of the two primers used in the first PCR and another, different primer. The second round PCR was run on 1 µL of the first round PCR product, using the same settings as above but 35 cycles. We amplified *FASTKD4* fragments using several pairs of primers (Tables S2 and S3) with the following PCR settings: one step of 5 min at 95°C, 38 cycles of 30 s at 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 performed a semi-nested PCR whenever the first amplification failed. We completed PCRs for 28S fragments with two independent pairs of primers in order to detect divergent copies. We used the following PCR settings: one step of 3 min at 95°C; 37 cycles of 30 s at 95°C, 30 s at 62°C, 30 s at 72°C and one step of 5 min at 72°C.

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

2.2.2 | Molecular operational taxonomic units

We delimited MOTUs based on a COI alignment of 1385 haplotypes, which were defined from the sequences obtained from 2093

specimens belonging to 299 morphospecies of Aselloidea (Tables S1 and S4). We used the following procedure to select specimens for which we obtained COI sequences. Whenever possible, we first obtained 16S sequences from three specimens of each morphospecies present at a site. Second, we obtained a COI sequence for each specimen whose 16S sequence differed by more than five nucleotides (i.e. 1% of the 16S-nucleotides) with any 16S sequence of the two other specimens.

We used the fixed COI threshold method (TH) implemented by Lefebvre et al. (2006) for crustaceans, and the Poisson tree processes (PTP) proposed by Zhang et al. (2013), to delimit MOTUs. The TH method was previously used in several studies for delimiting species of asellids (Eme et al., 2013, 2018; Morvan et al., 2013). It is based on the observation made from hundreds of COI sequences of crustaceans that two clades diverging by more than 0.16 substitution per site, as measured by patristic distances, have a strong probability (i.e. 99%) of belonging to different described morphospecies. It is a conservative method insofar as it identifies both fewer MOTUs and MOTUs that are more divergent than tree-based methods such as the PTP method (Eme et al., 2018). We applied the TH and PTP methods on a COI haplotype alignment in which the longest sequence with the fewest ambiguities was retained as the best representative sequence for any given haplotype. We constructed a COI haplotype phylogeny in maximum likelihood with PhyML (Guindon et al., 2010) using the general time reversible model of evolution (GTR; Tavaré, 1986) with the G (heterogeneous rates across sites following the gamma distribution) and I parameters (proportion of invariant sites) estimated by maximum likelihood. We used Stenasellidae as outgroup. We computed patristic distances from this phylogeny with the R package 'ape' (Paradis et al., 2004) and delimited MOTUs according to the TH method with the 'cluster' package (Maechler et al., 2002). To delimit MOTUs with the PTP method, we ran mPTP v0.2.2 (<https://github.com/Pas-Kapli/mptp>) using 400,000 MCMC generations, with a thinning of 400 and 0.1 (10%) burn-in. Using or discarding outgroup species (Stenasellidae) did not affect the delimitation of MOTUs within the Asellidae.

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

2.2.3 | Four-gene alignment and phylogeny.

We produced alignments of the COI, 16S, *FASTKD4* and 28S genes for 424 MOTUs of Aselloidea delimited with the TH method (Tables S5–S9). Using MOTUs delimited with the TH method rather than the PTP method limited the risk of considering two populations of the same species as belonging to two distinct MOTUs. In each alignment, we retained the longest sequence with

the fewest ambiguities to represent each MOTU, using the chimera assembler script (https://github.com/TristanLefebvre/chime_ra_assembler.pl). We aligned the 28S and 16S genes with MAFFT Q-INS-i using default parameters (Katoh & Standley, 2013) and the COI and FASTKD4 genes with PRANK codon (Löytynoja & Goldman, 2008). Sites ambiguously aligned were removed with Gblocks (Castresana, 2000). We checked all alignments by eye and made minor editing manually. We used the four genes for 424 MOTUs to build a phylogeny with PhyloBayes (Lartillot et al., 2009) under a CAT-GTR model of evolution. To guarantee the absence of polytomy, a threshold of 10% was set to obtain the majority consensus tree, meaning that each clade must be found in at least 10% of the trees of the Markov process after burn-in. We computed *posterior probabilities* to estimate the *support* of tree topologies and rooted the tree using species of Stenasellidae as outgroup.

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

2.2.4 | Time-scale phylogeny

In the absence of fossil records for the Aselloidea, we used well-identified paleobiogeographic events to constrain the age of 17 nodes in the phylogeny. We provide in Table S10 a detailed description of these events. Palaeobiogeographic calibration points spanned a period ranging from 300 to 2 Myr before present. They were distributed within the phylogeny as follows (Figure 3): one calibration point for the root of the tree (divergence between Stenasellidae and Asellidae <300 Myr), one within the Stenasellidae (≥29 Myr), one within the *Asellus* pattern (sensu Henry & Magniez, 1995; ≤27 Myr), one within the *Baicalasellus* (≤27 Myr), one within the *Bragasellus* (≤75 Myr), 11 within the *Proasellus* (10 points from ≤2 to ≤70 Myr and one point ≥29 Myr) and one at the divergence node between *Gallasellus* and *Conasellus* (≥54 Myr). We estimated divergence times among Aselloidea with PhyloBayes using a CAT-GTR+G+I model, the 17 calibration points as soft bounds, a birth-death prior on divergence time and a log-normal autocorrelation of the substitution rates among branches (Lepage et al., 2007). The effect of any given calibration point on divergence time estimates was assessed by removing that given calibration point during time tree reconstruction.

2.3 | Comparative phylogenetic analyses of BS

2.3.1 | BS and sexual BS dimorphism

Here, we provide a case study of BS and sexual BS dimorphism to show how WAD resources and the World Asellidae phylogeny allow exploring how phenotypic traits evolve upon groundwater colonization. We completed literature data with laboratory measurements made on specimen lots contained in WAD to quantify the maximum BS of adult males and females of 162 asellid MOTUs included in the World Asellidae phylogeny (Table S11). We defined BS as the distance between the anterior margin of the cephalon and the posterior margin of the pleotelson (Prevorčnik et al., 2004). Maximum BS provides an estimator of the size of full-grown specimens in a species: It avoids including immature specimens and is often the only measurement provided in publications. For each MOTU, we provide in Table S11 our best estimate of the number of specimens used for quantifying BS, as the exact number is not always reported in the source articles. For measurements made on specimen lots contained in WAD, we took photos of specimens with a DP25 Olympus camera connected to a dissecting microscope (SZX16 Olympus) and measured BS using ImageJ (Schneider et al., 2012).

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

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

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

2.3.2 | Habitat specialization and habitat size

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

We assessed the size of habitat or pore volume available to species because it is potentially a major determinant of maximum BS (Pipan & Culver, 2017). We used a fuzzy coding approach to assess habitat size because most groundwater ecological studies do not provide any quantitative estimates of pore volume available to species. Fuzzy coding is a coding procedure whereby scientific experts describe the affinity of species for the modalities of ecological variables on the basis of existing biological knowledge. Since its introduction by Chevenet

et al. (1994), fuzzy coding has been widely used to produce species trait databases for a number of taxa (Degen et al., 2018; Kunz et al., 2022; Sarremejane et al., 2020). For the 162 asellid MOTUs incorporated in the comparative analyses, we attributed positive scores (from 0=no affinity to 3=strong affinity) to three categories of habitat size (large, medium and small pore volumes). We attributed habitat size scores independently of the eye and pigmentation status of species. Hence, we assigned a high affinity for large size habitats to species living in the benthic layer of surface streams as well as to those living in the benthic layer of cave streams. Scores were attributed to all species separately by two of us (F.M. and J.J.L.) using species occurrence data per habitat category as guideline data (data extracted from WAD). Then, we corrected for inconsistencies between the two sets of scores to produce a single 'habitat trait categories per species' matrix. We provide the species habitat scores and the scoring procedure in Table S11 to ensure data traceability and reproducibility, and potential revision of scores in the event of new habitat data of species. Then, we performed a fuzzy correspondence analysis (COA) of the 'habitat trait categories per species' matrix (Chevenet et al., 1994) and used the coordinates of species along the first axis of the COA, representing 85% of total variability, as quantitative surrogates of their habitat size. The COA was performed using the R package 'ade4' (Thioulouse et al., 2018).

2.3.3 | Data analysis

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

3 | RESULTS

3.1 | The World Asellidae Database

The database contains 9438 distributional records for 163 surface water species and 285 groundwater species of Asellidae belonging to 23 genera (Tables 1 and 2). Asellids are widely distributed in the Northern Hemisphere with species belonging to four formerly recognized groups of morphospecies, which occupy distinct but partially overlapping distribution ranges (Figure 1). All four groups

include both surface and groundwater species, although in different proportions. (Table 2). The first group is the 'Asellus pattern', so named by Henry and Magniez (1995) in reference to the specific shape of copulatory organs shared by several genera of Asellidae. It has nine genera (61 species); all distributed in Asia and Northwestern America, except the genus *Asellus*, which is also represented in Europe by the widespread *A. aquaticus* species complex (Verovnik et al., 2005). The second group to which we refer as the North American asellids include seven genera (152 described species), all located in North America, except *Gallasellus* and *Baicalasellus*, which are endemic to western France and Lake Baikal (Russia) respectively. The third group containing the two genera *Bragasellus* and *Synasellus* (56 species) is endemic to the Iberian Peninsula. The fourth group corresponding to the genus *Proasellus* (174 species) extends from southern Scandinavia to northern Africa and from Portugal to Iran.

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

WAD provides metadata—from sampling of specimens to PCR and chromatogram settings—for 8914 validated sequences of Asellidae belonging to two mitochondrial and three nuclear genes (Table 1). Of these, we submitted 3698 sequences to NCBI as part of the present study (Table S12) and 3676 sequences as part of other studies. In WAD, COI sequences are assigned to MOTUs using different molecular species delimitation methods. The geographic distribution of MOTUs within morphospecies can be visualized using ready-to-use queries implemented in GOTIT application (Figure S3).

3.2 | The Asellidae timetree

3.2.1 | Molecular operational taxonomic units

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

3.2.2 | Time-scale phylogeny

The phylogeny included 384 MOTUs of asellids delimited with the TH method. They collectively represented 268 morphospecies,

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

Morphospecies groups/genera	WAD		Asellidae timetree				
	Described species		Morphospecies ^e		MOTUs		Species pairs ^f
	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> ^a	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 ^b	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— <i>Bragasellus</i> & <i>Synasellus</i>	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— <i>Proasellus</i> ^c	61	113	54	108	51	181	21
Others ^d		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	34

^a*Nomen nudum* in Hidding et al. (2003).^bGenera according to recent revision by Lewis et al. (2023).^cIncluding *Chthonasellus bodoni* (Argano & Messana, 1991).^dGenera that cannot be assigned to any of the four species groups.^eNumbers include undescribed morphospecies.^fNumber of independent species pairs containing a surface water and a groundwater asellid species.

among which 195 were formally described (Figure 3, Table 2). The tree topology recovered the monophyly of the four main species groups described above and that of all asellid genera, with posterior probabilities >0.9, except the genus *Conasellus* ($PP=0.43$; Figure 3, Figures S4 and S5). The *Asellus* pattern (group 1 in Figures 1 and 3) formed a sister clade to the rest of a larger clade comprised of the North American asellid clade (group 2), the *Bragasellus*+*Synasellus* clade (group 3) and the *Proasellus* clade (group 4). However, relationships among the later three clades were not resolved. The tree

topology for *Proasellus* was also consistent with earlier subdivisions of this species-rich genus into four clades (Morvan et al., 2013). Within that genus, the slavus clade was sister to a larger clade comprised of the ibero-aquitania, anopthalmus-coxalis and Alpine clades, but the relationships among the later three clades were not resolved (Figure 3).

Divergence time estimates were robust to the removal of any single palaeobiogeographic calibration point, except the deepest one that constrained the divergence between Stenasellidae and

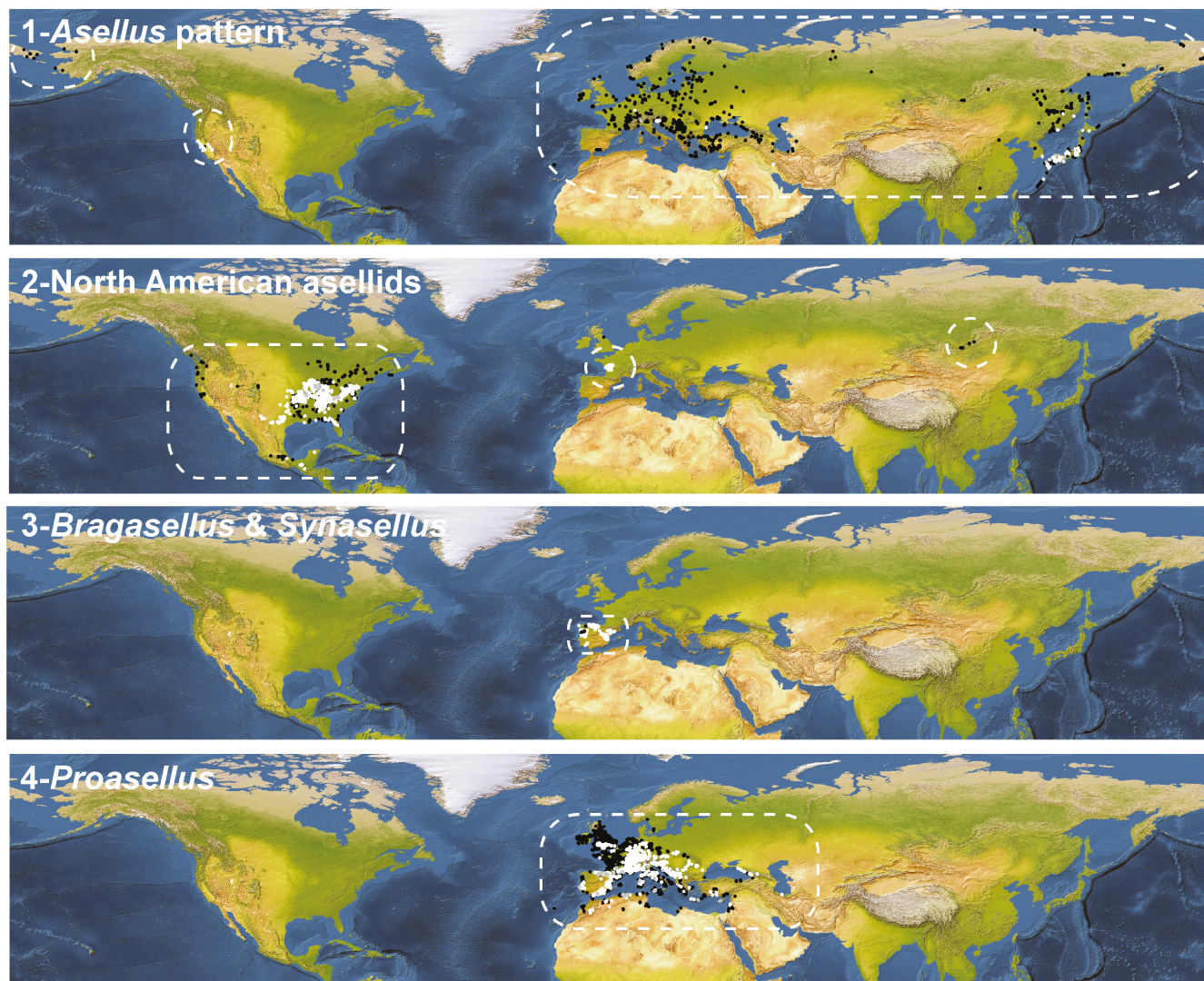


FIGURE 1 Distribution of four major species groups of Asellidae (Isopoda, Pancrustacea). Dots are species occurrence data contained in the World Asellidae Database (black dots: surface water species; white dots: groundwater species). Background map is extracted from version 2.1.0 of Natural Earth II: free vector and raster map data @[naturalearthdata.com](https://www.naturalearthdata.com).

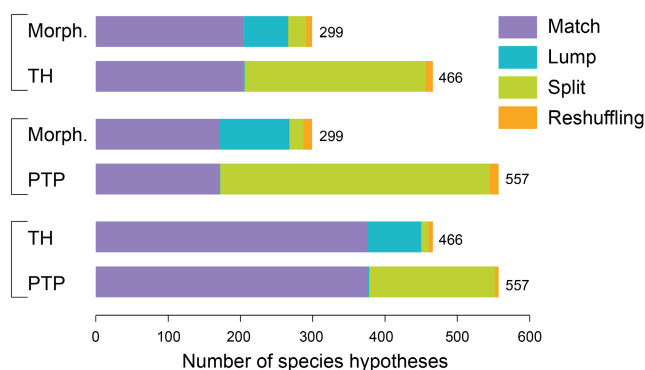


FIGURE 2 Pairwise taxonomic comparisons between the three different sets of Aselloidea species hypotheses delimited using morphology (Morph.), a COI divergence threshold (TH) and the Poisson tree processes model (PTP).

Asellidae to be more recent than 300 Myr (Figure S6). Removing this point yielded older divergence times, notably pushing back the divergence between the Stenasellidae and Asellidae to 300 Myr (95% credibility interval [CI]: 415–222 Myr) instead of 139 Myr (CI: 174–106 Myr), when including it (Figure 3, Figures S5 and S6). The diversification of Asellidae might have started in early Cretaceous (132 Myr, CI: 168–102 Myr) and that of *Proasellus* at the end of Cretaceous or beginning of the Paleogene (72 Myr, CI: 88–58 Myr).

We identified up to 34 independent pairs of surface and groundwater asellid species in the phylogeny (Figure 3, Figure S4). Species pairs were present in all four major groups of asellid species, although several species-rich clades were almost exclusively comprised of groundwater species, including the Alpine *Proasellus* clade, *Synasellus* and *Caecidotea*. The uneven distribution of species pairs among the *Proasellus* (21 pairs), North American asellids (10 pairs)

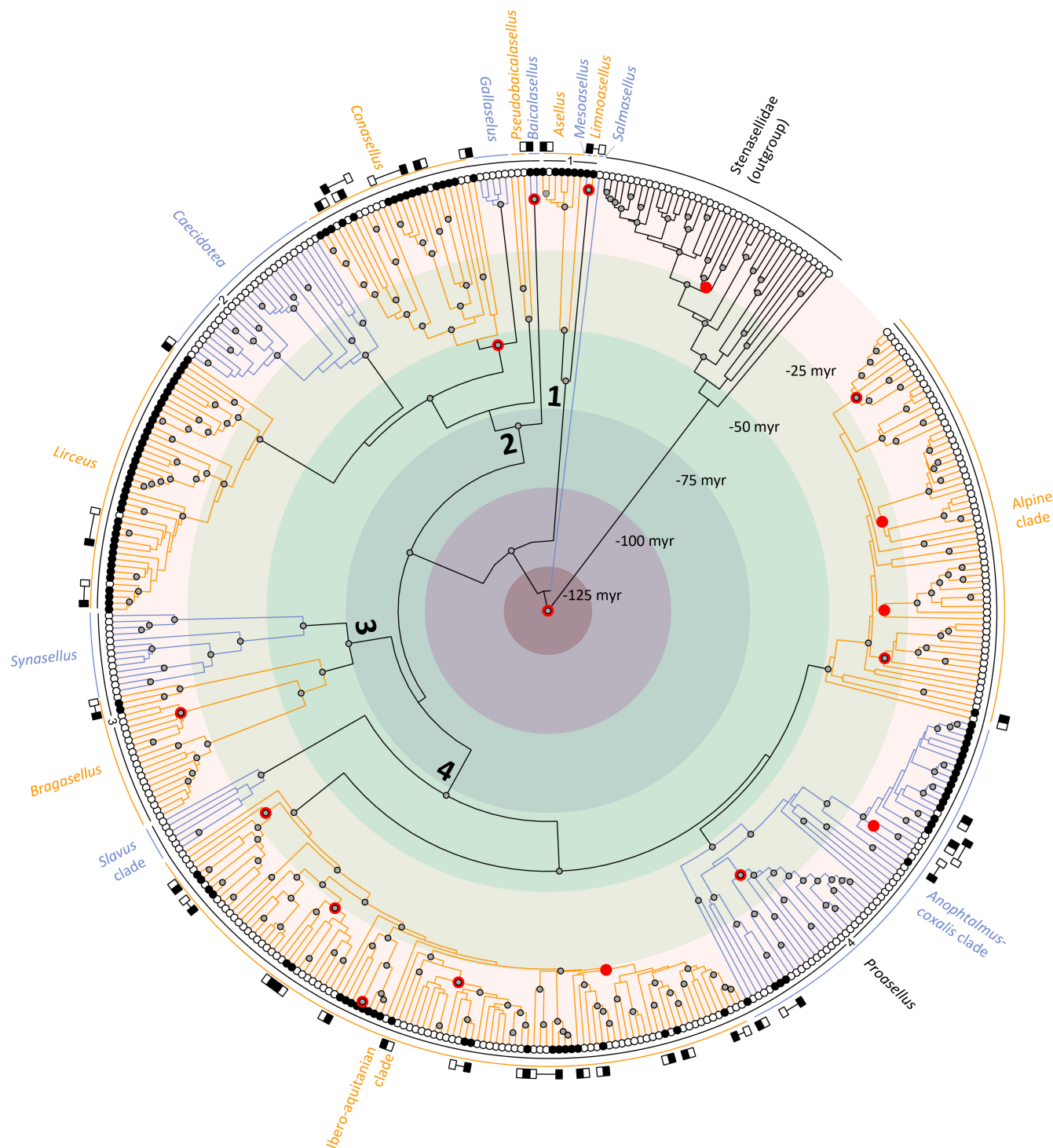


FIGURE 3 Timetree of Asellidae (Isopoda, Pancrustacea). The tree is rooted using Stenasellidae as outgroup. Terminal nodes are MOTUs as delimited with the fixed COI threshold method (TH) implemented by Lefebure et al. (2006). White terminal nodes are eyeless and depigmented MOTUs; black terminal nodes are oculated and/or pigmented MOTUs. Colour rings show time. Red and grey dots show palaeobiogeographic calibration points and node supports with posterior probabilities >0.9 respectively. Black and white squares on the outer ring show independent pairs of surface (black) and groundwater (white) asellid species (see definition of species pairs in materials and methods). Legends show genera and main species groups within the Asellidae family and *Proasellus* genus. Groups are as follows: for Asellidae: 1—*Asellus* pattern, 2—North American asellids, 3—*Bragasellus* + *Synasellus*, 4—*Proasellus*; for *Proasellus*: slavus—ibero-aquitania—*Anophtalmus-coxalis*—alpine.

and the *Asellus* pattern (one pair) essentially reflected a too low sampling in the latter two groups (Table 2). Transitions to groundwater have probably occurred throughout the evolutionary history of the

Asellidae (Figure 3, Figure S4). Using the speciation event leading to a congeneric species pair as a surrogate of the transition time to groundwater (but see Section 4), some transitions occurred less than

10 Ma (6.9, CI: 11.4–4.2 Myr for the *A. aquaticus*–*A. kosswigi* species pair), whereas others potentially occurred much longer ago (at most 40.4, CI: 50.9–30.4 Myr for the *Conasellus burkensis*–*Conasellus red-delli* species pair; Figure S4).

3.3 | Comparative phylogenetic analyses of BS

Female and male BS ranged from 2.3 to 18 mm (mean = 6.3 ± 2.7 mm, $n = 162$ MOTUs) and from 2.1 to 25 mm (mean = 7.3 ± 4.1 mm, $n = 162$ MOTUs) respectively (Table S11 and Figure S7). BS increased significantly with habitat size, both for females and males (Table 3, Figure 4, Table S13). Species colonizing open habitats, both above (e.g. surface lakes and streams) and below ground (e.g. cave streams), had larger BS than species colonizing interstitial habitats (e.g. groundwater in unconsolidated sediment). The effect of habitat specialization on BS was gender dependent (Table 3, Figure 4, Table S13). Male BS was significantly smaller in groundwater species than in surface water species, whereas we found no significant differences in female BS between surface water and groundwater species. However, habitat specialization accounted for a smaller proportion of variance in male BS (Cox–Snell $R^2 = .074$) than habitat size ($R^2 = .161$; Table S13). We found no interactions between the effects of habitat size and habitat specialization on male and female BS, indicating that constraints imposed by the size of habitats on BS applied similar to eyeless and depigmented species and oculated and/or pigmented species.

Asellids showed substantial variation in SBSD among species. Of the 162 species examined in this study, 94 (58%) exhibited male-biased dimorphism, 62 (38.3%) exhibited female-biased dimorphism and 6 (3.7%) were monomorphic for BS. We found a significant effect of habitat specialization on SBSD: mean SDI was -0.33 ± 0.22 ($n = 61$) and 0.01 ± 0.28 ($n = 101$) for surface water species and groundwater

species respectively (Table 3, Figure 4, Table S13). Males were larger than females in 57 of 61 (i.e. 93.4%) surface water species examined in this study, whereas they were larger than females in only 37 of 101 (i.e. 36.6%) groundwater species. The SDI decreased significantly with increasing habitat size (Table 3, Figure 4). Male-biased SBSD (SDI < 0 in Figure 4) predominated in open habitats whereas female-biased SBSD (i.e. SDI > 0) predominated in interstitial habitats. However, habitat size accounted for a smaller proportion of variance in SBSD (Cox–Snell $R^2 = .069$) than habitat specialization ($R^2 = .110$). We found no interactions between the effect of habitat size and habitat specialization on SBSD.

4 | DISCUSSION

4.1 | The World Asellidae Database

Collaborative databasing has become essential to biodiversity sciences because the amount of data and biological material needed to address broad-in-scope questions exceeds the production capabilities of even the most performing laboratories (Fišer, 2019b; Hobern et al., 2012; Nelson et al., 2011). The structure of the database used in the present study and its web application GOTIT have been conceived to provide scientists with an efficient tool to jointly produce multiple-species ecological and molecular resources to study life evolution in groundwater. The tool has been used with success since 2017 to amass worldwide data at an unprecedented rate for the Asellidae. WAD provides to date one of the most important resource of species occurrence, DNA sequences and biological material for testing eco-evolutionary hypotheses pertaining to groundwater colonization using comparative phylogenetic methods and evolutionary model fitting (Langille et al., 2022; Lefebvre et al., 2017; Saclier et al., 2018; Stern et al., 2017).

TABLE 3 Results of phylogenetic generalized least-squares regression models for testing the effects of habitat size and specialization (i.e. surface vs. groundwater habitats) on body size of females and males and sexual dimorphism index.

Dependent variable	Explanatory variable	Parameter estimate	SE	t	p
Male body size	Intercept (groundwater)	1.871	0.401	4.662	
	Habitat size	0.186	0.044	4.242	<.001
	Habitat specialization	0.166	0.080	2.084	.0387
	Habitat size × habitat specialization	−0.039	0.091	−0.427	.670
Female body size	Intercept (groundwater)	1.768	0.354	4.994	
	Habitat size	0.129	0.039	3.350	.001
	Habitat specialization	−0.008	0.070	−0.111	.912
	Habitat size × habitat specialization	−0.008	0.080	−0.098	.922
Sexual dimorphism index	Intercept (groundwater)	−0.134	0.289	−0.472	
	Habitat size	−0.072	0.032	−2.268	.025
	Habitat specialization	−0.203	0.057	−3.534	.001
	Habitat size × habitat specialization	0.044	0.065	0.676	.500

Note: Significant *p* values are in bold.

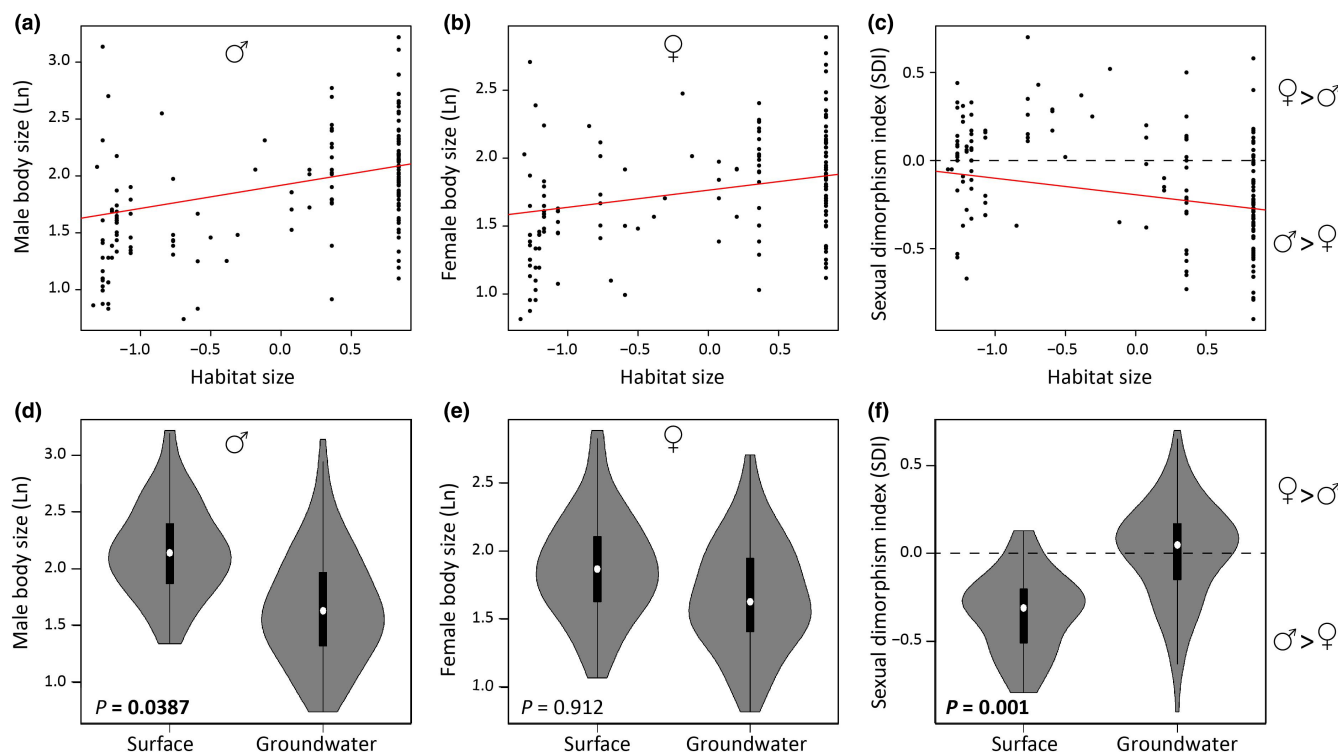


FIGURE 4 (a–c) Relationships between body size (males and females) and habitat size and between sexual dimorphism index (SDI) and habitat size. Data for habitat size correspond to the coordinates of species along the first axis of the fuzzy COA performed on the ‘habitat trait categories per species’ matrix (Table S11). SDI is negative when males are larger than females and positive when females are larger than males. The red lines represent the *phylogenetic generalized least square regressions*. All regressions are statistically significant. (d–f) Violin plots showing the difference in body size (males and females) and sexual dimorphism index (SDI) between surface water- and groundwater-habitat specialist species. The white dot, thick black bar and thin black line show the median value, interquartile range and 95% of all data respectively. Significant *p* values are in bold.

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

Current development efforts are following two main directions. The first direction is widening the database structure for housing

biological species trait data, including but not limited to morphological traits, which are measured on referenced specimens (see, e.g. Lefébure et al., 2017 and Saclier et al., 2018 for data on genome size and rate of molecular evolution respectively). The second direction consists in providing user-friendly tools to promote expertise sharing among users. One example expert tool could guide sequence producers in selecting the most appropriate primers for sequencing any given species from the hundred primers available in the database (see Tables S2 and S3).

4.2 | Species delimitation and dated phylogeny

The presence of highly genetically divergent units (i.e. MOTUs)—often referred to as cryptic species—within morphospecies is a common phenomenon across most animal taxa (Bickford et al., 2007; Pfenninger & Schwenk, 2007), and asellids do not escape the rule (Eme et al., 2013; Morvan et al., 2013). Hence, species molecular delimitation methods based on the *COI* gene typically provide more species hypotheses—in the present study 1.6–1.9 times as many—than morphological delimitation. Although molecular methods typically split asellid morphospecies into several MOTUs, they very rarely reshuffle MOTUs among morphospecies. Using different elementary species units in biodiversity

research can provide novel insights into the mechanisms underlying biodiversity patterns (Fišer et al., 2018). In their analysis of the range size pattern of groundwater Asellidae and Niphargidae (Amphipoda) in Europe, Eme et al. (2018) showed that using MOTUs instead of morphospecies reinforced the Rapoport effect of increasing range size at higher latitudes and increased the proportion of variance in range size explained by historical climates. In WAD, we are continuously updating the geographic coverages of COI sequences and MOTUs within morphospecies (see Figure S3), thereby accumulating data for rigorously testing the hypothesis that groundwater species have a reduced range size compared to their surface counterparts. Despite being a long-standing hypothesis (Malard et al., 2023), the crayfish study by Stern et al. (2017) remains the only phylogenetically controlled test to date, even though the authors used morphospecies rather than MOTU-level data. WAD also provides one of the most comprehensive reference barcode libraries of groundwater taxa for accurately assigning to existing known species the COI sequences that arise from a growing number of DNA-based biodiversity studies (Zagmajster et al., 2022). Such a WAD reference barcode library offers great opportunities to combine environmental DNA sampling, metabarcoding, DNA taxonomy and traditional taxonomy to speed up the acquisition of species occurrence data in difficult to access groundwater habitats (Fontaneto et al., 2015; Saccò et al., 2022; Verdier et al., 2022).

The World Asellidae phylogeny provides one of the most comprehensive phylogenetic frameworks available to date for undertaking comparative studies on evolution to groundwater life (but see also Stern et al., 2017). Here, we highlight key improvements to the phylogeny since a previous version published by Morvan et al. (2013). First, the present version of the phylogeny contains 2.5 and 2.4 times more MOTUs and morphological species of asellids, respectively, than its previous version. Its geographical coverage is also considerably wider, as it includes not only European species but also many North American and eastern Mediterranean species. Yet, the phylogeny is far from being complete since it presently contains 60% of described species of asellids, the most species-deficient group being the *Asellus* pattern in Asia with only 13% of described species included in the phylogeny. Second, we improved dating of divergence times in the phylogeny by adding 14 palaeobiogeographic calibration points to the three points originally used by Morvan et al. (2013). This addition resulted in overall younger divergence times. Thus, in the present phylogeny, the early diversification of the four *Proasellus* clades is dated to the Palaeogene and not to the Upper Jurassic, as estimated by Morvan et al. (2013). However, palaeobiogeographic calibration points are still relatively unevenly distributed across the phylogeny, with only a single point for the North American, albeit species-rich, clade. Adding new calibration points to this clade would require sampling US regions where species-rich clades might have diversified 'on place' following emergence of lands from the sea (e.g. eastern Texas, Florida and Chesapeake Bay). Third, in comparison with Morvan et al. (2013)'s phylogeny, we more than doubled the number of replicates of groundwater evolution by identifying 34 independent pairs of surface and groundwater asellid species, among which 21 occur within the genus *Proasellus*. Further sampling will likely provide additional species pairs within the *Asellus* pattern and North American

asellids, thereby providing a more even distribution of groundwater transitions among three of the four major groups of asellids. Obtaining many replicate species pairs is crucial to robust testing of common principles of groundwater evolution while accounting for the effects of local contingencies. Up to now, comparative studies have relied on few replicates of evolution to groundwater life—that is, on 3–13 independent species pairs—for assessing changes in the evolution of genome size and rate of molecular evolution in asellids (Lefébure et al., 2017; Saclier et al., 2018); vision genes in beetles, crayfishes and fishes (Langille et al., 2022; Policarpo et al., 2021; Stern & Crandall, 2018); and gene repertoires in beetles (Balart-García et al., 2023).

Another desirable attribute of a biological study system for understanding trait evolution in groundwater is to have species that have colonized groundwater for different lengths of time. Time is undoubtedly an important factor controlling the evolution of traits, especially those that evolve at least in part under relaxed selection, such as the regression of eyes in subterranean animals (Culver et al., 2023; Langille et al., 2022; Policarpo et al., 2021; Wilkens & Strecker, 2017). Among the asellids, depigmented and reduced eye subterranean populations of the surface species *A. aquaticus* colonized groundwater less than 100,000 years ago (Protas et al., 2023; Protas & Jeffery, 2012), whereas some eyeless and depigmented species of *Proasellus* have resided in groundwater for over 10 million years (Lefébure et al., 2017). However, in a phylogeny, it is usually unclear at which point along a terminal branch leading to a groundwater species colonization of groundwater exactly occurred. Specifically, groundwater colonization may be much more recent than the speciation event leading to a pair of surface and groundwater species if now-extinct surface species have persisted long after that speciation event. In asellids, a promising approach is to use the pseudogenization of genes coding for opsin light-sensitive proteins to estimate the groundwater colonization time, assuming that loss-of-function mutations accumulate early in the process of groundwater colonization. In a study by Lefébure et al. (2017), colonization time was measured for 19 asellid species as a function of the speciation time and an estimate of the pseudogenization of the *opsin* genes on branches leading to subterranean species. Increasingly sequencing the *opsin* genes across asellid species (see Table 2) paves the way for accounting for the effect of colonization time on the evolution of phenotype in comparative studies. Of note, however, the pseudogenization approach to dating groundwater colonization times reaches its limits when the genes fail to be detected in the genome, either by direct PCR, exon capture methods or whole genome sequencing. The genes decay and are completely lost over time in species that have presumably spent a too long period of time underground (Langille et al., 2022; Lefébure et al., 2017).

4.3 | Comparative phylogenetic analyses of BS and sexual BS dimorphism

Our phylogenetic comparative study of BS and SBS between surface- and groundwater species illustrates the usefulness of WAD

for documenting evolutionary changes during transition to novel habitats. We found that BS in asellids was constrained by the size of habitat in both sexes. This corroborates Pipan and Culver (2017)'s hypothesis that BS within clades containing subterranean species is in part controlled by habitat volume because pore size between rocks can set an upper limit to maximum BS.

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

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

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

A recent morphological study by Balázs et al. (2021) investigated sexual dimorphism in 17 morphological traits, including BS, using nine surface and six cave groundwater populations of *A. aquaticus* showing various degrees of reduction of eyes and body pigments (see also Biró et al., 2022). The authors showed that several

morphological traits were significantly less male-biased in cave than in surface populations (e.g. the shape of pereopods I). However, contrary to the present study, they found no significant reduction in male-biased dimorphism in BS upon cave groundwater colonization. A potential explanation is that the intraspecific comparative study by Balázs et al. (2021) may have been unable to detect a reduction in male-biased SBS in cave populations of *A. aquaticus* due to insufficient time for BS to evolve. Of note, males were reported to be smaller than females in several depigmented and eyeless subterranean *Asellus* species including *A. amamiensis*, *A. hyugaensis*, *A. primoryensis* and *A. tamaensis* (Henry & Magniez, 1993; Matsumoto, 1960, 1961, 1963). A potentially important proportion of the variance in SBS exhibited by groundwater species might be due to differences in groundwater colonization time among species. If so, using colonization time as a predictor instead of a qualitative present-day biological status (i.e. eyeless and depigmented vs. oculated and pigmented) would contribute to a better understanding of trait changes associated with groundwater transitions. This may become possible in the near future as sequences of genes accumulating loss-of-function mutations during colonization (e.g. *opsin* genes, see Lefébure et al., 2017) become available for a large number of species.

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

Beyond BS, WAD provides many of the necessary resources for testing predictions on how phenotypic traits linked to mating success, fecundity and survival evolve in response to changes in selective pressures and trade-offs during groundwater colonization. We provide below three example predictions. First, if searching for rather than fighting for mates is key to determining mating success of male groundwater species, then selection is likely to target sensory organs that improve the ability of males to find females. More specifically, the hundreds of specimens referenced in WAD can be used to test whether males of groundwater species have longer antennae, relative to BS, than surface males and groundwater females, because long antennae are advantageous for detecting receptive females (Balázs et al., 2021; Bertin & Cézilly, 2003). Second, life-history theory predicts that relative to their BS, groundwater, iteroparous species should produce fewer but larger eggs per reproductive event than surface, semelparous species (Fišer, 2019a; Venarsky et al., 2023). WAD keeps full records of the number of ovigerous females contained in hundreds of specimen lots for testing this hypothesis. Third, WAD resources can be used to test for

the occurrence of a trade-off between transient fecundity (i.e. the number of offspring produced per brood per single reproductive event) and adult survival in long-lived, iteroparous groundwater species. Fecundity selection favours increase in BS, whereas selection for survival may favour narrow and elongated body shapes that allow individuals to withdraw into tiny hiding places to escape predators (Fišer et al., 2013, 2019; Miller, 1933). A trade-off may arise because an elongated brood pouch prevents good ventilation of eggs beyond a certain BS. If such a trade-off exists, we predict variation in BS to be more evolutionarily constrained in groundwater females than in surface females. This prediction can be tested by comparing best-fit evolution models of BS and shape between habitats and sexes.

5 | CONCLUSION

The asellids fulfil many of the desirable attributes of a model animal system for studying evolution during colonization of a new environment, in particular here groundwater. Recently, Protas et al. (2023) synthesized the ecological and molecular resources available for studying microevolutionary dynamics of groundwater colonization from multiple cave and surface populations of *A. aquaticus*. Here, we make available to the scientific community a comprehensive set of taxonomic, distributional and molecular resources and biological material that have been acquired for studying macroevolutionary dynamics of groundwater colonization from multiple-species data. Looking at trait variation among multiple independent colonization events across a wide range of times since colonization can provide better understanding into the temporal dynamics of phenotypic and genomic evolution.

AUTHOR CONTRIBUTIONS

All authors revised the manuscript and approved the publication. Nathanaelle Saclier, Florian Malard, Christophe J. Douady and David Eme were involved in manuscript conception, writing, editing and illustrations. Leaders—Florian Malard, Philippe Grison, Louis Duchemin and contributors—Lara Konecny-Dupré, Tristan Lefébure, Christophe J. Douady, Chloé Martin, Cécile Callou, David Eme were involved in database structure conception. Philippe Grison and Louis Duchemin were involved in writing of code for the database and web application. Florian Malard and Lara Konecny-Dupré were involved in database management. Lara Konecny-Dupré, Christophe J. Douady, Tristan Lefébure and Clémentine François were involved in molecular data acquisition and management. Nathanaelle Saclier, Christophe J. Douady and Florian Malard were involved in phylogenetic and comparative analyses. Colin Issartel, Florian Malard and Nathanaelle Saclier were involved in acquisition of body size data. The authors hereinafter largely contributed to sampling and identification of biological material for the following groups and/or regions: North American asellids/North America: Julian J. Lewis; multiple sites in the European Union: Nathanaelle Saclier, Florian Malard, Christophe J. Douady, Clémentine François,

David Eme, Tristan Lefébure; coxalis group (*Proasellus*): Fabio Stoch; The Balkan Peninsula: Boris Sket, Sanja Gottstein, Teo Delić, Maja Zgarnajster, Michal Grabowski; Germany/Luxemburg: Dieter Weber; Portugal: Ana Sofia P. S. Reboreira; Russia: Dmitry Palatov; Crete: Kaloust Paragamian; United Kingdom: Lee R. F. D. Knight; Belgium: Georges Michel; Nouvelle Aquitaine (France): Francois Lefebvre; Iran: Mohammad-Javad Malek Hosseini; Spain: Begoña Gartzia De Bikuña, Ana I. Camacho; Cameroon/Stenasellidae: Raoul P. Tuekam Kayo; Algeria: Amina Taleb, Nouria Belaidi; Italy: Diana Maria Paola Galassi; Romania: Oana Teodora Moldovan.

AFFILIATIONS

¹Univ Lyon, Université Claude Bernard Lyon 1, CNRS, ENTPE, UMR 5023 LEHNA, Villeurbanne, France

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

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

⁴INRAE, UR-RiverLY, Centre Lyon-Grenoble Auvergne-Rhône-Alpes, Villeurbanne, France

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

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

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

⁸Department of Biology, SubBio Lab, Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia

⁹Faculty of Science, Department of Biology, University of Zagreb, Zagreb, Croatia

¹⁰Department of Invertebrate Zoology & Hydrobiology, Faculty of Biology & Environmental Protection, University of Lodz, Lodz, Poland

¹¹Musée National d'Histoire Naturelle de Luxembourg, Luxembourg City, Luxembourg

¹²Senckenberg Deutsches Entomologisches Institut, Müncheberg, Germany

¹³Departamento de Biologia Animal, and Centre for Ecology, Evolution and Environmental Changes & CHANGE – Global Change and Sustainability Institute, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal

¹⁴Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

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

¹⁶Hellenic Institute of Speleological Research, Crete, Greece

¹⁷Hypogean Crustacea Recording Scheme, Devon, UK

¹⁸CWEPSS, Commission Wallonne d'Etude et de Protection des Sites Souterrains, Bruxelles, Belgium

¹⁹SEPANSO Aquitaine, Bordeaux, France

²⁰Jovan Hadži Institute of Biology, Research Centre of the Slovenian Academy of Sciences and Arts (ZRC-SAZU), Ljubljana, Slovenia

²¹Department of Organisms and Ecosystems Research, National Institute of Biology (NIB), Ljubljana, Slovenia

²²Museo Nacional de Ciencias Naturales (CSIC). Dpto. Biodiversidad y Biología Evolutiva, Madrid, Spain

²³Anbiotek, Investigación científica y técnica del medio ambiente, Erandio, Bizkaia, Spain

²⁴Anbiolab, BIC Bizkaia Astondo bidea, Derio, Spain

²⁵Laboratoire d'Écologie et Gestion des Écosystèmes Naturels, University of Tlemcen, Tlemcen, Algeria

²⁶Faculty of Science, Department of Zoology, University of Bamenda, Bamenda, Cameroon

²⁷Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy

²⁸Emil Racovita Institute of Speleology, Cluj-Napoca, Romania

²⁹Centro Nacional de Investigación sobre la Evolución Humana, Burgos, Spain

³⁰Institut Universitaire de France, Paris, France

ACKNOWLEDGEMENTS

We thank all collectors, including the many speleologists and naturalists, who kindly provided specimens of asellids: their donation together with their names is gratefully acknowledged in the database. The World Asellidae database GOTIT is hosted by the CNRS/IN2P3 Computing Center (Villeurbanne, France). Sampling, taxonomic identification, DNA sequencing, database management, BS measurements and data analysis were supported financially by French National Research Agency projects CONVERGENOMICS (ANR-15-CE32-0005), EUR H2O'Lyon project (ANR-17-EURE-0018) and Biodiversa+ (Project DarCo), the European Biodiversity Partnership under the 2021-2022 BiodivProtect joint call for research proposals, co-funded by the European Commission (GA no. 101052342) (F.M., N.S., C.J.D., L.K.D., T.L., C.F., C.I.); the Slovenian Research Agency through the Research Core Programme Funding P1-0184 (T.D., B.S. and M.Z.); Spanish project PID2019-110243GB-I00 of MICINN/FEDER (A.I.C.); Romania Ministry of Research, Innovation and Digitization grant, CNCS/CCCDI-UEFISCDI, project 2/2019 (DARKFOOD), within PNCDI III (O.T.M.); the Polish Ministry of Science and Higher Education through projects no. N N303 579439 and 5818/B/P01/2010/39 (M.G.); the VILLUM FONDEN (research grant 15471) and Portuguese National Funds through 'Fundação para a Ciência e a Tecnologia' (FCT) within the cE3c Unit funding UIDB/00329/2020 (A.S.P.S.R.); Belgium exceptional grant ARES-CCD and French SCAC project (R.P.T.K.). We thank the associate editor and two anonymous reviewers for comments on an earlier draft of this manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

BENEFIT-SHARING STATEMENT

The work presented herein is from a collaborative group of researchers who is committed to international scientific partnerships, as well as institutional capacity building. Scientific collaborators are included as co-authors and the results of research are shared with the sample provider communities and the broader scientific community via the World Asellidae Database.

DATA AVAILABILITY STATEMENT

The sequence data generated as part of this study have been deposited in NCBI (accession numbers in Table S12). Sequence metadata, sequence alignments, the World Asellidae phylogeny and data used in comparative analyses are included as Supporting Information at the publisher's website or archived in Zenodo at <https://doi.org/10.5281/zenodo.8330212>. Access to metadata and data stored in WAD is available at <https://gotit.univ-lyon1.fr>. The web application GOTIT and structure of the World Asellidae Database are distributed with full documentation at <https://github.com/gotit-dev/gotit> under the terms of GNU General Public Licence. A demo version of GOTIT application is available at <https://gotit.cnrs.fr>.

ORCID

Nathanaele Saclier  <https://orcid.org/0000-0003-1522-9644>
 Louis Duchemin  <https://orcid.org/0000-0002-7984-9554>
 Lara Konecny-Dupré  <https://orcid.org/0009-0004-4604-687X>
 David Eme  <https://orcid.org/0000-0001-8790-0412>
 Cécile Callou  <https://orcid.org/0000-0002-8540-8114>
 Tristan Lefébure  <https://orcid.org/0000-0003-3923-8166>
 Clémentine François  <https://orcid.org/0000-0001-7781-8781>
 Fabio Stoch  <https://orcid.org/0000-0003-4535-3769>
 Boris Sket  <https://orcid.org/0000-0002-7398-5483>
 Sanja Gottstein  <https://orcid.org/0000-0002-1424-2911>
 Teo Delić  <https://orcid.org/0000-0003-4378-5269>
 Maja Zagmajster  <https://orcid.org/0000-0003-1323-9937>
 Michal Grabowski  <https://orcid.org/0000-0002-4551-3454>
 Dieter Weber  <https://orcid.org/0000-0001-7813-842X>
 Ana Sofia P. S. Reboleira  <https://orcid.org/0000-0002-4756-7034>
 Dmitry Palatov  <https://orcid.org/0000-0002-8826-9316>
 Kaloust Paragiamian  <https://orcid.org/0000-0001-7372-733X>
 Mohammad-Javad Malek Hosseini  <https://orcid.org/0000-0001-7411-2150>
 Ana I. Camacho  <https://orcid.org/0000-0003-0596-7678>
 Begoña Gartzia De Bikuña  <https://orcid.org/0000-0001-6089-9458>
 Amina Taleb  <https://orcid.org/0000-0003-4045-590X>
 Nouria Belaidi  <https://orcid.org/0009-0004-2780-8031>
 Raoul P. Tuekam Kayo  <https://orcid.org/0000-0002-2578-4898>
 Diana Maria Paola Galassi  <https://orcid.org/0000-0002-6448-2710>
 Oana Teodora Moldovan  <https://orcid.org/0000-0002-1262-0675>
 Christophe J. Douady  <https://orcid.org/0000-0002-4503-8040>
 Florian Malard  <https://orcid.org/0000-0001-8037-4464>

REFERENCES

- Adams, J., Greenwood, P., Pollitt, R., & Yonow, T. (1985). Loading constraints and sexual size dimorphism in *Asellus aquaticus*. *Behaviour*, 92(3/4), 277–287. <http://www.jstor.org/stable/4534415>
- Andersson, M. (1994). *Sexual selection*. Princeton University Press.
- Balart-García, P., Aristide, L., Bradford, T., Beasley-Hall, P., Polak, S., Ribera, I., Cooper, S., & Fernandez, R. (2023). Genomic exaptation and convergent evolution paved the way to independent subterranean colonization across beetle lineages. *Research Square*. Preprint under Review at Nature Portfolio. <https://doi.org/10.21203/rs.3.rs-2254102/v1>
- Balázs, G., Biró, A., Fišer, Ž., Fišer, C., & Herczeg, G. (2021). Parallel morphological evolution and habitat-dependent sexual dimorphism in cave- vs. surface populations of the *Asellus aquaticus* (crustacea: Isopoda: Asellidae) species complex. *Ecology and Evolution*, 11(21), 15389–15403. <https://doi.org/10.1002/ece3.8233>
- Bertin, A., & Cezilly, F. (2003). Sexual selection, antennae length and the mating advantage of large males in *Asellus aquaticus*. *Journal of Evolutionary Biology*, 16(4), 698–707. <https://doi.org/10.1046/j.1420-9101.2003.00565.x>
- Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K., & Das, L. (2007). Cryptic species as a window

- on diversity and conservation. *Trends in Ecology & Evolution*, 22(3), 148–155. <https://doi.org/10.1016/j.tree.2006.11.004>
- Biró, A., Balázs, G., Fišer, Ž., & Herczeg, G. (2022). Gender inequality in the dark: Are adaptations to the cave environment sex-specific? *Karstologia Memoires*, 21(1), 229–232.
- Blanckenhorn, W. U. (2000). The evolution of body size: What keeps organisms small? *The Quarterly Review of Biology*, 75(4), 385–407. <http://www.jstor.org/stable/2664968>
- Calvignac, S., Konecny, L., Malard, F., & Douady, C. J. (2011). Preventing the pollution of mitochondrial data sets with nuclear mitochondrial paralogs (numts). *Mitochondrion*, 11, 246–254. <https://doi.org/10.1016/j.mito.2010.10.004>
- Casquet, J., Thebaud, C., & Gillespie, R. G. (2012). Chelex without boiling, a rapid and easy technique to obtain stable amplifiable DNA from small amounts of ethanol stored spiders. *Molecular Ecology Resources*, 12, 136–141. <https://doi.org/10.1111/j.1755-0998.2011.03073.x>
- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17, 540–552. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>
- Chevenet, F., Dolédec, C., & Chessel, D. (1994). A fuzzy coding approach for the analysis of long-term ecological data. *Freshwater Biology*, 31, 295–309. <https://doi.org/10.1111/j.1365-2427.1994.tb01742.x>
- Christiansen, K. A. (1961). Convergence and parallelism in cave Entomobryinae. *Evolution*, 15, 288–301. <https://doi.org/10.2307/2406229>
- Cooper, S., Fišer, C., Zakšek, V., Delić, T., Borko, Š., Faille, A., & Humphreys, W. (2023). Phylogenies reveal speciation dynamics: Case studies from groundwater. In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater ecology and evolution* (pp. 165–183). Academic Press. <https://doi.org/10.1016/B978-0-12-819119-4.00020-2>
- Copilas-Ciocianu, D., Fišer, C., Borza, P., & Petrusek, A. (2018). Is subterranean lifestyle reversible? Independent and recent large-scale dispersal into surface waters by two species of the groundwater amphipod genus *Niphargus*. *Molecular Phylogenetics and Evolution*, 119, 37–49. <https://doi.org/10.1016/j.ympev.2017.10.023>
- Culver, D. C., Kowalko, J. E., & Pipan, T. (2023). Natural selection versus neutral mutation in the evolution of subterranean life: A false dichotomy? *Frontiers in Ecology and Evolution*, 11, 1080503. <https://doi.org/10.3389/fevo.2023.1080503>
- Culver, D. C., & Pipan, T. (2019). *The biology of caves and other subterranean habitats*. Oxford University Press.
- Degen, R., Aune, M., Bluhm, B. A., Cassidy, C., Kędra, M., Kraan, C., Vandepitte, L., Włodarska-Kowalczyk, M., Zhulay, I., Albano, P. G., Bremner, J., Grebmeier, J. M., Link, H., Morata, N., Nordström, M. C., Shojaei, M. G., Sutton, L., & Zuschin, M. (2018). Trait-based approaches in rapidly changing ecosystems: A roadmap to the future polar oceans. *Ecological Indicators*, 91, 722–736. <https://doi.org/10.1016/j.ecolind.2018.04.050>
- Eme, D., Malard, F., Konecny-Dupré, L., Lefébure, T., & Douady, C. J. (2013). Bayesian phylogeographic inferences reveal contrasting colonization dynamics among European groundwater isopods. *Molecular Ecology*, 22, 5685–5699. <https://doi.org/10.1111/mec.12520>
- Eme, D., Zagmajster, M., Delić, T., Fišer, C., Flot, J.-F., Konecny-Dupré, L., Pálsson, S., Stoch, F., Zakšek, V., Douady, C. J., & Malard, F. (2018). Do cryptic species matter in macroecology? Sequencing European groundwater crustaceans yields smaller ranges but does not challenge biodiversity determinants. *Ecography*, 41, 424–436. <https://doi.org/10.1111/ecog.02683>
- Faille, A. (2019). Beetles. In W. B. White, D. C. Culver, & T. Pipan (Eds.), *Encyclopedia of caves* (pp. 102–108). Academic Press. <https://doi.org/10.1016/B978-0-12-814124-3.00014-5>
- Fairbairn, D. J. (2007). Introduction: The enigma of sexual size dimorphism. In D. J. Fairbairn, W. U. Blanckenhorn, & T. Székely (Eds.), *Sex, size, and gender roles: Evolutionary studies of sexual size dimorphism* (pp. 1–10). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199208784.003.0001>
- Felsenstein, J. (1988). Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics*, 19, 445–471. <http://www.jstor.org/stable/2097162>
- Fišer, C. (2019a). Adaptation: morphological. In W. B. White, D. C. Culver, & T. Pipan (Eds.), *Encyclopedia of caves* (pp. 33–39). Academic Press. <https://doi.org/10.1016/B978-0-12-814124-3.00005-4>
- Fišer, C. (2019b). Collaborative databasing should be encouraged. *Trends in Ecology and Evolution*, 34, 184–185. <https://doi.org/10.1016/j.tree.2018.12.001>
- Fišer, C., Brancelj, A., Yoshizawa, M., Mammola, S., & Fišer, Ž. (2023). Dissolving morphological and behavioral traits of groundwater animals into a functional phenotype. In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater ecology and evolution* (pp. 415–438). Academic Press. <https://doi.org/10.1016/B978-0-12-819119-4.00012-3>
- Fišer, C., Robinson, C. T., & Malard, F. (2018). Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology*, 27, 613–635. <https://doi.org/10.1111/mec.14486>
- Fišer, C., Zagmajster, M., & Zakšek, V. (2013). Coevolution of life history traits and morphology in female subterranean amphipods. *Oikos*, 122, 770–778. <http://www.jstor.org/stable/41937725>
- Fišer, Ž., Prevorčnik, S., Lozej, N., & Trontelj, P. (2019). No need to hide in caves: Shelter-seeking behavior of surface and cave ecomorphs of *Asellus aquaticus* (Isopoda: Crustacea). *Zoology*, 134, 58–65. <https://doi.org/10.1016/j.zool.2019.03.001>
- Fong, D. W., & Carlini, D. B. (2023). Ecological and evolutionary perspectives on groundwater colonization by the amphipod crustacean *Gammarus minus*. In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater ecology and evolution* (pp. 373–392). Academic Press. <https://doi.org/10.1016/B978-0-12-819119-4.15007-3>
- Fontaneto, D., Flot, J.-F., & Tang, C. Q. (2015). Guidelines for DNA taxonomy, with a focus on the meiofauna. *Marine Biodiversity*, 45, 433–451. <https://doi.org/10.1007/s12526-015-0319-7>
- Friedrich, M. (2019). Adaptation to darkness. In W. B. White, D. C. Culver, & T. Pipan (Eds.), *Encyclopedia of caves* (pp. 16–23). Academic Press. <https://doi.org/10.1016/B978-0-12-814124-3.00003-0>
- Gouy, M., Guindon, S., & Gascuel, O. (2010). Seaview version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, 27, 221–224. <https://doi.org/10.1093/molbev/msp259>
- Gross, J. B., Boggs, T. E., Rétaux, S., & Torres-Paz, J. (2023). Developmental and genetic basis of troglomorphic traits in the teleost fish *Astyanax mexicanus*. In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater ecology and evolution* (pp. 351–371). Academic Press. <https://doi.org/10.1016/B978-0-12-819119-4.00004-4>
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of phylml 3.0. *Systematic Biology*, 59, 307–321. <https://doi.org/10.1093/sysbio/syq010>
- Henry, J.-P. (1976). Recherches sur les Asellidae hypogées de la lignée cavaticus (Crustacea, Isopoda, Asellota). PhD thesis, University of Dijon, Dijon, France.
- Henry, J.-P., & Magniez, G. (1993). Présence d'Asellides stygobies (Crustacea, Isopoda, Aselloidea) dans la région du Primorye, Sibérie Sud-Orientale. *Contributions to Zoology*, 62(3), 179–191. <https://doi.org/10.1163/26660644-06203003>
- Henry, J.-P., & Magniez, G. (1995). Nouvelles données sur les Asellidae épigés d'Extrême-Orient (Crustacea, Isopoda, Asellota). *Contributions to Zoology*, 65(2), 101–122. <https://doi.org/10.1163/26660644-06502003>
- Hidding, B., Michel, E., Natyaganova, A. V., & Sherbakov, D. Y. (2003). Molecular evidence reveals a polyphyletic origin and chromosomal

- speciation of Lake Baikal's endemic asellid isopods. *Molecular Ecology*, 12(6), 1509–1514. <https://doi.org/10.1046/j.1365-294X.2003.01821.x>
- Hobern, D., Apostolico, A., Arnaud, E., Bello, J. C., Canhos, D., Dubois, G., Field, D., Alonso Garcia, E., Hardisty, A., Harrison, J., Heidorn, B., Krishtalka, L., Mata, E., Page, R. D. M., Parr, C., Price, J., & Willoughby, S. (2012). Global biodiversity informatics outlook: Delivering biodiversity knowledge in the information age. *Global Biodiversity Information Facility*, 1–41. <https://doi.org/10.15468/6jxa-yb44>
- Hose, G., Chariton, A., Daam, M., Di Lorenzo, T., Galassi, D., Halse, S., Reboleira, A. S. P. S., Robertson, A., Schmidt, S., & Korb, K. (2022). Invertebrate traits, diversity and the vulnerability of groundwater ecosystems. *Functional Ecology*, 36(9), 2200–2214. <https://doi.org/10.1111/1365-2435.14125>
- Humphreys, W. F. (2000). Relict faunas and their derivation. In H. Wilkens, D. C. Culver, & W. F. Humphreys (Eds.), *Subterranean ecosystems* (pp. 417–432). Elsevier.
- Jormalainen, V. (2007). Mating strategies in isopods: From mate monopolization to conflicts. In J. E. Duffy & M. Thiel (Eds.), *Evolutionary ecology of social and sexual systems: Crustaceans as model organisms* (pp. 167–190). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195179927.003.0008>
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kelly, C. D., Bussière, L. F., & Gwynne, D. T. (2008). Sexual selection for male mobility in a giant insect with female-biased size dimorphism. *The American Naturalist*, 172(3), 417–423. <https://doi.org/10.1086/589894>
- Konec, M., Prevorčnik, S., Sarbu, S. M., Verovnik, R., & Trontelj, P. (2015). Parallels between two geographically and ecologically disparate cave invasions by the same species, *Asellus aquaticus* (isopoda, crustacea). *Journal of Evolutionary Biology*, 28(4), 864–875. <https://doi.org/10.1111/jeb.12610>
- Kostanjsek, R., Zaksek, V., Bizjak-Mali, L., & Trontelj, P. (2023). The olm (*Proteus anguinus*), a flagship groundwater species. In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater ecology and evolution* (pp. 305–327). Academic Press. <https://doi.org/10.1016/B978-0-12-819119-4.15006-1>
- Kowalko, J., Dickinson, M. H., Voshall, L. B., & Dow, J. A. T. (2020). Utilizing the blind cavefish *Astyanax mexicanus* to understand the genetic basis of behavioral evolution. *Journal of Experimental Biology*, 223(Suppl. 1), jeb208835. <https://doi.org/10.1242/jeb.208835>
- Kunz, S., Kefford, B. J., Schmidt-Kloiber, A., Matthaei, C. D., Usseglio-Polatera, P., Graf, W., Poff, N. L., Metzeling, L., Twardochleb, L., Hawkins, C. P., & Schäfer, R. B. (2022). Tackling inconsistencies among freshwater invertebrate trait databases: Harmonising across continents and aggregating taxonomic resolution. *Freshwater Biology*, 67, 275–291. <https://doi.org/10.1111/fwb.13840>
- Langille, B. L., Tierney, S. M., Bertozzi, T., Beasley-Hall, P. G., Bradford, T. M., Fagan-Jeffries, E. P., Hyde, J., Leijes, R., Richardson, M., Saint, K. M., Stringer, D. N., Villastrigo, A., Humphreys, W. F., Austin, A. D., & Cooper, S. J. B. (2022). Parallel decay of vision genes in subterranean water beetles. *Molecular Phylogenetics and Evolution*, 173, 107522. <https://doi.org/10.1016/j.ympev.2022.107522>
- Lartillot, N., Lepage, T., & Blanquart, S. (2009). PhyloBayes 3: A Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*, 25(17), 2286–2288. <https://doi.org/10.1093/bioinformatics/btp368>
- Ledford, J., Paquin, P., Cokendolpher, J., Campbell, J., & Griswold, C. (2011). Systematics of the spider genus *Neoleptoneta* Brignoli, 1972 (Araneae: Leptonetidae) with a discussion of the morphology and relationships for the north American Leptonetidae. *Invertebrate Systematics*, 25, 334–388. <https://www.publish.csiro.au/is/IS11014>
- Lefebure, T., Douady, C. J., Gouy, M., & Gibert, J. (2006). Relationship between morphological taxonomy and molecular divergence within crustacea: Proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution*, 40, 435–447. <https://doi.org/10.1016/j.ympev.2006.03.014>
- Lefebure, T., Morvan, C., Malard, F., François, C., Konecny-Dupré, L., Guéguen, L., Weiss-Gayet, M., Seguin-Orlando, A., Ermini, L., Der Sarkissian, C., Charrier, N. P., Eme, D., Mermillod-Blondin, F., Duret, L., Vieira, C., Orlando, L., & Douady, C. J. (2017). Less effective selection leads to larger genomes. *Genome Research*, 27, 1016–1028. <https://doi.org/10.1101/gr.212589.116>
- Lepage, T., Bryant, D., Philippe, H., & Lartillot, N. (2007). A general comparison of relaxed molecular clock models. *Molecular Biology and Evolution*, 24(12), 2669–2680. <https://doi.org/10.1093/molbev/msm193>
- Lewis, J. J., Lewis, S. L., Orndorff, W., Orndorff, Z., Malard, F., Konecny-Dupré, L., Saclier, N., & Douady, C. J. (2023). *The groundwater isopods of Virginia (Isopoda: Asellidae and Cirolanidae)*. Virginia Museum of Natural History. Special Publication 19, in press.
- Lovich, J. E., & Gibbons, J. W. (1992). A review of techniques for quantifying sexual size dimorphism. *Growth Development and Aging*, 56, 269–281.
- Löytynoja, A., & Goldman, N. (2008). Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science*, 320(5883), 1632–1635. <https://doi.org/10.1126/science.1158395>
- Lukić, M. (2019). Collembola. In W. B. White, D. C. Culver, & T. Pipan (Eds.), *Encyclopedia of caves* (pp. 308–319). Academic Press. <https://doi.org/10.1016/B978-0-12-814124-3.00034-0>
- Maechler, M., Rousseeuw, P., Struyf, A., & Hubert, M. (2022). Package 'cluster': Cluster analysis extended. R Foundation for Statistical Computing, Vienna. <https://svn.r-project.org/R-packages/trunk/cluster/>
- Malard, F. (2022). Groundwater metazoans. In T. Mehner & K. Tockner (Eds.), *Encyclopedia of inland waters* (2nd ed., pp. 474–487). Elsevier. <https://doi.org/10.1016/B978-0-12-819166-8.00176-6>
- Malard, F., Grison, P., Duchemin, L., Konecny-Dupré, L., Lefebure, T., Saclier, N., Eme, D., Martin, C., Callou, C., & Douady, C. J. (2020). GOTIT: A laboratory application software for optimizing multi-criteria species-based research. *Methods in Ecology and Evolution*, 11, 159–167. <https://doi.org/10.1111/2041-210X.13307>
- Malard, F., Machado, E. G., Casane, D., Cooper, S., Fišer, C., & Eme, D. (2023). Dispersal and geographic range size in groundwater. In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater ecology and evolution* (pp. 185–207). Academic Press. <https://doi.org/10.1016/B978-0-12-819119-4.15003-6>
- Mammola, S., Amorim, I. R., Bichuette, M. E., Borges, P. A. V., Cheeptham, N., Cooper, S. J. B., Culver, D. C., Deharveng, L., Eme, D., Ferreira, R. L., Fišer, C., Fišer, Ž., Fong, D. W., Griebler, C., Jeffery, W. R., Jugovic, J., Kowalko, J. E., Lilley, T. M., Malard, F., ... Cardoso, P. (2020). Fundamental research questions in subterranean biology. *Biological Reviews*, 95, 1855–1872. <https://doi.org/10.1111/brv.12642>
- Mammola, S., Cardoso, P., Angyal, D., Balázs, G., Blick, T., Brustel, H., Carter, J., Čurčić, S., Danflous, S., Dányi, L., Déjean, S., Deltchev, C., Elverici, M., Fernández, J., Gasparo, F., Komnenov, M., Komposch, C., Kováč, L., Kunt, K. B., ... Isaia, M. (2019). Continental data on cave-dwelling spider communities across Europe (Arachnida: Araneae). *Biodiversity Data Journal*, 7, e38492. <https://doi.org/10.3897/BDJ.7.e38492>
- Mammola, S., & Isaia, M. (2017). Spiders in caves. *Proceedings of the Royal Society B: Biological Sciences*, 284, 20170193. <https://doi.org/10.1098/rspb.2017.0193>

- Mammola, S., & Martínez, A. (2020). Let research on subterranean habitats resonate! *Subterranean Biology*, 36, 63–71. <https://doi.org/10.3897/subtbiol.36.59960>
- Mammola, S., Pavlek, M., Huber, B. A., Isaia, M., Ballarin, F., Tolve, M., Čupić, I., Hesselberg, T., Lunghi, E., Mouron, S., Graco-Roza, C., & Cardoso, P. (2022). A trait database and updated checklist for European subterranean spiders. *Scientific Data*, 9(1), 236. <https://doi.org/10.1038/s41597-022-01316-3>
- Martins, E. P., & Hansen, T. F. (1997). Phylogenies and the comparative method: A general approach to incorporating phylogenetic information into the analysis of interspecific data. *The American Naturalist*, 149(4), 646–667. <https://doi.org/10.1086/286013>
- Matsumoto, K. (1960). Subterranean isopods of the Kyushu District, with the descriptions of three new species. *Bulletin of the Biogeographical Society of Japan*, 22(3), 27–44.
- Matsumoto, K. (1961). Two subterranean isopods from the Amami group (Ryukyu Islands), with a description of a new species. *Annotationes Zoologicae Japonenses*, 34(4), 208–215.
- Matsumoto, K. (1963). Studies on the subterranean isopoda of Japan with notes on the well-water fauna of Japan. In *Part I: Studies on the subterranean isopoda of Japan (No. 1). Supplement of the annual report* (pp. 1–77). XIII Tokyo-to Laboratories for Medical Sciences.
- Miller, M. A. (1933). A new blind isopod, *Asellus californicus*, and a revision of the subterranean asellids. *University of California Publications in Zoology*, 39(4), 97–110.
- Morvan, C., Malard, F., Paradis, E., Lefébure, T., Konecny-Dupré, L., & Douady, C. J. (2013). Timetree of Aselloidea reveals species diversification dynamics in groundwater. *Systematic Biology*, 62(4), 512–522. <https://doi.org/10.1093/sysbio/syt015>
- Nelson, E. K., Piehler, B., Eckels, J., Rauch, A., Bellew, M., Hussey, P., Ramsay, S., Nathe, C., Lum, K., Krouse, K., Stearns, D., Connolly, B., Skillman, T., & Igra, M. (2011). LabKey server: An open source platform for scientific data integration, analysis and collaboration. *BMC Bioinformatics*, 12, 71. <https://doi.org/10.1186/1471-2105-12-71>
- Niemiller, M. L., Fitzpatrick, B. M., Shah, P., Schmitz, L., & Near, T. J. (2013). Evidence for repeated loss of selective constraint in rhodopsin of amblyopsid cavefishes (Teleostei: Amblyopsidae). *Evolution*, 67(3), 732–748. <https://doi.org/10.1111/j.1558-5646.2012.01822.x>
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20, 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- Pfenninger, M., & Schwenk, K. (2007). Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evolutionary Biology*, 7, 121. <https://doi.org/10.1186/1471-2148-7-121>
- Pincheira-Donoso, D., & Hunt, J. (2017). Fecundity selection theory: Concepts and evidence. *Biological Reviews*, 92(1), 341–356. <https://doi.org/10.1111/brv.12232>
- Pinheiro, J., Bates, D., & R Core Team. (2022). nlme: Linear and nonlinear mixed effects models. R Package Version 3.1-160. <https://CRAN.R-project.org/package=nlme>
- Pipán, T., & Culver, D. C. (2012). Convergence and divergence in the subterranean realm: A reassessment. *Biological Journal of the Linnean Society*, 107, 1–14. <https://doi.org/10.1111/j.1095-8312.2012.01964.x>
- Pipán, T., & Culver, D. C. (2017). The unity and diversity of the subterranean realm with respect to invertebrate body size. *Journal of Cave and Karst Studies*, 79(1), 1–9. https://digitalcommons.usf.edu/kip_articles/5638
- Policarpo, M., Fumey, J., Lafargeas, P., Naquin, D., Thermes, C., Naville, M., Dechaud, C., Volff, J. N., Cabau, C., Klopp, C., Möller, P. R., Bernatchez, L., García-Machado, E., Rétaux, S., & Casane, D. (2021). Contrasting gene decay in subterranean vertebrates: Insights from cavefishes and fossorial mammals. *Molecular Biology and Evolution*, 38(2), 589–605. <https://doi.org/10.1093/molbev/msaa249>
- Poulson, T. L., & White, W. B. (1969). The cave environment. *Science*, 165, 971–981. <https://www.science.org/doi.org/10.1126/science.165.3897.971>
- Prendini, L., Francke, O. F., & Vignoli, V. (2010). Troglomorphism, trichobothriotaxy and typhlochactid phylogeny (Scorpiones, Chactioidea): More evidence that troglotism is not an evolutionary dead-end. *Cladistics*, 26, 117–142. <https://doi.org/10.1111/j.1096-0031.2009.00277.x>
- Prevorčnik, S., Blejec, A., & Sket, B. (2004). Racial differentiation in *Asellus aquaticus* (L.) (crustacea isopoda Asellidae). *Archiv für Hydrobiologie*, 160(2), 193–214. <https://doi.org/10.1127/0003-9136/2004/0160-0193>
- Protas, M., & Jeffery, W. R. (2012). Evolution and development in cave animals: From fish to crustaceans. *Wiley Interdisciplinary Reviews: Developmental Biology*, 1, 823–845. <https://doi.org/10.1002/wdev.61>
- Protas, M., Trontelj, P., Prevorčnik, S., & Fišer, Ž. (2023). The *Asellus aquaticus* species complex: An invertebrate model in subterranean evolution. In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater ecology and evolution* (pp. 329–350). Academic Press. <https://doi.org/10.1016/B978-0-12-819119-4.00016-0>
- Rétaux, S., & Jeffery, W. R. (2023). Voices from the underground: Animal models for the study of trait evolution during groundwater colonization and adaptation. In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater ecology and evolution* (pp. 285–304). Academic Press. <https://doi.org/10.1016/B978-0-12-819119-4.00002-0>
- Ridley, M., & Thompson, D. J. (1979). Size and mating in *Asellus aquaticus* (crustacea: Isopoda). *Zeitschrift für Tierpsychologie*, 51, 380–397. <https://doi.org/10.1111/j.1439-0310.1979.tb00697.x>
- Saccò, M., Guzik, M. T., van der Heyde, M., Nevill, P., Cooper, S. J. B., Austin, A. D., Coates, P. J., Allentoft, M. E., & White, N. E. (2022). eDNA in subterranean ecosystems: Applications, technical aspects, and future prospects. *Science of the Total Environment*, 820, 153223. <https://doi.org/10.1016/j.scitotenv.2022.153223>
- Saclier, N. (2019). Origine des variations de taux d'évolution moléculaire inter-spécifiques: apport d'un modèle génomique en milieu souterrain. PhD thesis, Université Claude Bernard Lyon 1, Lyon, France.
- Saclier, N., François, C. M., Konecny-Dupré, L., Lartillot, N., Guéguen, L., Duret, L., Malard, F., Douady, C. J., & Lefébure, T. (2018). Life history traits impact the nuclear rate of substitution but not the mitochondrial rate in isopods. *Molecular Biology and Evolution*, 35, 2900–2912. <https://doi.org/10.1093/molbev/msy184>
- Sarremejane, R., Cid, N., Stubbington, R., Datry, T., Alp, M., Cañedo-Argüelles, M., Cordero-Rivera, A., Csabai, Z., Gutiérrez-Cánovas, C., Heino, J., Forcellini, M., Millán, A., Paillex, A., Paril, P., Polásek, M., de Figueroa, J. M. T., Usseglio-Polatera, P., Zamora-Muñoz, C., & Bonada, N. (2020). DISPERSE, a trait database to assess the dispersal potential of European aquatic macroinvertebrates. *Scientific Data*, 7, 386. <https://doi.org/10.1038/s41597-020-00732-7>
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH image to ImageJ: 25 years of image analysis. *Nature Methods*, 9, 671–675. <https://doi.org/10.1038/nmeth.2089>
- Shuster, S. M., Briggs, W. R., & Dennis, P. A. (2013). How multiple mating by females affects sexual selection. *Philosophical Transactions of the Royal Society B*, 368, 20120046. <https://doi.org/10.1098/rstb.2012.0046>
- Shuster, S. M., & Wade, M. J. (2003). *Mating systems and strategies*. Princeton University Press.
- Steel, E. A. (1961). Some observations on the life history of *Asellus aquaticus* (L.) and *Asellus meridianus* Racovitza (crustacea: Isopoda). *Proceedings of the Zoological Society of London*, 137, 71–87. <https://doi.org/10.1111/j.1469-7998.1961.tb06162.x>
- Stern, D. B., Breinholt, J., Pedraza-Lara, C., López-Mejía, M., Owen, C. L., Bracken-Grissom, H., Fetzner, J. W., Jr., & Crandall, K. A. (2017). Phylogenetic evidence from freshwater crayfishes that cave adaptation is not an evolutionary dead-end. *Evolution*, 71(10), 2522–2532. <https://doi.org/10.1111/evo.13326>

- Stern, D. B., & Crandall, K. A. (2018). Phototransduction gene expression and evolution in cave and surface crayfishes. *Integrative and Comparative Biology*, 58(3), 398–410. <https://doi.org/10.1093/icb/icy029>
- Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences*, 17, 57–86.
- Thioulouse, J., Dray, S., Dufour, A.-B., Siberchicot, A., Jombart, T., & Pavoine, S. (2018). *Multivariate analysis of ecological data with ade4*. Springer. <https://doi.org/10.1007/978-1-4939-8850-1>
- Trontelj, P., Blejec, A., & Fišer, C. (2012). Ecomorphological convergence of cave communities. *Evolution*, 66(12), 3852–3865. <https://doi.org/10.1111/j.1558-5646.2012.01734.x>
- Venarsky, M., Niemiller, M. L., Fišer, C., Saclier, N., & Moldovan, O. T. (2023). Life histories in groundwater organisms. In F. Malard, C. Griebler, & S. Rétaux (Eds.), *Groundwater ecology and evolution* (pp. 439–456). Academic Press. <https://doi.org/10.1016/B978-0-12-819119-4.00013-5>
- Verdier, H., Konecny-Dupré, L., Marquette, C., Reveron, H., Tadier, S., Grémillard, L., Barthès, A., Datry, T., Bouchez, A., & Lefébure, T. (2022). Passive sampling of environmental DNA in aquatic environments using 3D-printed hydroxyapatite samplers. *Molecular Ecology Resources*, 22(6), 2158–2170. <https://doi.org/10.1111/1755-0998.13604>
- Verovnik, R., Sket, B., & Trontelj, P. (2005). The colonization of Europe by the freshwater crustacean *Asellus aquaticus* (crustacea: Isopoda) proceeded from ancient refugia and was directed by habitat connectivity. *Molecular Ecology*, 14(14), 4355–4369. <https://doi.org/10.1111/j.1365-294X.2005.02745.x>
- Wilkens, H., & Strecker, U. (2017). Evolution in the dark. In *Darwin's loss without selection*. Springer. <https://doi.org/10.1007/978-3-662-54512-6>
- Zagmajster, M., Borko, Š., Delić, T., Douady, C. J., Eme, D., Malard, F., Trontelj, P., & Fišer, C. (2022). Availability of DNA barcodes in subterranean amphipods of Europe. *Karstologia Memoires*, 21(1), 361–364.
- Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, 29, 2869–2876. <https://doi.org/10.1093/bioinformatics/btt499>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Saclier, N., Duchemin, L., Konecny-Dupré, L., Grison, P., Eme, D., Martin, C., Callou, C., Lefébure, T., François, C., Issartel, C., Lewis, J. J., Stoch, F., Sket, B., Gottstein, S., Delić, T., Zagmajster, M., Grabowski, M., Weber, D., Reboleira, A. S. P. S. ... Malard, F. (2023). A collaborative backbone resource for comparative studies of subterranean evolution: The World Asellidae database. *Molecular Ecology Resources*, 00, 1–19. <https://doi.org/10.1111/1755-0998.13882>