Manuscript Details

Manuscript number	MPE_2019_406_R2
Title	Phylogeny and biogeography of the Cavernicola (Platyhelminthes: Tricladida): relicts of an epigean group sheltering in caves?
Article type	Research Paper

Abstract

The planarian suborder Cavernicola Sluys, 1990 was originally created to house five species of triclad flatworms with special morphological features and a surprisingly discontinuous and broad geographic distribution. These five species could not be accommodated with any degree of certainty in any of the three taxonomic groups existing at that moment, viz., Paludicola Hallez, 1892, Terricola Hallez, 1892, and Maricola Hallez, 1892. The scarce representation of the group and the peculiarities of the morphological features of the species, including several described more recently, have complicated new tests of the monophyly of the Cavernicola, the assessment of its taxonomic status, as well as the resolution of its internal relationships. Here we present the first molecular study including all genera currently known for the group, excepting one. We analysed newly generated 18S and 28S rDNA data for these species, together with a broad representation of other triclad flatworms. The resulting phylogenetic trees supported the monophyly of the Cavernicola, as well as its sister-group relationship to the Maricola. The sister-group relationship to the Maricola and affinities within the Cavernicola falsify the morphology-based phylogeny of the latter that was proposed previously. The relatively high diversity of some cavernicolan genera suggests that the presumed rarity of the group actually may in part be due to a collecting artefact. Ancestral state reconstruction analyses suggest that the ancestral habitat of the group concerned epigean freshwater conditions. Our results point to an evolutionary scenario in which the Cavernicola (a) originated in a freshwater habitat (b) as the sister clade of the marine triclads, and (c) subsequently radiated and colonized both epigean and hypogean environments. Competition with other planarians, notably members of the Continenticola, or changes in epigean habitat conditions are two possible explanations -still to be tested- for the loss of most epigean diversity of the Cavernicola, which is currently reflected in their highly disjunct distributions.

Keywords	Gondwana, Pangea, molecular phylogeny, taxonomy, troglobitic flatworms			
Taxonomy	Zoology, Evolutionary Biology			
Corresponding Author	Marta Riutort			
Corresponding Author's Institution	Universitat de Barcelona			
Order of Authors	LISANDRA BENÍTEZ ALVAREZ, Ana Leal-Zanchet, Alejandro Oceguera Figueroa, Rodrigo Ferreira, Diego de Medeiros Bento, João Alberto Leão Braccini, Ronald Sluys, Marta Riutort			
Suggested reviewers	Christopher Laumer, Fernando Carbayo, Masaharu Kawakatsu			

Submission Files Included in this PDF

File Name [File Type] Benitez_et_al_Cover-letter_051219.pdf [Cover Letter] Responses to the reviewers.docx [Response to Reviewers] Highligths.docx [Highlights] Graphical_abstract.pdf [Graphical Abstract] Benitez&al revised 051219-track-changes.docx [Manuscript File] Figure_1_5nov19.pdf [Figure] Figure 2.pdf [Figure] Figure_3.pdf [Figure] Figure 4 3dec19.pdf [Figure] Figure_5_new.pdf [Figure] AuthorsContributionStatement.docx [Author Statement] Appendix_A_5nov19.pdf [Supplementary Material] Appendix_B_5nov19.pdf [Supplementary Material] Appendix_C_5nov19.pdf [Supplementary Material] Appendix_D_3dic19.docx [Supplementary Material] Appendix_E_5nov19.pdf [Supplementary Material]

Submission Files Not Included in this PDF

File Name [File Type]

Supplementary-tables_S1.xlsx [Supplementary Material]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Highligths

- First molecular phylogeny of a rare (11 species) planarian suborder, Cavernicola
- confirms monophyly, Maricola sistership and their status as a distinct suborder
- freshwater and epigean in origin, colonized both epigean and hypogean habitats
- disjunct distribution may be due to Gondwana origin, and loss of most epigean diversity
- genetic diversity and new findings foreshadow more species to be discovered



1 hylogeny and biogeography of the Cavernicola (Platyhelminthes:

2 Tricladida): relicts of an epigean group sheltering in caves?

3

4	Lisandra Benítez-Alvarez ¹ , Ana Maria Leal-Zanchet ² , Alejandro Oceguera-Figueroa ³ ,
5	Rodrigo Lopes Ferreira ⁴ , Diego de Medeiros Bento ⁵ , João Braccini ² , Ronald Sluys ⁶ , Marta
6	Riutort ^{1, a}
7	
8 9	¹ Departament de Genètica, Microbiologia i Estadística and Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Catalonia, Spain.
10 11	² Instituto de Pesquisas de Planárias and Programa de Pós-Graduação em Biologia, Universidade do Vale do Rio dos Sinos (UNISINOS), 93022-750 São Leopoldo, Rio Grande do Sul, Brazil.
12 13	³ Laboratorio de Helmintología. Instituto de Biología, Universidad Nacional Autónoma de México. Tercer Circuito s/n, Ciudad Universitaria, Copilco, Coyoacán. C.P. 04510. A.P. 70–153. Ciudad de México, México.
14 15	⁴ Centro de Estudos em Biologia Subterrânea, Setor de Biodiversidade Subterrânea, Departamento de Biologia, Universidade Federal de Lavras, Campus Universitário, Caixa Postal 3037, Lavras, Minas Gerais, Brazil.
16 17	⁵ ICMBio/CECAV-RN - Centro Nacional de Pesquisa e Conservação de Cavernas, 59015-350 Natal, Rio Grande do Norte, Brazil.
18	⁶ Naturalis Biodiversity Center, P.O. Box 9517, 2300 RA Leiden, The Netherlands.
19	
20	^a Corresponding author: Marta Riutort, mriutort@ub.edu
21	
22	Keywords: Gondwana, Pangea, molecular phylogeny, taxonomy, troglobitic flatworms

24 Abstract

The planarian suborder Cavernicola Sluys, 1990 was originally created to house five species 25 of triclad flatworms with special morphological features and a surprisingly discontinuous and 26 broad geographic distribution. These five species could not be accommodated with any degree 27 of certainty in any of the three taxonomic groups existing at that moment, viz., Paludicola 28 Hallez, 1892, Terricola Hallez, 1892, and Maricola Hallez, 1892. The scarce representation of 29 the group and the peculiarities of the morphological features of the species, including several 30 described more recently, have complicated new tests of the monophyly of the Cavernicola, the 31 assessment of its taxonomic status, as well as the resolution of its internal relationships. Here 32 we present the first molecular study including all genera currently known for the group, 33 excepting one. We analysed newly generated 18S and 28S rDNA data for these species, 34 together with a broad representation of other triclad flatworms. The resulting phylogenetic 35 36 trees supported the monophyly of the Cavernicola, as well as its sister-group relationship to the Maricola. The sister-group relationship to the Maricola and affinities within the 37 38 Cavernicola falsify the morphology-based phylogeny of the latter that was proposed 39 previously. The relatively high diversity of some cavernicolan genera suggests that the presumed rarity of the group actually may in part be due to a collecting artefact. Ancestral 40 state reconstruction analyses suggest that the ancestral habitat of the group concerned epigean 41 freshwater conditions. Our results point to an evolutionary scenario in which the Cavernicola 42 (a) originated in a freshwater habitat (b) as the sister clade of the marine triclads, and (c) 43 subsequently radiated and colonized both epigean and hypogean environments. Competition 44 with other planarians, notably members of the Continenticola, or changes in epigean habitat 45 conditions are two possible explanations -still to be tested- for the loss of most epigean 46 diversity of the Cavernicola, which is currently reflected in their highly disjunct distributions. 47

48 **1. INTRODUCTION**

Between 1946 and 1983 five species of planarian flatworms (Platyhelminthes, 49 Tricladida) had been described that consistently defied the taxonomic schemes developed by 50 planarian systematists. Four out of these five species (*Opisthobursa mexicana* Benazzi, 1972; 51 52 O. josephinae Benazzi 1975; Balliania thetisae Gourbault, 1978; Novomitchellia sarawakana (Kawakatsu & Chapman, 1983) usually had been assigned to the marine triclads of the 53 Suborder Maricola Hallez, 1892. The fifth species, Rhodax evelinae Marcus, 1946, was 54 55 considered to belong to the freshwater triclads or Paludicola Hallez, 1892. It should be noted that the Suborder Paludicola is no longer valid; its representatives, together with terrestrial 56 planarians -the now obsolete Suborder Terricola Hallez, 1892- are currently classified in the 57 Suborder Continenticola Carranza et al., 1998 (Sluys et al. 2009; Riutort et al. 2012). In all 58 cases, however, some doubt was expressed about the taxonomic assignments of these five 59 species (Ball 1974, Sluys 1990). At long last, Sluys (1990) resolved the taxonomic confusion 60 surrounding these five species by showing, on the basis of morphological characters, that they 61 formed a monophyletic group that represented a new and different clade in the phylogenetic 62 63 tree of the triclad flatworms. At that time three major clades, at the level of suborder or 64 infraorder, were recognized within the Tricladida Lang, 1884, viz., Paludicola, Maricola, and Terricola. For his new, fourth branch on the tree of the planarian flatworms Sluys (1990) 65 erected a new taxon for which he coined the name Cavernicola Sluys, 1990, presently being 66 ranked as a Suborder (Sluys et al. 2009). Although most of its constituent species had a 67 hypogean habitat and exhibited adaptations to life in caves (unpigmented body, lack of eyes), 68 Sluys (1990) stressed the notion that the name of the new taxon had no ecological 69 connotation. 70

With respect to the phylogenetic position of the new suborder within the Tricladida,
Sluys (1990) suggested a possible close relationship between the Cavernicola and the

Paludicola, based on the fact that the cavernicolan *Opisthobursa josephinae* exhibits one of the three presumed autapomorphies of the Paludicola, viz., sperm transfer by means of a spermatophore. However, he considered that character distribution as too weak to formally propose presence of a spermatophore as a synapomorphy for the Cavernicola and the Paludicola. Relationships within the Cavernicola were analysed also by Sluys (1990). The fact that the species possess a mixture of primitive features (Marcus, 1946, Sluys 1990) greatly complicated resolution of their phylogenetic affinities.

80 After this, it took a long time before the number of species for the Cavernicola started to increase slowly. Two new species and one new genus were described in recent years, viz., 81 Hausera hauseri Leal-Zanchet & Souza, 2014 from Brazil, and Novomitchellia bursaelongata 82 Harrath, Sluys & Riutort, 2016 from Africa; both species live in a hypogean habitat (Leal-83 Zanchet et al., 2014; Harrath et al., 2016). In addition, Laumer and Giribet (2014) reported 84 18S and 28S rRNA sequences for a new, undescribed species of Cavernicola. It was only 85 recently that this new species acquired its proper taxonomic designation when it was 86 described as the new genus and species Kawakatsua pumila Sluys, 2019 (Sluys and Laumer, 87 88 2019). It is noteworthy that this species was found in a basically terrestrial habitat. Addition 89 of these new species to the Cavernicola made even more evident a conspicuous feature of this small group of species, i.e., their highly disjunct distributions (Fig. 1). 90

The present study is the first to include molecular data for all cavernicolan taxa, excepting *Balliania* Gourbault, 1978. In our analyses we have incorporated also representatives of 15 genera of triclads belonging to the other two suborders, thus allowing us to test for the first time the previously hypothesized monophyly of the Cavernicola, to analyse its relationships within the Tricladida, as well as the affinities between its constituent taxa.

96

2. MATERIALS AND METHODS

97 2.1 Taxon sampling and identification

98 We obtained samples from six either new or already known localities from South and North America (Southern Mexico), and combined our data with sequences obtained from 99 100 GenBank, thus including all genera of the Cavernicola presently known, excepting Balliania 101 (Table 1, Fig. 1). New specimens of Opisthobursa mexicana and Hausera hauseri were sampled at the original type localities of these two species, viz., Las Grutas de Coconá, 102 Tabasco, Mexico and Crotes cave, Rio Grande do Norte, Brazil, respectively. In the case of 103 104 *Rhodax*, the type-locality of *Rhodax evelinae* (the only described species for the genus) no longer exists as it was dramatically transformed due to urbanization, hence representatives of 105 this genus in our study come from other localities. In first instance, we assigned these new 106 representatives to the genus *Rhodax* on the basis of their external features, combined with 107 anatomical and histological features. All new specimens present the following characteristics 108 of the genus *Rhodax*: rounded anterior tip with an adhesive organ; eyes present; pigmented 109 body; short, cylindrical pharynx (Appendix A). Specimens of *Rhodax* spp. 1 and 2, sampled 110 in surface waters located in Tavares and Pinheirinhos, respectively, in southern Brazil, did not 111 have reproductive organs. Specimens of *Rhodax* sp. 2 showed asexual reproduction in the 112 113 laboratory, similarly to what was described by Marcus (1946) in the original description of the species. Specimens of *Rhodax* sp. 3 from surface water in Tramandaí, southern Brazil, 114 presented a reproductive system, which is characterized by the presence of testes tubes, a 115 large common spermiducal vesicle, and a connection between the copulatory apparatus and 116 the intestine (Leal-Zanchet et al., unpublished results). With respect to their female copulatory 117 organs, these animals show some differences with R. evelinae, such as a longer female genital 118 duct, which may be due to intraspecific variation. Hausera sp., which was sampled in a cave 119 from northeastern Brazil, is a typical troglobitic animal with an unpigmented body and 120 121 absence of eyes (Appendix A), similar to H. hauseri. Hausera sp. also matches other

diagnostic features of the genus, such as sperm ducts separately penetrating the penis bulb, the
female genital duct communicating with the intestine, ovovitelline ducts without caudal
dichotomy, uniting to form a common ovovitelline duct. *Hausera* sp. differs from *H. hauseri*in the shape of the penis papilla and bulbar cavity, the course of the sperm ducts when
approaching the penis bulb, and the shape and length of the female genital duct (Hellmann et
al., unpublished results).

In order to determine the phylogenetic position of the Cavernicola within Tricladida, we 128 included in our analyses one representative sequence of each genus of the Cavernicola and 129 also sequences of representative taxa of the suborders Maricola and Continenticola. We also 130 included as outgroup species belonging to groups most closely related to the Tricladida, 131 according to previous studies (Laumer et al., 2015; Norén and Jondelius, 2002; Riutort et al., 132 2012), viz., Fecampiidae and Prolecitophora (Table 2). For determining the relationships 133 within the suborder Cavernicola we used as ingroup all available sequences assigned to this 134 suborder and as outgroup two maricolan taxa. In order to reconstruct ancestral character states 135 related to habitat (epigean / hypogean) and salinity tolerance (freshwater / marine) some 136 recently published sequences were included of marine triclads that show a tolerance to 137 138 freshwater (Table 2).

139 2.2 DNA extraction, gene amplification and sequencing

140 Genomic DNA was extracted from specimens preserved in absolute ethanol by

141 Wizard® Genomic DNA Purification Kit (Promega), according to the manufacturer's

142 instructions. The extraction product was quantified using a NanoDrop 2000c

143 spectrophotometer (Thermo Fisher Scientific, USA). Genomic DNA was used to amplify a

144 fragment of the nuclear genes 18S rRNA (18S) and 28S rRNA (28S) through a polymerase

145 chain reaction (PCR). For 18S amplification we used the primers 18S1F, 18S4F, 18S5R and

18S9R (Carranza et al., 1996) to amplify two overlapping fragments. For 28S amplification 146 we used the primers 28S1F, 28S2F, 28S4R and 28S6R (Álvarez-Presas et al., 2008). The PCR 147 reactions were performed in a final volume of 25µl, with final concentrations as follows: 148 MgCl₂ 2.5mM, dNTPs 30µM, primers 0.4µM each, 0.75U Go Tag® DNA polymerase 149 enzyme (Promega Madison, Wisconsin, USA) with its corresponding buffer (1X), and 150 approximately 100 ng of template DNA. The amplification program for both fragments of 151 18S consisted of 30 cycles in the following manner: 30 seconds at 94°C, 45 seconds at 50°C 152 (AT: annealing temperature), and 1 minute at 72°C, with 2 minutes for initial denaturation at 153 95°C and 4 minutes for final extension at 72°C. The program for both fragments of 28S was 154 155 35 cycles in the following manner: 45 seconds at 94°C, 45 seconds at 55°C (AT), 45 seconds at 72°C, as well as 1 minute of initial denaturalization at 94°C and 3 minutes of final 156 extension at 72°C. The PCR products were purified by ultrafiltration in the Merck Millipore 157 158 MultiScreen System (Darmstadt, Germany). Both chains of purified fragments were sequenced by Macrogen Inc., (Macrogen Europe, Amsterdam). The chromatograms were 159 160 revised and edited with Geneious v. 10 (https://www.geneious.com) to obtain the final 161 contigs.

162 **2.3 Sequence alignment**

Sequences of both genes were independently aligned with MAFFT v7 (Katoh and 163 Standley, 2013) using the web server http://mafft.cbr.jp/alignment/server/ (last visited January 164 15th, 2019) with the G–INS–i algorithm. The following two principal sets of species were 165 166 allocated for the phylogenetic analyses. The first (Sp-dataset -I), was designed to test the monophyly of the Cavernicola, as well as its taxonomic position within the Tricladida. This 167 168 Sp-dataset -I included one representative of each cavernicolan genus, as well as one 169 representative per genus for a series of genera belonging to other suborders of the Tricladida 170 and the outgroup taxa (Table 2). From Sp-dataset -I three subsequent datasets were generated,

viz., two sets for the alignment of the individual genes 18S and 28S (named as Sp-set<u>dataset</u>I-[18S] and Sp-set<u>dataset</u>-I-[28S], respectively), and another for the concatenation of both
genes (named Sp-dataset -I-[18S+28S]).

The second major species set, Sp-dataset_-II, comprised all available sequences of cavernicolan representatives, as well as sequences of two maricolan taxa that were used as outgroups (Table 2). This species set was used to infer phylogenetic relationships within the Cavernicola. From Sp-setdataset_-II three subsequent datasets were generated, two for the individual genes 18S and 28S (called Sp-setdataset_-II-[18S] and Sp-setdataset_-II-[28S], respectively), and a third dataset for the concatenation of both genes (called Sp-setdataset_-II-[18S+28S]).

For the ancestral state reconstructions, another major dataset was created on the basis of an 18S alignment, including as ingroup taxa (a) one representative for each cavernicolan genus, and (b) one representative per genus for a series of genera belonging to other suborders of Tricladida; in the following this dataset is abbreviated as ASR-18S. In this dataset some recently published maricolan taxa were included (Table 2) that live in freshwater or are freshwater-tolerant, in order to have a better representation in this group of the feature "salinity tolerance"; representatives of the Fecampiidae were included as outgroup taxa.

Gblocks (Castresana, 2000) was used to remove regions of ambiguous alignment; the parameters used for each alignment are shown in Supplementary Material Table S1. Each alignment was edited by hand to trim the ends and the code N was assigned to sites with missing data. The concatenated alignments were obtained from the alignments of each gene in Mesquite v. 3.04 (Maddison and Maddison, 2008); in these alignments missing sequences were assigned the code N.

We used Xia's method (Xia and Lemey, 2009), implemented in DAMBE6
software (Xia, 2017), to assess the nucleotide substitution saturation. This test is based on the

concept of entropy in information theory and calculates an index of substitution saturation
(Iss), which is statistically compared to a critical value (Iss.c) that defines a threshold for
significant saturation in the data at which the sequences will begin to fail to recover the true
tree (Xia and Lemey, 2009). We analyzed each alignment by including all sites and using the
proportion of invariant sites previously calculated by the same program.

201 2.4 Phylogenetic Inference

In order to infer the best sequence evolution model for our datasets we used the 202 jModeltest 2.1 software (Darriba et al., 2012), taking into account the scores of the Bayesian 203 204 Information Criterion (BIC). The best model for both genes, calculated independently, was $GTR + \Gamma + I$ (General Time Reversible + Gamma Distribution + Invariable Sites). A gene 205 partition was defined in all the concatenated datasets analyses, so that the estimation of the 206 parameters for each partition was independent. We used two phylogenetic inference methods, 207 viz., Maximum Likelihood (ML) and Bayesian Inference (BI). Both approaches were used to 208 209 analyze each gene independently, as well as for analyzing the concatenated datasets. ML 210 analyses were performed with RAxML v8.2.4 software (Stamatakis, 2014) under the GTRGAMMA model, taking into account the author's recommendations. A rapid bootstrap 211 212 analysis with 2000 pseudoreplicates was conducted to obtain bootstrap support values (bs) for the nodes. We ran Bayesian analyses in MrBayes v3.2.2 software (Ronquist et al., 2012) with 213 5 million generations, sampling frequency every 1000 and 25% burn-in to obtain the 214 consensus tree and posterior probability values (pp). Convergence of the topology and 215 216 parameter values for the two runs was examined by observing that the average standard deviation of split frequencies was below 0.01. Furthermore, the .p file of each run was 217 inspected in Tracer v1.5 software (Rambaut and Drummond, 2007) to ensure that the 218 effective sample size (ESS) values of each parameter were above 200. 219

220 2.5 Ancestral states reconstruction

For the Ancestral States Reconstruction analysis (ASR), we obtained a phylogenetic 221 tree with BI using the ASR-18S dataset dataset III-[18S]. Because the Maricola clade shows a 222 good number of species with various degrees of salinity tolerance, we included in this ASR-223 224 18S dataset dataset III-[18S] (Table 2) 15 species covering the freshwater tolerance diversity in this group. In the salinity tolerance state reconstruction analysis the state for terrestrial 225 planarians was coded as freshwater, since the animals generally depend on the humidity of 226 227 forests soils, which usually will be formed by freshwater. This tree was used to independently estimate the ancestral states for habitat (epigean/hypogean) and salinity tolerance 228 (freshwater/freshwater-marine/marine) by using the Phytools package v.0.6.60 of R (Revell, 229 2012). The posterior probability for each state on the nodes was determined from stochastic 230 character-state maps by applying the empirical Bayes method (Bollback, 2006). For this, we 231 used an all-rates-different (ARD) model and applied the make.simmap function with Markov 232 Chain Monte Carlo (MCMC), and ran 10000 simulations. In the case of salinity tolerance, 233 which has polymorphic states (freshwater/freshwater-marine/marine), we used the fitpolyMk 234 235 function integrated with the make.simmap function.

236 **3. RESULTS**

237 **3.1. Datasets**

The length of the amplified 18S and 28S fragments was approximately 1800 base pairs (bp) and 1500 bp, respectively. For unknown reasons, which may range from problems in the fixation of the specimens, conditions of preservation during transport to intrinsic characteristics of these animals, many of our attempts to obtain good quality DNA for amplification of the genes were unsuccessful. Fortunately, eventually a total of seven new sequences of 18S and of 28S were obtained. After Gblocks processing, the Sp-setdataset - I-

[18S] dataset contained 29 sequences with a total length of 1602 sites, while the Sp-set-I-244 245 28Sdataset I-[28S] data set included 27 sequences with a total length of 1336 sites (Table 2 and Supplementary Material Table S1); these two alignments were concatenated in a Sp-set-I-246 18S+28S dataset I-[18S+28S] dataset (2939 bp and 29 OTUs) and used to infer phylogenetic 247 relationships within the Tricladida. Another dataset with concatenated alignments (Sp-set-II-248 18S+28S dataset II-[18S+28S], with 3199 positions: 1710 for 18S; 1489 for 28S) with 11 249 250 OTUs including only cavernicolan taxa, as well as Procerodes dohrni Wilhelmi, 1909 and Bdelloura candida (Girard, 1850) as outgroups, was obtained and analyzed to infer the 251 252 phylogenetic relationships within the Cavernicola. Finally, the ASR-18S dataset dataset III-253 [18S] included 28 sequences (1709 bp after Gblocks processing; Supplementary material 254 Table S1) and was used for the ancestral character state reconstruction analysis.

255 Saturation analysis of the alignments for each gene, including outgroups, showed no 256 saturation for our datasets; the proportion of invariant sites was 0.17 and 0.22 for 18S and 257 28S, respectively.

258 3.2 Phylogeny

259 The trees obtained from the three Sp-setdatsets -I datasets (Sp-set-I-18S[18S], Sp-set-I-285[285], and Sp-set-I-18S+285[18S+285]) by both inference strategies (ML and Bayesian), 260 all group the cavernicolan taxa into a well-supported monophyletic clade that is sister to 261 Maricola; in turn, Cavernicola + Maricola is sister group of phylogenetically separated from 262 the Continenticola and the Maricola (Fig. 2; Appendix B). However, the topology of the 263 phylogenetic tree inferred from the 28S alone differs in two points from the results obtained 264 265 with the 18S and concatenated datasets. While the Sp-set-I-18S dataset I-[18S] and the Sp-set-I-18S+28S dataset I-[18S+28S] trees (Appendix B.1 Fig. 2) position the Cavernicola as the 266 sister-group of the Maricola with high support (94% bs; 1.00 pp for the 18S tree and 100% bs; 267

1.00 pp for the concatenated), the Sp-set-I-28S dataset I-[28S] tree (Appendix B.2) fails to 268 resolve the relationships between the three triclad suborders, as the Cavernicola groups with 269 270 Continenticola at a poorly supported node (49% bs; 0.6 pp). Further, in the Sp-set-I-271 18Sdataset I-[18S] and the Sp-set-I-18S+28Sdataset I-[18S+28S] trees the genus 272 Opisthobursa groups with Novomitchellia with high support, and together with Hausera and *Kawakatsua* they form a monophylum that is highly supported for 18S (87% bs; 1.00 pp) and 273 receives a low or reasonably good support, depending on the method, for the concatenated 274 275 trees (58% bs; 0.96 pp). In the Sp-set-I-28S dataset I-[28S] tree there is no information on Novomitchellia bursaelongata (no data in GenBank), while in this phylogeny Opisthobursa is 276 277 the sister genus of *Rhodax* with a reasonable good support (73% bs; 0.96 pp).

The analyses of the three Sp-datasets-II datasets (Sp-set-II-18S_[18S], Sp-set-II-28S
[28S], and Sp-set-II-18S+28S[18S+28S]) resulted in phylogenetic trees in which *Opisthobursa* and *Novomitchellia* (only *Opisthobursa* in the case of Sp-set-II-28Sdataset II[28S]) were positioned as the sister-group of a clade formed by *Hausera* and *Kawakatsua*,
with good support in the 18S tree (76% bs; 0.97 pp) and in the concatenated tree (76% bs;
0.99 pp), but with low support in the 28S tree (59% bs; 0.65 pp) (Appendix C, Fig.ure 3).

284 In summary, with respect to the phylogenetic position of the Cavernicola within 285 Tricladida, the Sp-set-I-28S dataset I-[28S] tree showed a polytomy for the three suborders. In 286 contrast, analyses of the 18S and the concatenated datasets for Sp-set-I returned a highly 287 supported clade for Cavernicola + Maricola. With respect to relationships within the 288 Cavernicola, analysis of the Sp-set-I-28Sdataset I-[28S] dataset yielded moderate support for the clade *Rhodax* + *Opisthobursa*, taking into account that *Novomitchelia* is missing from that 289 290 dataset. The other five trees resulting from analyses of the Sp-set-I and Sp-set-II datasets 291 (Fig.ure 2, Fig.ure 3, Appendix B and C) show *Rhodax* as the sister-group of a clade including *Opisthobursa* + *Novomitchelia* and *Hausera* + *Kawakatsua*, generally with high support. 292

293 These results suggest, in our opinion, that the data at hand strongly support the topology294 shown in Fig_ure 2.

295 **3.3 Ancestral habitat**

296 We inferred the ancestral character states for the habitat types (epigean/hypogean) and for salinity tolerance (marine/freshwater) on the nodes of the phylogenetic tree obtained with 297 298 BI from the ASR-18S dataset dataset III-[18S] (Fig. 4; Appendix D). This dataset was used because it renders the same topology as shown in Fig. 2 and does not contain missing data, so 299 that branch lengths will be more accurate than those in the tree resulting from the 300 concatenated dataset. The hypothesis of an epigean habitat for the ancestor of the clade 301 Maricola + Cavernicola was strongly supported by the ancestral state reconstruction analyses 302 (pp=0.97; node 10, Fig 4<u>A</u>, Appendix D<u>.1</u>). In addition, an epigean habitat of the ancestor of 303 304 the Cavernicola is supported with a high posterior probability value (pp=0.80, node 11, Fig. 4A, Appendix D.1). Furthermore, a high support value (pp=0.98; node 11, Fig. 4B, Appendix 305 D.27) suggests that this ancestor lived in a freshwater habitat, while the common ancestor of 306 Maricola + Cavernicola has a nearly equal probability for being either a freshwater animal or 307 308 exhibiting a tolerance to changes in salinity (pp=0.595 freshwater; 0.401 freshwater/marine; 309 node 10, Fig. 4B, Appendix D.2).

310 4. DISCUSSION

4.1 Monophyly of the Cavernicola and its relationship to other Suborders of the

312 Tricladida

The phylogenetic trees obtained in the present study corroborate the monophyly of the Cavernicola, as proposed by Sluys (1990). Monophyly of the Cavernicola was proposed on the basis of three apomorphic features: (a) penis bulb with gland cells, (b) horizontal orientation of the bursal canal or female genital duct, combined with the dorsal opening of the

common oviduct, or diverticulum, and (c) location of the ovaries at some distance posterior to 317 318 the brain (Sluys 1990; Fig. 6). Two of the new species described since Sluys' (1990) monographic study also possess these three features, reinforcing their value as diagnostic 319 characters for the suborder (Leal-Zanchet et al., 2014; Harrath et al., 2016). However, 320 Kawakatsua pumila does not possess a penis bulb with gland cells, while Rhodax does neither 321 exhibit the character "gland cells in the penis bulb" (character 1 in Sluys 1990; see also Fig. 322 6). Thus, currently there are two species of cavernicolans, among the eight species known at 323 present, that lack this presumed apomorphic character state of the Cavernicola, one species 324 (*Rhodax*) being positioned as sister to the rest of genera in the phylogenetic tree and the other 325 326 species (Kawakatsua) positioned at one of its tips. Under present conditions absence of this 327 character state in these two taxa is probably best interpreted as being the result of secondary 328 loss.

With respect to the third presumed apomorphic character for the Cavernicola postulated 329 by Sluys (1990), viz., "ovaries situated at some distance posterior to the brain", this character 330 condition is present in B. thetisae, R. evelinae, O. mexicana, N. bursaelongata, and 331 Kawakatsua pumila (Sluys, 1990; Harrath et al., 2016; Sluys and Laumer 2019), while it is 332 333 absent in O. josephinae and H. hauseri (Sluys, 1990; Leal-Zanchet et al., 2014) and ambiguous for N. sarawakana (Sluys, 1990). This character state distribution casts some 334 doubt on the presumed synapomorphy for the Cavernicola, as the condition for the 335 cavernicolan ancestor, in view of the topology of our tree (Fig. 4), probably downgrades to 336 being equivocal. 337

The five currently known cavernicolan genera are housed in a single family, Dimarcusidae Mitchell & Kawakatsu, 1972, which is supported by the fact that these genera together form a monophyletic clade in our analysis (Fig. 2). In addition, its sister-group relationship to the Maricola and the sister-group relationship shared between Dimarcusidae +

Maricola and the Continenticola in our phylogenetic trees lends further support to Sluys' 342 (1990) proposal of including all cavernicolan species in a separate suborder. However, in 343 contrast to his hypothesis that the Cavernicola is more closely related to freshwater planarians 344 than to marine triclads, our present phylogeny (Fig. 2) corroborates the conclusions of two 345 previous molecular studies (Laumer and Giribet, 2014; Harrath et al., 2016) that the 346 Cavernicola is most closely related to the Maricola. Our results do not support inclusion of the 347 Dimarcusidae in the suborder Maricola, as suggested by Mitchell and Kawakatsu (1972), but 348 do agree with Sluys' (1990) arguments that the Dimarcusidae does not possess the presumed 349 apomorphous character state of the Maricola, nor the derived features of more restrictive 350 351 groups of marine triclads.

Our results falsify the presumed close relationship between the Cavernicola and the 352 freshwater planarians. This is in agreement with the fact that cavernicolans lack two out of the 353 three apomorphies hypothesized for the Paludicola (see Sluys, 1989), viz., the probursal 354 condition, and body musculature with an extra outer layer of subepidermal longitudinal fibers. 355 Furthermore, although the third apomorphic feature proposed for the Paludicola, presence of a 356 spermatophore, has been described for O. josephinae by Benazzi (1975), it has not been 357 358 observed in any other cavernicolan species (Sluys 1990; AMLZ, pers. obs.). In point of fact, Sluys (1990, p. 26) himself recognized that "... the data set at hand suggests little else than a 359 sistergroup relationship between the Paludicola and the Dimarcusidae, although this presumed 360 affinity remains poorly supported by apomorphous characters". 361

362 In summary, currently we can recognize three suborders within the Tricladida, viz.,

363 Cavernicola, Maricola, and Continenticola, which show clear differences in their morphology

364 (Sluys 1990) and are genetically highly differentiated.

365 4.2 Relationships and diversity within the Cavernicola

The phylogenetic relationships within the Cavernicola as revealed by molecular data (Figs 2 and 3) differ from those proposed on the basis of morphological apomorphies (Fig. 5). In the phylogeny presented here, *Rhodax* is sister to all other taxa of the Cavernicola. This implies some changes in our interpretation of the evolution of several morphological features, as will be discussed below.

The clade formed by *Opisthobursa* and *Novomitchellia* is highly supported (Fig. 3), as 371 well as its sister-group relationship to the clade comprising Hausera and Kawakatsua. This 372 373 casts some doubt on the four synapomorphies proposed for the sister-group relationship between Rhodax and Opisthobursa as proposed by Sluys (1990): (a) ciliation being confined 374 to the posterior section of the bursal canal or female genital duct, (b) vitellaria being situated 375 medially to the testes, (c) loss of the primary copulatory bursa, and (d) presence of testes 376 tubes, i.e., fused testicular follicles (characters 5-8 in Fig. 5). In fact, Opisthobursa presents a 377 copulatory bursa, but Sluys (1990) pointed out that it is a secondary bursa and not a primary 378 one. In view of the present phylogenetic results, loss of the primary copulatory bursa in both 379 *Rhodax* and *Opisthobursa* may be interpreted as having evolved independently. 380

Fused testicular follicles or testis tubes, as present in *R. evelinae* and *O. josephinae* (character 8 in Fig. 5), represent a rare condition among triclad flatworms (see Sluys and Riutort, 2018). In view of the fact that *O. mexicana* has discrete testes follicles (Sluys 1990), presence of testis tubes in *R. evelinae* and *O. josephinae* is presently best interpreted as the result of convergent evolution.

Sluys (1990) hypothesized that oviducts running medially to the ventral nerve cords
represent a derived character linking the genera *Rhodax*, *Opisthobursa* and *Novomitchellia*,
not shared with *Balliania*, which situated the latter genus as sister to the rest of Cavernicola.
Unfortunately, molecular data for *Balliania* is not available and, therefore, we cannot put

forward a hypothesis on its position within the Cavernicola. However, for the genus Hausera 390 391 the situation is different, in that the oviducts are exactly dorsal to the nerve cords (Leal-Zanchet et al., 2014); the precise character state for the course of the oviducts in relation to 392 the ventral nerve cords is not known for *Kawakatsua pumila* (Sluys and Laumer, 2019). 393 Under the present phylogeny "oviducts running medially to the ventral nerve cords" may still 394 be postulated as a synapomorphy for all cavernicolans included in our analysis (Fig. 3), under 395 the assumption that at least Hausera has evolved another character state in which the oviducts 396 397 run dorsally to the ventral nerve cords.

The high differentiation found in the tree (Fig. 3) between some of the representatives 398 of *Rhodax* included in our study may point to the presence of more than one species, but this 399 issue may be resolved only by a thorough study that also includes morphological data, which 400 currently is unavailable. Even more surprising is the low genetic differentiation between the 401 genera Kawakatsua and Hausera, while these are well differentiated at the morphological 402 level (Sluys and Laumer, 2019) and are geographically distant. This situation clearly shows, 403 on the one hand, that a broader sampling most probably will reveal new species for the 404 Cavernicola, and, on the other hand, that within this suborder levels of genetic diversification 405 may vary among different groups. 406

407 **4.3 Origin and evolution of the Cavernicola**

One of the most conspicuous features of this group is its rarity (only a handful of genera, each with merely one or two species, mostly present at a single locality), together with their disjunct distributions (Fig. 1). The fact that currently known species occur at tropical latitudes and that cavernicolans thus far have not been reported from relatively well sampled areas such as Europe and North America suggests that the Cavernicola fauna has a predominantly intertropical distribution. However, it is important to realize that the current

distribution, including its disjunctions, may simply reflect a collector's artefact, due to
comparatively low investments in research of subterranean habitats in most regions of the
world, excepting Europe and North America.

Our phylogenetic trees suggest that the Cavernicola forms an old group. Although no 417 418 strict time calibration of the entire order Tricladida has been performed, the few molecular timetrees published for *Dugesia* (Solà, 2014; Solà et al., 2013) have situated the 419 diversification of this continenticolan freshwater genus at approximately 150 million years 420 421 ago (Mya), which implies that the origin of the Continenticola lies even much further back in time. From this perspective, the present phylogenetic tree suggests a great antiquity also for 422 423 the Cavernicola. This agrees with one of the possibilities for the evolution of the Cavernicola proposed by Sluys (1990), i.e., that the group had differentiated on Gondwana and perhaps 424 already on Pangea. This hypothesis seems plausible given the sister-group relationship 425 between Opisthobursa and Novomitchellia, genera present in Mexico and Benin, respectively; 426 427 in turn, this clade is sister to the clade formed by Hausera and Kawakatsua, from northeastern Brazil and Panama, respectively (Figs 1, 3). This suggests that the common ancestor of these 428 429 four lineages may have lived on Gondwana before this supercontinent started to breakup 430 around 200 Mya (McLoughlin, 2001; Storey, 1995). As Panama and Mexico were not part of Gondwanaland, this scenario presumes that after breakage some descendants of the American 431 lineage dispersed to North America via the Panamanian isthmus when North and South 432 America were eventually connected with each other. 433

Despite the presumed antiquity of the group, one could argue, alternatively, that it exhibits signs of recent dispersal since both Panama and Tahiti are the result of recent (in the last 5 million years) volcanism, with the latter being a true oceanic island with no connection to any continental landmass. Therefore, an alternative hypothesis might be that there must have been a mechanism for dispersal. But if cavernicolans could disperse, relatively recently.

to such habitats, it may similarly have been possible for them to disperse away from 439 Gondwana in ancient times, subsequent to its breakup, or even to Gondwana from elsewhere. 440 One may be tempted to favour such dispersal explanations in view of the situation that an 441 hypothesis about Gondwanan origins of the Cavernicola currently lacks any representation 442 from South Africa, southern South America, Australia, and New Zealand. Generally, much of 443 the evidence for true Gondwanan relicts in other taxa hinges on representation from such 444 areas and even then molecular timetrees may falsify presumed evolution on Gondwana, as 445 446 was the case with ratite birds (Reilly, 2019 and references therein). However, it may well 447 have been the case that a taxon had a distribution that was restricted already to a certain 448 region of Gondwana, which may have applied also to the Cavernicola.

Although calibrated timetrees are as yet not available for the Cavernicola, what is 449 known about the absolute age of triclad flatworms (see above) suggests that the group is 450 ancient. Accepting the antiquity of the group, one may wonder whether the cavernicolans had 451 already evolved their troglobitic adaptations on Pangea or Gondwana. Harrath et al. (2016) 452 put forward two alternative hypotheses for the origin of the Cavernicola. According to their 453 454 first hypothesis, which was based on the close phylogenetic relationship of the Cavernicola to 455 the Maricola, marine ancestors were forced to invade the underground habitat due to gradual recession of the sea, after which the worms adapted to the hypogean freshwater habitat. A 456 similar scenario was suggested to explain the ecology of Hausera hauseri from the karstic 457 Jandaíra formation in northeastern Brazil (Leal-Zanchet et al., 2014). Under their first 458 scenario, Harrath et al. (2016) proposed that epigean R. evelinae would have evolved from 459 stygobiont populations and have again acquired the eyes that were lost in its underground 460 ancestors, a possibility that has also been hypothesized for some crustaceans (Humphreys, 461 2000, and references therein). The second scenario sketched by Harrath et al. (2016) for the 462 evolution of the Cavernicola proposed that an ancestral brackish- and freshwater-tolerant 463

464 epigean maricolan species led to a brackish water-tolerant *Rhodax* ancestor and to a lineage
465 that invaded the phreatic habitat, possibly to escape presumed competition with
466 continenticolans. In this scenario presence of pigmentation and eyes in *Rhodax* simply reflects
467 retention of the ancestral character states.

468 Our ancestral states analyses revealed that the character conditions for the ancestor of the Cavernicola are most probably epigean and freshwater (0.80 and 0.987 posterior 469 470 probability, respectively, appendix D, Fig. 4, Appendix D), implying diversification of the 471 cavernicolans from worms originally adapted to continental epigean freshwater habitats. In this scenario presence of pigmentation and eyes in *Rhodax* then most probably reflects 472 473 retention of the ancestral conditions, while for *Kawakatsua pumila* the probability of having retained the ancestral epigean character state is lower (its ancestor with Hausera has a 0.51 474 probability of having been epigean). With respect to salinity tolerance, our analyses show that 475 the ancestor of the Maricola + Cavernicola has a higher probability of being a freshwater 476 animal, or at least being tolerant to freshwater, than that it was a marine species (0.55 and 477 478 0.41 vs. 0.03, node 10, Fig. 4A, Appendix D.1, Figure 4); therefore, the Cavernicola may not 479 have had a marine ancestor. This lends support to the second scenario for the evolution of the 480 Cavernicola sketched by Harrath et al. (2016). On the other hand, our results falsify the scenario suggested by Leal-Zanchet et al. (2014) that Hausera evolved directly from marine 481 ancestors that entered the caves and then adapted to the freshwater environment. 482

We hypothesize here that the scenario for the Cavernicola, with dispersal of freshwater animals and colonization of hypogean habitats, resembles cases currently known for the Continenticola, such as representatives of *Girardia* from South America (Souza et al., 2015, 2016; Hellmann et al., 2018), and many species of the Dendrocoelidae in Europe (Stocchino et al., 2017, and references therein) that occur in caves or in surface waters. However, in contrast to *Girardia*, in which epigean species outnumber hypogean taxa, epigean species of

the Cavernicola presently represent a minority, as compared with hypogean members of the 489 490 same suborder (for which at least three more undescribed species occur in the Jandaíra formation; AMLZ, unpublished data), or with epigean species of the Continenticola. This 491 scarcity of cavernicolans in general and that of epigean ones in particular, may be the result of 492 a loss of diversity due to climatic changes or competition with other groups. Specifically, the 493 loss of suitable epigean habitats, and/or competition with continenticolan species, the latter 494 group showing a broad radiation in the same biogeographic regions that also house 495 cavernicolans, could underlie the paucity of epigean cavernicolans. Although these 496 explanations may seem highly speculative with the data at hand, the karstic Jandaíra 497 498 formation in northeastern Brazil, where *Hausera* species inhabit (Appendix E1 and E2), may be an example of such a loss of suitable epigean habitat. In this area, surface karstification 499 forms recharge zones, favouring the storage and flow of subterranean water (Miranda, 2012), 500 501 constituting the only water source in most of the region, where epigean streams are scarce (Fernandes et al., 2005). Therefore, no epigean species are expected to be able to survive 502 under these conditions. Thus, under this tentative scenario, caves may have become a refuge 503 for the cavernicolans. 504

505 In summary, our results lend support to the hypothesis of a freshwater ancestor of the 506 Cavernicola that colonized continental epigean and phreatic habitats and, subsequently, 507 radiated to form a diverse group with a broad distribution, probably on Gondwanaland. Under 508 this scenario the evolution of the Cavernicola constitutes a classical example of evolutionary 509 diversification, followed by independent adaptations to hypogean habitats, where caves may 510 have become a refuge habitat for the group for reasons still not fully understood.

511

512 **5. ACKNOWLEDGEMENTS**

We thank Christopher Laumer for kindly providing information on the locality and some
external characteristics of *Kawakatsua pumila* before the species was formally described. We
also thank Halim Harrath for making available to us a photograph of the type locality of *Novomitchellia bursaelongata*. We acknowledge Lindsey Hellmann for the Appendix A
figure.

518

519 **6. FUNDING**

- 520 This work was supported by the Brazilian Research Council CNPq [grants 306853/2015-9
- and 304682/2014-4]; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
- 522 (CAPES); and Ministerio de Economía y Competitividad of Spain [project CGL2015-63527-
- 523 1P]. Fieldwork in Mexico was funded by PAPIIT [IN210318] and SEP-CONACYT Ciencia
- Básica [220408]. AO-F and MR are grateful for support from the Unión Iberoamericana de
- 525 Universidades (UIU) [Project BIO02/2017].
- 526

527 **7. REFERENCES**

- Álvarez-Presas, M., Baguñà, J., Riutort, M., 2008. Molecular phylogeny of land and
 freshwater planarians (Tricladida, Platyhelminthes): From freshwater to land and back.
 Mol. Phylogenet. Evol. 47, 555–568.
- Ball, I.R., 1977. On the phylogenetic classification of aquatic planarians. Acta Zool. Fenn.
 154, 21–35.
- 533 Ball, I.R., 1974. A contribution to the Phylogeny and Biogeography of the Freshwater
- 534 Triclads (Platyhelminthes: Turbellaria), in: Riser, N.W., Morse, M.P. (Eds.), Biology of
- the Turbellaria, McGraw-Hill, NY. New York, pp. 339-401.

- 536 Benazzi, M., 1975. *Opisthobursa josephinae*, a new troglobitic planarian from Chiapas,
- 537 Mexico. Atti della Accad. Nazi onale dei Lincei LIX, 533–536.
- 538 Bollback, J., 2006. SIMMAP: Stochastic character mapping of discrete traits on phylogenies.
- 539 BMC Bioinformatics 7, 88. https://doi.org/10.1186/1471-2105-7-88
- 540 Carranza, S., Giribet, G., Ribera, C., Riutort, M., 1996. Evidence that two types of 18S rDNA
- 541 coexist in the genome of *Dugesia* (Schmidtea) *mediterranea* (Platyhelminthes,
- 542 Turbellaria, Tricladida). Mol. Biol. Evol. 13, 824–832.
- 543 https://doi.org/10.1093/oxfordjournals.molbev.a025643
- Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in
 phylogenetic analysis. Mol Biol Evol 17, 540–552.
- 546 Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: More models, new
 547 heuristics and parallel computing. Nat. Methods 9, 772.
- 548 https://doi.org/10.1038/nmeth.2109
- 549 Fernandes, M.A.B., Santiago, M.F., Gomes, D.F., Filho, J.M., Frischkorn, H., Lima, J.G.,
- 550 2005. A origm dos cloretos nas águas subterrâneas na Chapada do apodi-Ceará. Águas
 551 subterrâneas 19, 25–34.
- Harrath, A.H., Mansour, L., Lagnika, M., Sluys, R., Boutin, C., Alwasel, S., Poch, A., Riutort,
- 553 M., 2016. A molecular analysis of the phylogenetic position of the suborder Cavernicola
- within the Tricladida (Platyhelminthes), with the description of a new species of
- stygobiont flatworm from Benin. Zool. J. Linn. Soc. 178, 482-491.
- 556 https://doi.org/10.1111/zoj.12430
- 557 Humphreys, W., 2000. Relict fauna and their derivation, in: Wilkens, H., Culver, D.,
- Humphreys, W. (Eds.), Subterranean Ecosystems. Elsevier, Amsterdam, pp. 417–432.

559	Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7:
560	Improvements in Performance and Usability. Mol. Biol. Evol. 30, 772-780.

- 561 https://doi.org/10.1093/molbev/mst010
- 562 Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular
- 563 Evolutionary Genetics Analysis across Computing Platforms. Mol. Biol. Evol. 35, 1547–
- 564 1549. https://doi.org/10.1093/molbev/msy096
- Laumer, C.E., Giribet, G., 2014. Inclusive taxon sampling suggests a single, stepwise origin
- of ectolecithality in Platyhelminthes. Biol. J. Linn. Soc. 111, 570–588.
- 567 https://doi.org/10.1111/bij.12236
- Laumer, C.E., Hejnol, A., Giribet, G., 2015. Nuclear genomic signals of the
- ⁵⁶⁹ "microturbellarian" roots of platyhelminth evolutionary innovation. Elife 4, 1–31.
- 570 https://doi.org/10.7554/eLife.05503
- 571 Leal-Zanchet, A.M., De Souza, S.T., Ferreira, R.L., 2014. A new genus and species for the
- 572 first recorded cave-dwelling Cavernicola (Platyhelminthes) from South America.
- 573 Zookeys 15, 1–15. https://doi.org/10.3897/zookeys.442.8199
- 574 Maddison, W.P., Maddison, D.R., 2008. Mesquite: A modular system for evolutionary

analysis. Evolution (N. Y). 62, 1103–1118. https://doi.org/http://mesquiteproject.org

- 576 Marcus, E., 1946. Sôbre Turbellaria límnicos brasileiros. Bol. Fac. Filos. Ciênc. Let., Zool.
 577 Univ. São Paulo 11, 5–254.
- McLoughlin, S., 2001. The breakup history of Gondwana and its impact on pre-Cenozoic
 floristic provincialism. Aust. J. Bot. 49, 271-300. https://doi.org/10.1071/BT00023
- 580 Miranda, T.S. De, 2012. Estudo do comportamento geológico-estrutural dos depósitos

- 581 Cretácicos do sistema aquífero Apodi, Bacia Potiguar, Ne do Brasil. Estud. Geológicos
 582 22, 3–19.
- Mitchell, R.W., Kawakatsu, M., 1972. A new family, genus, and species of cave-adapted
 planarian from Mexico (Turbellaria, Tricladida, Maricola). Occas. Pap. Museum Texas
 Tech Univ. 8, 1–16.
- 586 Norén, M., Jondelius, U., 2002. The phylogenetic position of the Prolecithophora
- 587 (Rhabditophora, 'Platyhelminthes'). Zool. Scr. 31, 403–414.
- 588 https://doi.org/10.1046/j.1463-6409.2002.00082.x
- 589 Rambaut, A., Drummond, A.J., 2007. Tracer v1.4. Disponible en: beast.bio.ed.ac.uk/Tracer.
- 590 Reilly, J., 2019. The Ascent of Birds How Modern Science is Revealing their Story. Pelagic
 591 Publishing, Exeter.
- Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other
 things). Methods Ecol. Evol. 3, 217–223. https://doi.org/10.1111/j.2041-
- 594 210X.2011.00169.x
- Riutort, M., Álvarez-Presas, M., Lazaro, E.M., Solà, E., Paps, J., 2012. Evolutionary history
 of the Tricladida and the Platyhelminthes: an up-to-date phylogenetic and systematic
 account. Int. J. Dev. Biol. 56, 5–17. https://doi.org/10.1387/ijdb.113441mr
- 598 Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B.,
- Liu, L., Suchard, M. a, Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian
- 600 phylogenetic inference and model choice across a large model space. Syst. Biol. 61,
- 601 539–42. https://doi.org/10.1093/sysbio/sys029
- 602 Sluys, R., 1990. A monograph of the Dimarcusidae (Platyhelminthes, Seriata, Tricladida).

603 Zool. Scr. 19, 13–29.

- Sluys, R., 1989. Phylogenetic relationships of the triclads (Platyhelminthes, Seriata,
 Tricladida). Bijdr. Dierk. 59, 3–25.
- 606 Sluys, R., Kawakatsu, M., Riutort, M., Baguñà, J., 2009. A new higher classification of
- 607 planarian flatworms (Platyhelminthes, Tricladida). J. Nat. Hist. 43, 1763–1777.

608 https://doi.org/10.1080/00222930902741669

- 609 Sluys, R., Laumer, C., 2019. A new genus and species of a terrestrial cavernicolan planarian
- 610 from Barro Colorado Island, Panama (Platyhelminthes, Tricladida, Cavernicola).
- 611 Zootaxa 4586, 187–193.
- 612 Sluys, R., Riutort, M., 2018. Planarian Diversity and Phylogeny, in: Rink, J.C. (Ed.),
- 613 Planarian Regeneration: Methods and Protocols. Methods in Molecular Biology, Vol.
- 614 1774. Springer, New York, pp. 1–56. https://doi.org/10.1007/978-1-4020-6682-5
- 615 Sluys, R., Vila-Farré, M., Álvarez-Presas, M., Riutort, M., Kawakatsu, M., Tulp, A.S., 2015.
- 616 The diet and distribution of *Pentacoelum kazukolinda* (Kawakatsu & Mitchell, 1984), a
- 617 maricolan planarian with a freshwater ecology. Zool. Scr. 44, 72–91.
- 618 https://doi.org/10.1111/zsc.12084
- Solà, E., 2014. Systematics and historical biogeography of the genus Dugesia. Unpublished
 PhD thesis. Universitat de Barcelona.
- Solà, E., Sluys, R., Gritzalis, K., Riutort, M., 2013. Fluvial basin history in the northeastern
 Mediterranean region underlies dispersal and speciation patterns in the genus *Dugesia*(Platyhelminthes, Tricladida, Dugesiidae). Mol. Phylogenet. Evol. 66, 877–888.
- 624 Souza, S., Morais, A.L., Bichuette, M.E., Leal-Zanchet, A.M., 2016. Two new species of

- 625 freshwater flatworms (Platyhelminthes: Tricladida: Continenticola) from South
- 626 American caves. Zootaxa 4092, 107–121.
- 627 Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
 628 large phylogenies. Bioinformatics 30, 1312–1313.
- 629 https://doi.org/10.1093/bioinformatics/btu033
- 630 Stocchino, G.A., Sluys, R., Kawakatsu, M., Sarbu, S.M., Manconi, R., 2017. A new species of

631 freshwater flatworm (Platyhelminthes, Tricladida, Dendrocoelidae) inhabiting a

632 chemoautotrophic groundwater ecosystem in Romania. Eur. J. Taxon. 342, 1–21.

633 Storey, B.C., 1995. The role of mantle plumes in continental breakup: case histories from

634 Gondwanaland. Nature 377, 301–308. https://doi.org/10.1038/377301a0

- Xia, X., 2017. DAMBE6: New tools for microbial genomics, phylogenetics, and molecular
 evolution. J. Hered. 108, 431–437. https://doi.org/10.1093/jhered/esx033
- Kia, X., Lemey, P., 2009. Assessing substitution saturation with DAMBE., in: Lemey, P.,
- 638 Salemi, M., Vandamme, A.-M. (Eds.), The Phylogenetic Handbook: A Practical
- Approach to DNA and Protein Phylogeny. Cambridge University Press, Cambridge, pp.
- 640 615–630. https://doi.org/10.1017/CBO9780511819049.022

641

642 LEGENDS TO THE FIGURES AND TABLES

Figure 1. Distribution of cavernicolan taxa; all known sites are shown. The taxa included inthis study are highlighted in boldface.

Figure 2. Bayesian Inference tree inferred from the dataset including 18S and 28S sequences
of representatives of the various Suborders of the Tricladida (Sp-set-I-18S+28Sdataset I-

[18S+28S]). Values at nodes correspond to posterior probability/bootstrap support. *: 1.00
and 100% values for BI and ML, respectively. Scale bar: number of substitutions per
nucleotide position.

650 Figure 3. Relationships within the Cavernicola inferred by Bayesian Inference from the Sp-

651 <u>set-II-18S+28S</u><u>dataset II-[18S+28S]</u> dataset. Numbers at nodes indicate posterior

probability/bootstrap supports for BI and ML, respectively. *: 1.00 and 100% values for BI

and ML, respectively. Scale bar indicates number of substitutions per site.

Figure 4. Results of the Ancestral States Reconstruction analysis based on the Bayesian

655 Inference tree obtained from the ASR-18S dataset dataset III-[18S]. Pie charts at nodes

represent the posterior probabilities of ASR analysis for A: epigean (green) and hypogean

657 (yellow) habitat and B: freshwater (red), freshwater-marine (purple) and marine (blue) salinity

tolerance. Terrestrial species were scored for the freshwater condition, as they are only able to

659 survive in a humid habitat. For exact posterior probabilities obtained at each node, see

660 Appendix D.

Figure 5. Morphological phylogeny of Cavernicola based on Sluys (1990). Black rectangles
represent morphological apomorphies.

663

Table 1. Geographic coordinates of the type localities, and habitat as well as stygobiontconditions of cavernicolan taxa.

Table 2. Species names, higher taxa, and GenBank accession numbers for sequences used in
the analyses. Species selected to represent each of the cavernicolan genera highlighted in
boldface.

670 LEGENDS TO THE APPENDICES

Appendix A. Dorsal view of live specimens of *Rhodax* sp. 3 (A1) from surface water in

- 672 Tramandaí, southern Brazil, and *Hausera* sp. (A2) from Furna Feia cave, northeastern Brazil.
- 673 Anterior end to the left.
- Appendix B. Bayesian Inference tree of Tricladida suborders based on the Sp-set-I-18Sdataset
 I-[18S] (B1) and Sp-set-I-28Sdataset I-[28S] (B2) datasets. Numbers at nodes indicate
 posterior probability/bootstrap supports for BI and ML, respectively. *: 1.00/100% values for
- 677 BI/ML. Scale bar indicates number of substitutions per site.
- 678 Appendix C. Bayesian Inference tree of Suborder Cavernicola inferred from the Sp-set-II-
- 679 <u>18Sdataset II-[18S]</u> (C1) and <u>Sp-set-II-28Sdataset II-[28S]</u> (C2) data sets. Numbers at nodes
- 680 indicate posterior probability/bootstrap supports for BI and ML, respectively. *: 1.00/100%
- values for BI/ML. Scale bar indicates number of substitutions per site.
- Appendix D. Posterior probability values at each node in Figure 4 for the conditions habitat
- 683 (epigean/hypogean) (D1) and salinity tolerance (freshwater/freshwater-marine/marine) (D2),
- obtained in the Ancestral States Reconstruction analysis using the make.simmap fuction from
- 685 the R package Phytools.
- Appendix E. Habitats of cavernicolan taxa included in this study. E1: Crotes cave,
- 687 northeastern Brazil (type locality of *Hausera hauseri*); E2: Furna Feia cave, northeastern
- 688 Brazil (*Hausera* sp.); E3: Parakou, Benin (type locality of *Novomitchellia bursaelongata*);
- 689 E4: Grutas de Coconá cave, México (type locality of *Opisthobursa mexicana*); E5: Mostardas,
- 690 southern Brazil (*Rhodax* sp.1); E6: Tavares, southern Brazil (*Rhodax* sp.1); E7: Santo
- 691 Antonio de Patrulha, southern Brazil (*Rhodax* sp.2); E8: Tramandaí, southern Brazil (*Rhodax*
- 692 sp.3).

Table 1. Geographic coordinates of the type localities, and habitat as well as stygobiont conditions of cavernicolan taxa.

Taxon	Locality Geographic coordinates		Habitat	Stygobiont features	
<i>Rhodax evelinae</i> Marcus, 1946	Sao Paulo city, São Paulo, Brazil	-	Lotic and lentic superficial waters	No	
<i>Opisthobursa mexicana</i> Benazzi, 1972	Las Grutas de Coconá, Teapa, Tabasco, Mexico	17.616667, -92.966667	Cave	Yes	
<i>Opisthobursa josephinae</i> Benazzi, 1975	Pozza Casa Bell, San Cristobal de las Casas, Estado de Chiapas, Mexico	16.716667, -92.666667	Cave	Yes	
<i>Balliania thetisae</i> Gourbault, 1978	Maraa, Paea, Tahiti ^a	-	Phreatic layer	Yes	
<i>Opisthobursa</i> sp. Kawakatsu & Mitchell, 1983	Grutas de Languín, Alta Verapaz, Guatemala	15.573611, -89.980556	Cave	Yes	
<i>Kawakatsua pumila</i> Sluys, 2019	Barro Colorado, Panama	9.15265, -79.85172	Large pile of humid leaf mulch between two buttress roots of an old broadleaf tree	Yes	
<i>Hausera hauseri</i> Leal-Zanchet & Souza, 2014	Crotes cave, Felipe Guerra, Rio Grande do Norte, Brazil	-5.592344, -37.686183	Cave	Yes	
<i>Novomitchellia sarawakana</i> (Kawakatsu & Chapman, 1983)	Water Polo Cave, Gunung Mulu National Park, Sarawak, Malaysia	4.000000, 114.852778	Cave	Yes	
<i>Novomitchellia bursaelongata</i> Harrath, Sluys & Riutort, 2016	Parakou, Benin Republic	9.272972, 2.581861	Waterhole (7.8 m depth and 6.6 m height)	Yes	
<i>Hausera</i> sp. 1	Furna Feia cave, Baraúna, Rio Grande do Norte, Brazil	-5.036877, -37.560177	Cave	Yes	
Rhodax sp. 1_1	Tavares, Rio Grande do Sul, Brazil	-31.280277, -51.060555	Coastal wetland	No	
Rhodax sp. 1_2	Tavares, Rio Grande do Sul, Brazil	-31.317500, -51.122777	Coastal wetland	No	
Rhodax sp. 2	Pinheirinhos, Santo Antônio da Patrulha, Rio Grande do Sul, Brazil	-29.71205, -50.638233	Rice field	No	
Rhodax sp. 3	Tramandaí, Rio Grande do Sul, Brazil	-30.087777, -50.170833	Coastal wetland	No	

695 ^a Temporary water course close to the Insectarium of the Institut de Recherches médicales «Louis Malardé»

Table 2. Species names, taxonomic classification, and GenBank accession numbers for the Plathyheminthes sequences used in the analyses.
 Species selected to represent each of the cavernicolan genera are highlighted in boldface.

Species name	Taxon	18S	28S
Protomonotresidae sp.ª	Rhabditophora: Prolecithophora: Protomonostresidae	KC869820	KC869873
Acanthiella sp.ª	Rhabditophora: Prolecithophora: Protomonotresidae	KC869786	KC869839
Reisingeria hexaoculataª	Rhabditophora: Prolecithophora: Pseudostomidae	AF065426	AY157157
Plagiostomum stellatum ^a	Rhabditophora: Prolecithophora: Plagiostomidae	KC869819	KC869872
Plagiostomum whitmaniª	Rhabditophora: Prolecithophora: Plagiostomidae	KC869818	KC869871
Plicastoma cuticulataª	Rhabditophora: Prolecithophora: Plagiostomidae	AF065422	AY157158
Notentera ivanoviaª	Rhabditophora: Rhabdocoela: Dalyellioida: Fecampiidae	AJ287546.1	AY157167.1
Kronborgia isopodicolaª	Rhabditophora: Rhabdocoela: Dalyellioida: Fecampiidae	AJ012513.1	AF022862.1
Urastoma cyprinaeª	Rhabditophora: Mediofusata: Urastomidae	AF065428.2	AY157165.1
Bdelloura candida	Tricladida: Maricola: Bdellouroidea: Bdellouridae	Z99947.1	AJ270167.1
Nerpa fistulata ^ь	Tricladida: Maricola: Bdellouroidea: Bdellouridae	MH916614.1	-
Palombiella stephensoni	Tricladida: Maricola: Bdellouroidea: Bdellouridae	DQ666008.2	DQ665988.1
Pentacoelum kasukolinda	Tricladida: Maricola: Bdellouroidea: Bdellouridae	JN009784.1	JN009787.1
Pentacoelum sinensis ^ь	Tricladida: Maricola: Bdellouroidea: Bdellouridae	MK140782.1	-
Sluysia triapertura	Tricladida: Maricola: Bdellouroidea: Uteriporidae	MF383119.1	MF383122.1
Paucumara falcata ^b	Tricladida: Maricola; Bdellouroidea; Uteriporidae	MH916612.1	-
Miroplana shenzhensis ^b	Tricladida: Maricola; Bdellouroidea; Uteriporidae	MK140778.1	-
Ectoplana limuli ^b	Tricladida: Maricola; Bdellouroidea; Uteriporidae	D85088.1	-
Obrimoposthia wandeli ^b	Tricladida: Maricola; Bdellouroidea; Uteriporidae	MH108586.1	-
Uteriporus sp	Tricladida: Maricola: Bdellouroidea; Uteriporidae	AF013148.1	-
Cercyra hastata	Tricladida: Maricola: Cercyroidea: Cercyridae	KM200902.1	DQ665962.1
Sabussowia dioica	Tricladida: Maricola: Cercyroidea: Cercyridae	JN009785.1	JN009788.1
Oregoniplana geniculata ^b	Tricladida: Maricola: Cercyroidea: Cercyridae	MH916614.1	-

Procerodes littoralis	Tricladida: Maricola: Procerodoidea: Procerodidae	Z99950.1	DQ665985.1
Hausera hauseri *	Tricladida: Cavernicola: Dimarcusidae	<u>MN719501</u>	<u>MN719494</u>
<i>Hausera</i> sp. 1*	Tricladida: Cavernicola: Dimarcusidae	<u>MN719502</u>	<u>MN719495</u>
Kawakatsua pumila	Tricladida: Cavernicola: Dimarcusidae	KC869823.1	KC869876.1
Novomitchellia bursaelongata	Tricladida: Cavernicola: Dimarcusidae	KU096054.2	-
Opisthobursa mexicana*	Tricladida: Cavernicola: Dimarcusidae	<u>MN719503</u>	<u>MN719496</u>
<i>Rhodax</i> sp. 1_1*	Tricladida: Cavernicola: Dimarcusidae	<u>MN719504</u>	<u>MN719497</u>
Rhodax sp. 1_2*	Tricladida: Cavernicola: Dimarcusidae	<u>MN719505</u>	<u>MN719498</u>
<i>Rhodax</i> sp. 2*	Tricladida: Cavernicola: Dimarcusidae	<u>MN719506</u>	<u>MN719499</u>
<i>Rhodax</i> sp. 3*	Tricladida: Cavernicola: Dimarcusidae	<u>MN719507</u>	<u>MN719500</u>
Dugesia gonocephala	Tricladida: Continenticola: Geoplanoidea: Dugesiidae	DQ666002.1	DQ665965.1
<i>Girardia</i> sp.	Tricladida: Continenticola: Geoplanoidea: Dugesiidae	AF013156.1	DQ665977.1
Schmidtea polychroa	Tricladida: Continenticola: Geoplanoidea: Dugesiidae	AF013154.1	DQ665993.1
Geoplana quagga	Tricladida: Continenticola: Geoplanoidea: Geoplanidae	KC608497.1	KC608380.1
Cephaloflexa bergi	Tricladida: Continenticola: Geoplanoidea: Geoplanidae	KJ599712.1	KC608355.1
Phagocata vitta	Tricladida: Continenticola: Planarioidea: Planariidae	DQ665998.1	DQ665989.1
Crenobia alpina	Tricladida: Continenticola: Planarioidea: Planariidae	M58345.1	DQ665960.1

 *: sequences obtained in this study, Gen Bank accession number pending
 ^a: outgroup sequences
 ^b: Maricola species included only in the Ancestral Reconstruction States analysis 700 701











- 1: penis bulb with gland cells
- 2: T-junction bursal canal/female genital duct and common oviduct/diverticulum
- 3: ovaries at some distance posterior to the brain
- 4: oviducts medially to ventral nerve cords
- 5: posterior part bursal canal/female genital duct ciliated
- 6: vitellaria medially to the testes
- 7: absence of primary copulatory bursa
- 8: testes tubes
- 9: two gonopores

Author contibution statement

MR and AML-Z conceived the project. MR, AML-Z and AO-F obtained the funding suport. AO-F, RLF, DMB, JB contributed to the samplings. LB-A did all the wet laboratory and data analyses. MR and LB-A wrote the first draft of the manuscript. All the authors participated in the discussions of the results and on the writing of the manuscript. All authors read and approved the final version of the manuscript.











0.04

0.1





Appendix D. Posterior probabilities of each condition (D.1: habitat and D.2: salinity) obtained in the ancestral reconstruction states analysis.

Probability				Pro	bability	
Node	Epigean	Hypogean	Node	Freshwater	Freshwater / Marine	Marine
1	0,88	0,13	1	0.22	0.52	0.26
2	0,93	0,07	2	0.03	0.22	0.75
3	0,98	0,02	3	0.64	0.35	0.01
4	1,00	0,00	4	0.97	0.03	0.00
5	1,00	0,00	5	0.99	0.01	0.00
6	1,00	0,00	6	0.99	0.01	0.00
7	1,00	0,00	7	1.00	0.00	0.00
8	1,00	0,00	8	1.00	0.00	0.00
9	1,00	0,00	9	1.00	0.00	0.00
10	0,97	0,03	10	0.59	0.40	0.01
11	0,80	0,20	11	0.98	0.03	0.00
12	0,50	0,50	12	0.99	0.01	0.00
13	0,05	0,95	13	1.00	0.00	0.00
14	0,51	0,49	14	1.00	0.00	0.00
15	0,99	0,01	15	0.15	0.81	0.04
16	0,99	0,01	16	0.15	0.82	0.03
17	0,99	0,01	17	0.22	0.76	0.01
18	0,99	0,01	18	0.02	0.50	0.48
19	0,99	0,01	19	0.01	0.45	0.54
20	0,99	0,01	20	0.01	0.49	0.51
21	1,00	0,00	21	0.00	0.52	0.48
22	0,99	0,01	22	0.01	0.48	0.51
23	1,00	0,00	23	0.00	0.38	0.62
24	1,00	0,00	24	0.00	0.13	0.87
25	1,00	0,00	25	0.00	0.09	0.91
26	1,00	0,00	26	0.02	0.40	0.58
27	1,00	0,00	27	0.00	0.03	0.97

D.2

D.1















