



OPEN Parasites specific to centipedes form a new major lineage of terrestrial gregarines

Tatiana S. Miroljubova^{1,2}✉, Kirill V. Mikhailov^{3,4}, Timur G. Simdyanov⁵, Vladimir V. Aleoshin^{3,4,5}, Đinh Thế Dũng² & Aleksandra I. Kudriavkina^{1,2}

Gregarines from the families Dactylophoridae and Trichorhynchidae parasitize exclusively centipedes and have a distinct morphology among other terrestrial eugregarines, but their evolutionary relationships have not yet been studied with molecular methods. Here we obtain rDNA operon sequences for the dactylophorids and trichorhynchids. We describe a new species *Trichorhynchus efeykini* sp. n. from a scutigermorph *Thereuopoda longicornis* from Vietnam. Phylogenetic analyses with combined SSU, 5.8S and LSU rDNA dataset support the previously proposed separation of *Trichorhynchus* to the Trichorhynchidae based on morphology and recover the dactylophorids and trichorhynchids as sister groups in a monophyletic clade. This clade appears sister to the clade of the Actinocephaloidea and Stylocephaloidea, and represents a new major lineage of terrestrial gregarines that we designate as a new superfamily Dactylophoroidea.

Terrestrial gregarines are parasites that have been known for almost 200 years¹. They infect a wide range of terrestrial and freshwater invertebrates, including annelids and arthropods². Metagenomic data shows their high abundance in soil communities, which points to their significant yet currently underestimated role in terrestrial ecosystems^{3,4}. However, their evolutionary relationships remain largely unclear and taxonomy is still mainly based on morphological data, since many genera and even families do not have a single sequenced representative. Two of such families are Dactylophoridae Léger, 1892 and Trichorhynchidae Ormières, Marquès & Puisségur, 1977. Recent phylogenies have outlined four lineages of terrestrial gregarines, which have been designated as superfamilies Actinocephaloidea Léger, 1892, Gregarinoidea Labbé, 1899, Stenophoroidea Clopton, 2009, and Stylocephaloidea Clopton, 2009^{5–8}, but it is unclear how dactylophorids and trichorhynchids are related to these superfamilies.

The members of Dactylophoridae and Trichorhynchidae parasitize exclusively centipedes (Chilopoda). The dactylophorids infect Geophilomorpha, Lithobiomorpha, and Scolopendromorpha centipedes, while the only described trichorhynchid, *Trichorhynchus pulcher* Schneider, 1882, infects Scutigermorpha centipedes².

Dactylophorid trophozoites have a remarkable distinctive feature—an asymmetrical and dilated protomerite carrying an attachment apparatus in the form of a number of digitiform or filiform processes—rhizoids, instead of a single epimerite as in the majority of eugregarines^{9–11}. At first, *Trichorhynchus pulcher* was placed in the family Dactylophoridae as a centipede parasite¹². Later, Ormières and colleagues studied the ultrastructure of *T. pulcher* trophozoites^{13,14} and found significant morphological differences from other dactylophorids, including the absence of rhizoids, a lobed epimerite anchored close to the basal lamina of the host intestinal epithelium, a well-developed neck—an elongated anterior end of protomerite deeply immersed in the host intestinal epithelium, and the presence of multiple cytopilia on the cell surface instead of a typical eugregarine epicyte (multiple pellicular folds)². Based on this, they separated *Trichorhynchus* into a monotypic family Trichorhynchidae¹³. Yet in addition to the closely related hosts, there is another feature bringing Trichorhynchidae closer to Dactylophoridae: they have similar (but not identical) mechanisms of the gametocyst dehiscence involving a dilating cytoplasmic residuum, the pseudocyst^{13,15}. In order to determine the relationship between the Dactylophoridae and Trichorhynchidae and their position among other gregarines, we obtained rDNA operon sequences of the dactylophorids *Echinomera hispida* (Schneider, 1876) Labbé, 1899 and *Grebnickiella*

¹Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninsky Ave. 33, Moscow, Russian Federation 119071. ²Joint Vietnam-Russia Tropical Science and Technology Research Center, Hanoi, Vietnam. ³Belozersky Institute for Physico-Chemical Biology, Lomonosov Moscow State University, Leninskiye Gory 1, bldg. 40, Moscow, Russian Federation 119991. ⁴Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Bolshoy Karetny Ln. 19, bldg. 1, Moscow, Russian Federation 127051. ⁵Faculty of Biology, Lomonosov Moscow State University, Leninskiye Gory 1, bldg. 12, Moscow, Russian Federation 119234. ✉email: provorosенок@gmail.com

gracilis (Grebnicki, 1873) Bhatia, 1938 and a trichorhynchid *Trichorhynchus efeykini* sp. n. described here, and reconstructed a phylogeny of terrestrial gregarines.

Results

Hundreds of trophozoites of *Trichorhynchus efeykini* sp. n. were found in the intestine of one individual of *Thereuopoda longicornis* Fabricius, 1793. Narrowly obpyriform gregarines had a relatively large protomerite (the protomerite length to total length ratio was 1:2.5 to 1:3) and a deutomerite of the same width as the protomerite near the septum (Fig. 1A–G). There was a slight constriction along the septum (Fig. 1D,G). Trophozoites had an attachment apparatus in the form of a lobed epimerite located at the anterior end of an elongated neck and anchored close to the basal lamina of the host intestinal epithelium (Fig. 1B,E,F, and Supplementary fig. S1). Parasites had a spherical nucleus with a large eccentrically placed nucleolus, with the nucleolus to nucleus diameter ratio of 1:2 (Fig. 1D). The nucleus had variable position among individuals, and could be observed in any part of the deutomerite. The cell surface was covered with multiple cytopilia that were absent at the front part of protomerite adjacent to the host intestinal epithelium (Fig. 1G–J). Detached gregarines demonstrated gliding and sometimes bending motility. Detailed measurements of gregarines are given in Table 1.

Three out of four collected individuals of *Scolopendra cingulata* were infected with 3–100 trophozoites of *Grebnickiella gracilis*. Nine out of nineteen individuals of *Lithobius forficatus* from Timiryazevsky Park were infected with 1–50 trophozoites of *Echinomera hispida* (TP isolate). Two out of five individuals of *L. forficatus* from the Zvenigorod Biological Station were infected with 1–7 trophozoites of *E. hispida* (ZBS isolate). The TP and ZBS isolates of *E. hispida* had an identical morphology. The trophozoites of *Grebnickiella gracilis* and *Echinomera hispida* had an appearance typical for the species, with asymmetrical protomerite and epicytic folds on the cell surface (Fig. 2).

Near-complete sequences of the rRNA operon (SSU rDNA, ITS1, 5.8S rDNA, ITS2, and LSU rDNA) were obtained for *Echinomera hispida* ZBS isolate (5522 bp), *Grebnickiella gracilis* (5145 bp), and *Trichorhynchus efeykini* sp. n. (5563 bp); GenBank accessions: PP657469, PP654190, PP657470. Partial sequence of SSU rDNA (1595 bp) was obtained for *Echinomera hispida* TP isolate; GenBank accession: PP659854.

Echinomera hispida TP and ZBS isolates differed slightly in the SSU rDNA sequences (the identity percentage was 99.62%). In order to determine the degree of their relationship, we obtained their ITS2 sequences and searched for compensatory base changes in the secondary structures of the ITS2. The presence of compensatory base changes in the secondary structures of ITS2 regions is a criterion for separating cryptic species^{16–19}. We identified a total of four ribotypes in *E. hispida*: two in the TP isolate and two in the ZBS isolate (Fig. 3); GenBank accessions: PP659855, PP659857, PP657469, PP659856. The secondary structure of the ITS2 region of *E. hispida* isolates consists of three helices, similarly to some other sporozoans^{8,20}, and the second helix has a typical U-U mismatch²¹. Polymorphism in the ZBS isolates could only be searched from nucleotide position 120 to the end of the ITS2 due to low coverage in the transcriptomic data. The handful of differences found in the secondary structure of the ITS2 regions did not represent compensatory base changes, indicating both TP and ZBS isolates belong to the same species.

The Bayesian (Fig. 4) and ML (Supplementary fig. S2) trees for the rDNA concatenate had identical topologies and showed the monophyly of major terrestrial and marine eugregarine groups: Actinocephaloidea, Gregarinoidea, Stenophoroidea, Stylocephaloidea, Ancoroidea, Cephaloidophoroidea, and Lecudinoidea, although some of them received low support. The gregarines from centipedes obtained in this study formed a fully supported (posterior probabilities, PP=1 and bootstrap percentages, BP=100) monophyletic group together with unidentified gregarines whose sequences were obtained from either environmental sequencing data or centipede transcriptomic data contaminated with gregarine parasites (see M&M). The dactylophorids (Dactylophoridae) formed a compact and fully supported clade (PP=1, BP=100) sister to the Trichorhynchidae clade (PP=1, BP=100) that included *Trichorhynchus efeykini* and another gregarine from a scutigermorph *Scutigera weberi*. We designate this highly supported union of dactylophorids and trichorhynchids as a new superfamily—the Dactylophoroidea. In our phylogenies the Dactylophoroidea were sister to the clade uniting Actinocephaloidea and Stylocephaloidea with high to moderate support (PP=1, BP=85).

Discussion

The number of gregarine species described from centipedes is less than a third of that from millipedes². This disparity might be an indication of greater morphological similarity of gregarines from centipedes, which might mask their true diversity, but it might also be simply due to the fact that centipedes are more difficult to catch and identify than millipedes, leading to a bias in their representation. Thus far, *Trichorhynchus pulcher* was the only described member of the Trichorhynchidae, with several reported findings being attributed to the same species. This likely is an underestimate of the true diversity in the family, since the reported findings of *T. pulcher* demonstrate some morphological differences. Several authors describe and depict the gregarines with ovoid nuclei^{13,22,23}. Hoshida²⁴ does not describe the appearance of *Trichorhynchus* gregarines found in *Thereuonema chunifera* and *T. tuberculata* in Japan, but according to the Fig. 256 in the article, the author observed a *Trichorhynchus* sp. with apparently a round nucleus. The gregarines found in *Scutigera coleoptrata* in Slovenia also had apparently round nuclei (Fig. 1–3 in²⁵). The gregarine described in the current study was assigned to the genus *Trichorhynchus* because it shares the key features with *Trichorhynchus pulcher*, such as a lobed epimerite anchored close to the basal lamina of the host epithelium, a well-developed neck—an elongated anterior end of protomerite deeply immersed in the host intestinal epithelium, cell surface covered with multiple cytopilia, and both parasites infect scutigermorph centipedes. However, we classified *Trichorhynchus efeykini* sp. n. as a new species because its appearance differs from the previously described *T. pulcher*, including the shape of trophozoites and the protomerite length to total length ratios: 1:4 to 1:7 for *T. pulcher*²² versus 1:2.5 to 1:3 for

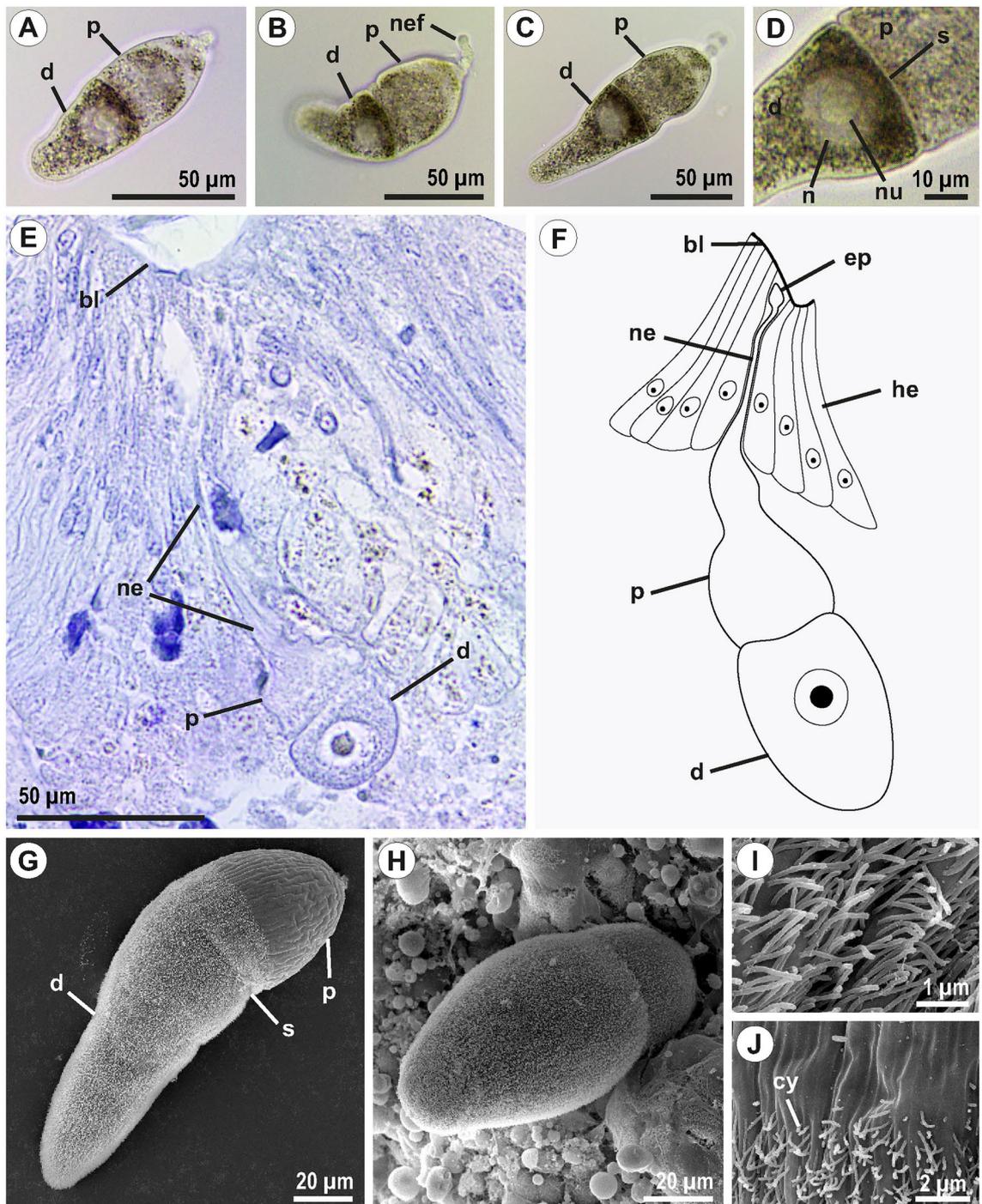


Fig. 1. Morphology of *Trichorhynchus efeykini* sp. n. (A, B) young trophozoite; (C) trophozoite; (D) middle part of a trophozoite cell; (E) histological section of *Thereuopoda longicornis* intestinal epithelium with a *T. efeykini* trophozoite attached; (F) schematic drawing of an attached trophozoite reconstructed from a series of histological sections (see Supplementary fig. S1); (G) trophozoite, SEM; (H) attached trophozoite; (I) cell surface with numerous cytopilia; (J) the area of protomerite surface where the cytopilic carpet ends. bl—basal lamina; cy—cytopilia; d—deutomerite; ep—epimerite; he—host entherocyte; n—nucleus; nu—nucleolus; ne—neck; nef—neck fragment; p—protomerite; s—septum;

T. efeykini. Although molecular data are still lacking to adequately assess the diversity of trichorhynchids, we can assume that gregarine parasites of scutigermorphs may differ greatly, since the rDNA sequences assembled here for *T. efeykini* (5563 bp) and another putative *Trichorhynchus* species found in the transcriptomic data of a South African scutigermorph centipede *Scutigera weberi* (5588 bp) display only 90% identity across

<i>Trichorhynchus efeykini</i> , trophozoites							
n = 24	TL	PL	DL	PWM	PWE	DWM	DWE
Mean	149.7	53.9	95.8	50.4	48.4	48.8	39.85
Standard deviation	44.6	14.6	34.3	10.9	10.7	11.8	12.4
Standard error of the mean	9.1	3.0	7.0	2.2	2.2	2.4	2.5
Max	242.9	90.1	157	69.4	68.8	69	61.8
Min	85	37.5	46.8	32.4	29	29.5	23.2

Table 1. Measurements of *Trichorhynchus efeykini* (in μm). TL—total length; PL—length of protomerite; DL—length of deutomerite; PWM—maximum width of protomerite; PWE—width of protomerite at equatorial axis; DWM—maximum width of deutomerite; DWE—width of deutomerite at equatorial axis; n—number of individuals.

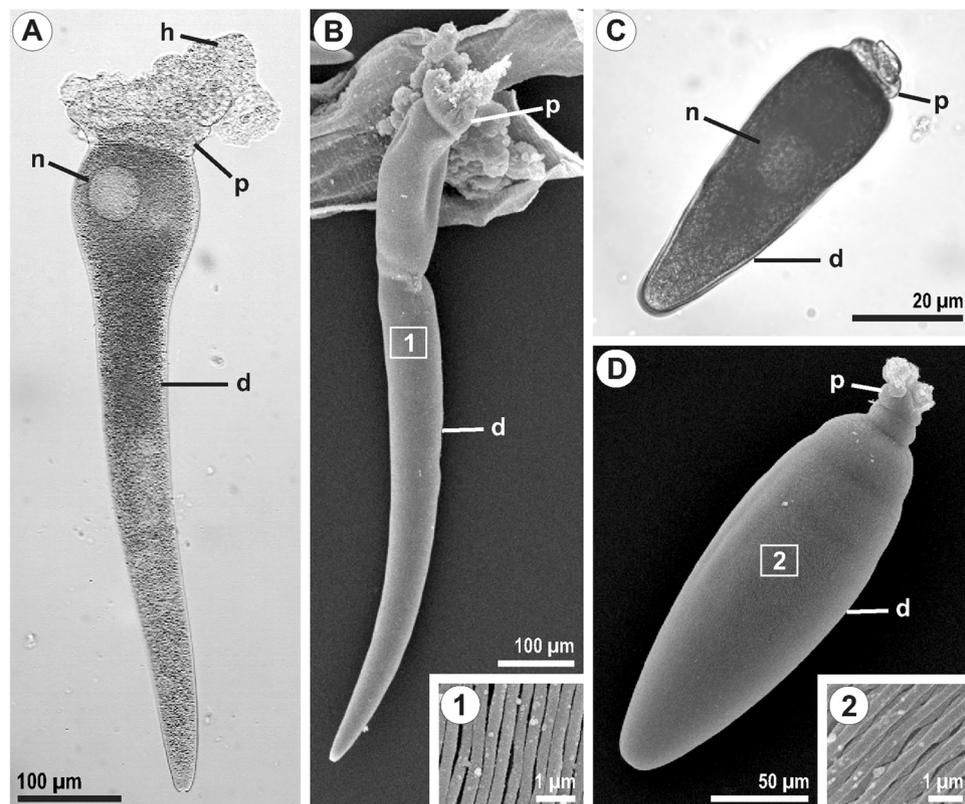


Fig. 2. Morphology of *Grebnickiella gracilis* and *Echinomera hispida*. (A, B) trophozoite of *G. gracilis*; inset 1—epicyclic folds of *G. gracilis*; (C, D) trophozoite of *E. hispida*, isolate TP; inset 2—epicyclic folds of *E. hispida*. d—deutomerite; h—host gut epithelium residuum attached to rhizoids; n—nucleus; p—protomerite;

the whole operon. Therefore, more sequenced isolates from different locations and host species are needed to elucidate the true diversity and distribution of trichorhynchids.

In our phylogenies (Fig. 4, Supplementary fig. S2) the trichorhynchids formed a fully supported group that branched sister to the dactylophorids *E. hispida*, *G. gracilis* and related sequences. The trichorhynchid branch diverged closer to the root of the Trichorhynchidae + Dactylophoridae lineage, with both dactylophorids and trichorhynchids forming relatively tight clusters of more closely-related sequences. Thus, our molecular data support the separation of *Trichorhynchus* from the Dactylophoridae proposed by Ormières and colleagues on the basis of morphological characteristics¹³.

Nevertheless, the fully supported sister relationship between Trichorhynchidae and Dactylophoridae (Fig. 4), an evolutionary lineage of exclusive parasites of centipedes, deserves the taxonomical rank of superfamily, similarly to other major lineages of terrestrial gregarines^{5,6}. We propose the name Dactylophoroidea after the first described family Dactylophoridae. The Dactylophoroidea have radiated through carnivorous hosts—centipedes, which may have conditioned the evolution of their well-developed attachment apparatus: protoepimerite²—a lobate epimerite and a well-developed protomerite neck deeply immersed in the host intestinal epithelium in trichorhynchids, and rhizoids in dactylophorids (Fig. 1, Supplementary fig. S1,^{2,9,11,13}). We can cautiously

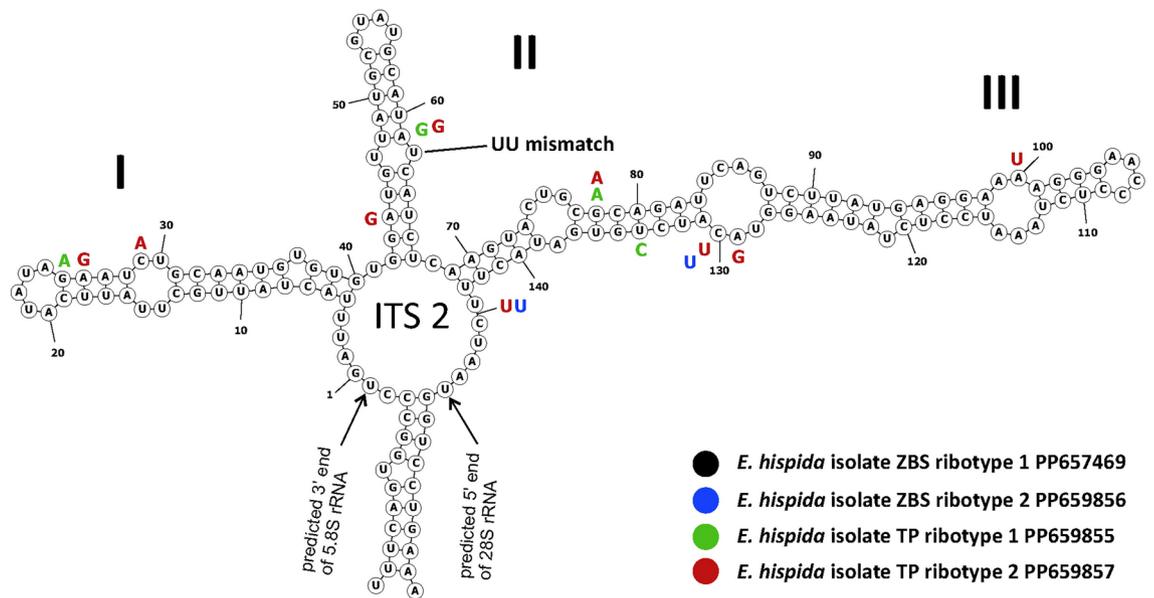


Fig. 3. Predicted secondary structure of the ITS2 transcripts of *Echinomera hispida* demonstrating differences between the four ribotypes (in color). Helices are numbered I–III.

assume that the syzygy in Dactylophoroidea should form late, since late frontal syzygies have been observed in some dactylophorids^{9,26,27} and no syzygies have been reported in trichorhynchids to date. The Dactylophoroidea have radiated through hosts most commonly found in moist environments. This is reflected in their method of the gametocyst dehiscence that occurs with the help of a pseudocyst, a cytoplasmic residuum left over from sporogony. The pseudocyst dilates when exposed to a humid atmosphere. In the dactylophorids the pseudocyst causes the rupture of the hyaline gametocyst wall, while in the trichorhynchids there is no hyaline wall and the pseudocyst opens the gametocyst along the dark equatorial ring^{2,13,15}. Oocysts are reported cylindrical in the dactylophorid and trichorhynchid species studied in this respect^{9,14,15,26–28}.

Ribosomal DNA phylogeny shows that gregarines from millipedes and centipedes form different lineages that are not sister to each other (Fig. 4). All the gregarine sequences found in millipede transcriptomes and the described gregarines from millipedes sequenced to date are placed in the Stenophoroidea clade that is sister to the Gregarinoidea^{7,29}. The centipede-specific parasites—dactylophorids and trichorhynchids are more closely related to the Stylocephaloidea and Actinocephaloidea. The latter superfamily is also known to include gregarines from centipedes in addition to parasites of other invertebrates^{10,24}. The gregarines sequenced by Yazaki and colleagues³⁰ confirm this observation, since the scolopendromorph *Scolopocryptops rubiginosus* parasite misidentified as *Stenophora* sp. (“gregarine ex *Scolopocryptops rubiginosus*” in our tree) was in fact a dactylophorid closely related to *Echinomera*, and the two gregarines from millipedes *Orthomorpha* sp. and *Parafontaria laminata armigera* both grouped within the Stenophoroidea (Fig. 4). The closer relationship between Dactylophoroidea and Actinocephaloidea is also confirmed by multiprotein phylogenies^{30,31} where the dactylophorid misidentified as *Stenophora* sp. grouped with actinocephaloids *Monocystis agilis*, *Ascogregarina taiwanensis* and a gregarine from *Teleopsis dalmanni* with a full support.

Based on all of the above, we propose to establish a new superfamily Dactylophoroidea that includes two families: Dactylophoridae and Trichorhynchidae.

Taxonomy

Phylum Apicomplexa Levine, 1970
 Subphylum Sporozoa Leuckart, 1879
 Class Gregarinomorpha Grassé, 1953
 Order Eugregarinida Léger, 1900
 Superfamily Dactylophoroidea superfam. nov.

Syzygy late frontal in dactylophorids^{9,26,27}, unknown in trichorhynchids. Gametocyst dehiscence by pseudocyst^{2,13,15}. Oocysts cylindrical^{9,14,15,26–28}. In centipedes.

Family Dactylophoridae Léger, 1892
 Family Trichorhynchidae Ormières, Marquès & Puisségur, 1977
Trichorhynchus Schneider, 1882
Trichorhynchus efeykini Miroljubova et Kudriavkina sp. n.

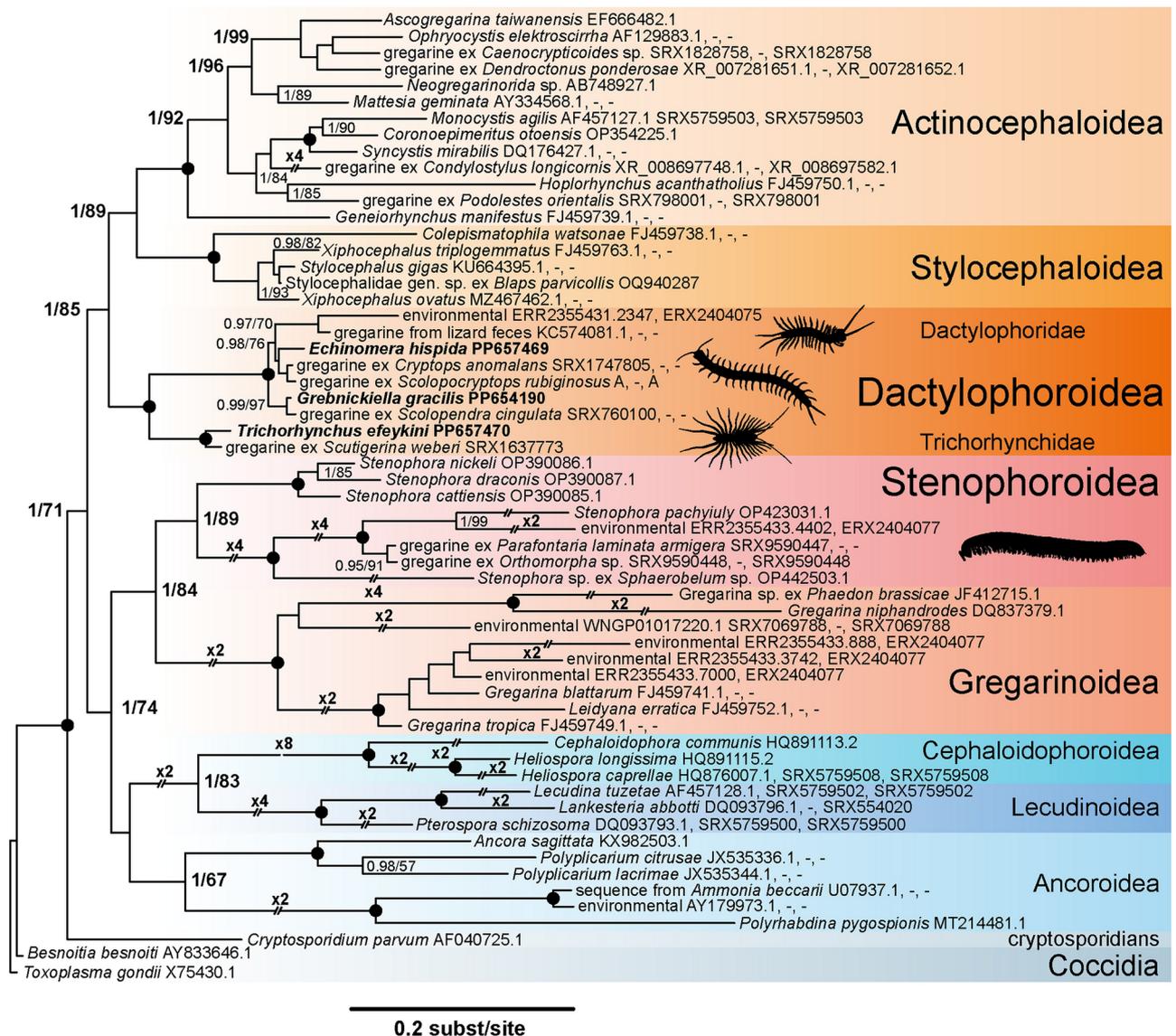


Fig. 4. Bayesian inference tree of eugregarines with an alignment of 59 concatenated SSU, 5.8S and LSU rDNA sequences (4329 sites). Tree node support values: Bayesian posterior probabilities (PP, left) and ML bootstrap percentage (BP, right). Black dots at the tree nodes indicate PP = 1 and BP = 100%. Support values of PP < 0.95 and BP < 70% are omitted. Oblique transverse lines on the branches mark how many times the latter were shortened for the illustration. The newly obtained sequences are in bold. GenBank accession numbers for SSU, 5.8S and LSU rDNAs are separated by comma. Solid contigs of rDNA operon are marked with only one accession number. Sequences assembled from the available transcriptomic data are marked with SRA numbers or A if accession is not in SRA (see M&M). Absent sequences are marked with -.

Trophozoites are narrowly obpyriform, $149.7 \pm 9.1 \mu\text{m}$ long (min. $85 \mu\text{m}$ and max. $242.9 \mu\text{m}$) and $48.8 \pm 2.4 \mu\text{m}$ wide (min. $29.5 \mu\text{m}$ and max. $69 \mu\text{m}$); the protomerite length to total length ratio is 1:2.5 to 1:3; deutomerite of the same width as protomerite near the septum; a lobed epimerite at the anterior end of an elongated protomerite neck; nucleus spherical, in any part of the deutomerite; nucleolus eccentrically placed; the nucleolus diameter to the nucleus diameter ratio is 1:2; cell surface with multiple cytopilia.

DNA SEQUENCE. GenBank PP657470.

TYPE LOCALITY. Hòn Giao mount, Khánh Hòa, Vietnam ($12^\circ 13' 10''$ N $108^\circ 43' 5''$ E).

TYPE HABITAT. Terrestrial.

TYPE HOST. *Thereuopoda longicornis* Fabricius, 1793 (Chilopoda: Scutigeraomorpha).

LOCATION IN HOST. Intestine.

TYPE (syntype) MATERIAL. A gold sputter-coated SEM stub with host gut pieces with attached and detached gregarines, 9 slides with histological sections stained with Ehrlich's hematoxylin, a specimen of parasite cells and host material (intestine) fixed in 96% ethanol, and the host body (without viscera) fixed in 96% ethanol have been deposited in the collection of The Center for Parasitology IPEE RAS; Fig. 1 (this publication) shows some of the syntypes.

LSID: urn:lsid:zoobank.org:act:F975009C-2150-455D-B2AA-0DF15A850482.

ETYMOLOGY: named in honor of Boris Efeykin (junior researcher, IPEE RAS), who collected the infected individual of *Thereuopoda longicornis*.

Materials and methods

Four individuals of *Scolopendra cingulata* Latreille, 1829 were collected in Koktebel (44°57' 14.1" N 35° 14' 09.5" E) in May 2021. Nineteen individuals of *Lithobius forficatus* Linnaeus, 1758 were collected from Timiryazevsky Park in Moscow (55° 48' 29.7" N 37° 33' 04.4" E) in May–July 2021. Five individuals of *L. forficatus* were collected near the Zvenigorod Biological Station of Moscow State University, Moscow region (55° 41' 53.9" N 36° 43' 14.2" E) in July 2021. One individual of *Thereuopoda longicornis* Fabricius, 1793 (Scutigermorpha) was collected on the Hòn Giao mount, Khánh Hòa, Vietnam (12°13' 10" N 108°43' 5" E) in June 2023.

Centipedes were dissected in the insect-adapted Ringer's solution (0.650% NaCl, 0.025% KCl, 0.025% CaCl₂, 0.025% NaHCO₃). Gregarine parasites were isolated from the host's intestines with fine tip needles and plastic pipettes under MBS-1 and MBS-10 stereomicroscopes (LOMO, Russia). For electron microscopy, small fragments of the host intestine with attached gregarines and free gregarines were fixed with 1% glutaraldehyde (v/v) in 0.05 M cacodylate buffer, pH 7.35, final osmolarity 210 mOsm (4 °C, 2 h), then rinsed with the cacodylate buffer and post-fixed with 1% osmium tetroxide (w/v) in the cacodylate buffer (4 °C, 2 h). After fixation, the sample was dehydrated in an ethanol series (15%, 30%, 50%, 70%, 80%, 96%). For histological preparations, fragments of the host intestine with attached gregarines were fixed with Bouin's solution, rinsed two times with distilled water and dehydrated in ethanol series (30%, 50%, 70%, 96%, 100%). Using paraffin-celloidin method 5 µm thick sections were made with the Leica RM-2265 microtome (Leica Microsystems, Germany) and stained with Ehrlich's hematoxylin³².

For molecular phylogenetic studies, all isolated gregarines (forcedly detached trophozoites and gamonts) were washed thrice in fresh portions of the insect-adapted Ringer's solution before fixation/lysis. Twenty-five individuals of each *Echinomera hispida* from Timiryazevsky Park (TP isolate), twenty-five individuals *Grebniackiella gracilis*, and one hundred individuals of *Trichorhynchus efeykini* sp. n. were fixed in 96% ethanol. Seven individuals of *E. hispida* from the Zvenigorod Biological Station (ZBS isolate) were lysed in 1X reaction buffer (SMART-Seq v4 Ultra Low Input RNA Kit, Japan). The samples were then used entirely for DNA or RNA extraction.

Live parasites were photographed with different cameras by means of bright-field microscopy: a DCM35 digital camera (SCOPETEK, PRC) connected to a MBI-3 light microscope (LOMO, Russia) for *G. gracilis*, a Carl Zeiss AxioCam b/w R1.1 412–311 digital camera connected to a Zeiss Axiostar plus light microscope (Carl Zeiss, Germany) for *E. hispida*, and an ADF STD16 digital camera connected to an ADF B30 light microscope (ADF, China) for *Trichorhynchus efeykini* sp. n. The following measurements were taken for *Trichorhynchus efeykini* sp. n.: TL—total length; PL—length of protomerite; DL—length of deutomerite; PWM—maximum width of protomerite; PWE—width of protomerite at equatorial axis; DWM—maximum width of deutomerite; DWE—width of deutomerite at equatorial axis³³.

The gregarines fixed with glutaraldehyde and osmium tetroxide were transferred from 96% ethanol to a mixture of 96% ethanol and acetone (1:1, v/v), followed by three changes of pure acetone and displacement of acetone by liquid carbon dioxide, followed by critical point drying. After that, the samples were mounted on stubs, sputter-coated with gold/palladium, and examined under scanning electron microscopes: a JSM-6380LA (JEOL, Japan) for *G. gracilis* and *E. hispida*, and a Tescan MIRA3 LMU (TESCAN, Czech Republic) for *Trichorhynchus efeykini* sp. n.

For *E. hispida* TP isolate and *G. gracilis*, DNA was extracted with WLB + lysis buffer and proteinase K³⁴ and further purified with the Diatom DMA Prep 100 kit (Galart-Diagnostikum, Russia). The rDNA fragments (18S rDNA, ITS1, 5.8S rDNA, ITS2, 28S rDNA) were amplified with Encyclo PCR kit (Eurogen, Russia) in a total volume of 20 µl in a thermal cycler Applied Biosystems MiniAmp (Thermo Scientific, USA).

18S rDNA were amplified using forward 5'-GTATCTGGTTGATCCTGCCAGT-3' and reverse 5'-GGAAA CCTTGTACGACTTCTC-3' primers. ITS2 regions were amplified using forward 5'-GTACACACCGCCCG TCGCTC-3' and reverse 5'-GACTCCTTGGTCCGTGTTTCAAGACG-3' primers. 28S rDNA was amplified using two pairs of primers: forward 5'-GACCCGCTGAAAYTTAAGCATAT-3' and reverse 5'-ACATTCAGAG CACTGGGCAG-3', and forward 5'-GTCACTTCGGGAWAAGGATTGGCT-3' and reverse 5'-TTCTGACTT AGAGCGTTCAG-3'. PCR protocol: block lid temperature—100 °C, initial denaturation 95 °C—2 min 30 s, denaturation 95 °C—30 s, annealing 60–65 °C—30 s, elongation 72 °C—2 min 30 s; steps 2–4 were repeated 44 times, final extension 72 °C—10 min.

For *Trichorhynchus efeykini* sp. n., total RNA was extracted with RNAqueous-Micro Total RNA Isolation Kit (Invitrogen, USA). For *E. hispida* ZBS isolate and *Trichorhynchus efeykini* sp. n., the cDNA libraries were prepared using the SMART-Seq v4 Ultra Low Input RNA Kit (Takara Bio, Japan). 15.5 million paired-end 100 bp reads for *E. hispida* ZBS isolate were generated with Illumina HiSeq 4000. 79.7 million paired-end 150 bp reads for *Trichorhynchus efeykini* sp. n. were generated with Illumina NovaSeq 6000. The rDNA operon sequences (SSU rDNA, ITS1, 5.8S rDNA, ITS2, LSU rDNA) of *E. hispida* ZBS isolate and *Trichorhynchus efeykini* sp. n. were assembled from the transcriptomic data using NOVOPlasty assembler³⁵. Ribotypes were identified in the rDNA assemblies by mapping library reads with Bowtie2³⁶.

In order to improve the sampling of gregarines related to the known dactylophorids, we performed BLAST searches in the nr database (NCBI) and available metagenomic data (NCBI Sequence Read Archive, accessions: PRJEB25197) for environmental sequences related to these gregarines. Further, we searched for gregarine contamination in the NCBI Sequence Read Archive available transcriptomic data of different centipedes. The contamination was found and gregarine rDNA sequences were extracted from three libraries: *Cryptops anomalans* (SRX1747805), *Scolopendra cingulata* (SRX760100), and *Scutigera weberi* (SRX1637773). In order to increase the number of terrestrial gregarine operational taxonomic units (OTUs) that contain 5.8S and LSU sequences, we extracted rDNA sequences from available gregarine transcriptomic data: SRX9590447, SRX9590448, SRX5759500, SRX554020, SRX5759502, SRX5759508, and <https://doi.org/https://doi.org/10.5061/dryad.6m905qfz6> for a gregarine from a centipede *Scolopocryptops rubiginosus*. In addition, we extracted gregarine rDNA sequences from available metagenomic data (SRX7069788) and transcriptomic data of two insect hosts (SRX7069788 and SRX1828758).

Three rDNA alignments (the SSU, 5.8S, and LSU) with different representations of apicomplexans (222, 107, and 102 OTUs respectively) were generated using the MAFFT online service employing the E-INS-i alignment method³⁷. Poorly aligned regions and columns containing few nucleotides were removed from the alignments using a custom mask resulting in three alignments: 1,530 bp for the SSU, 145 bp for the 5.8S, and 2654 bp for the LSU rDNA. A total of 59 OTUs, including eugregarines, related environmental sequences, a cryptosporidian, and coccidians as an outgroup were selected for phylogenetic inference; their rDNAs were concatenated resulting in a 4,329 nucleotide site alignment.

Bayesian inference (BI) was performed with MrBayes 3.2.6³⁸ utilizing the resources of the CIPRES web server³⁹. The analysis was performed under the GTR + Γ + I model with 8 rate categories for the SSU and LSU rDNA partitions, and with 4 rate categories for the 5.8S rDNA partition. The chain heating coefficient (temp) was set to 0.2. The inference was done with two independent runs of four MCMC, and the consensus tree was built with a 50% burn-in after 10 million generations, and the tree sampling frequency of 0.001. The average standard deviation of split frequencies at the end of computations was 0.002945. Maximum likelihood (ML) analysis was performed with the IQ-TREE 2.1.2⁴⁰ using the GTR + I + G8 + F model for the SSU and LSU rDNA partitions and the GTR + I + G4 + F for the 5.8S rDNA partition. The computations were performed with 1000 nonparametric bootstrap⁴¹ replicates for estimation of branch support. The Bayesian and ML consensus trees had an identical topology; the ML support values were superimposed on the Bayesian tree for the illustration.

The predicted secondary structures of ITS2 sequences of *E. hispida* isolates were generated with the RNAfold web server (<http://rna.tbi.univie.ac.at/>) at the temperature setting of 37 °C under default parameters. To determine whether *E. hispida* isolates belong to different or the same species, compensatory base changes in the secondary structures of the ITS2 sequences were searched^{16–19}.

Data availability

Sequence data that support the findings of this study have been deposited in the GenBank with the primary accession codes: PP657469 (*Echinomera hispida* ZBS isolate SSU rDNA, ITS1, 5.8S rDNA, ITS2, and LSU rDNA), PP654190 (*Grebniackiella gracilis* SSU rDNA, ITS1, 5.8S rDNA, ITS2, and LSU rDNA), PP657470 (*Trichorhynchus efeykini* sp. n. SSU rDNA, ITS1, 5.8S rDNA, ITS2, and LSU rDNA); PP659854 (*Echinomera hispida* TP isolate SSU rDNA); PP659855 (*Echinomera hispida* TP isolate ribotype 1 ITS2), PP659857 (*Echinomera hispida* TP isolate ribotype 2 ITS2), PP659856 (*Echinomera hispida* ZBS isolate ribotype 2 ITS2). The datasets generated and analyzed during the study are available from the corresponding author on request.

Received: 19 April 2024; Accepted: 18 December 2024

Published online: 02 January 2025

References

- Dufour, L. Note sur la Grégarine, nouveau genre de ver qui vit en troupeau dans les intestins de divers insectes. *Ann. Sci. Nat.* **13**, 366–378 (1828).
- Desportes, I. & Schrével, J. *The gregarines: the early branching Apicomplexa* (ed. Desportes, I. & Schrével, J.) 1–781 (Brill, Leiden, 2013).
- Jamy, M. et al. Long-read metabarcoding of the eukaryotic rDNA operon to phylogenetically and taxonomically resolve environmental diversity. *Mol. Ecol. Resour.* **20**, 429–443 (2020).
- Mahé, F. et al. Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. *Nat. Ecol. Evol.* **1**, 1–8 (2017).
- Clopton, R. E. Phylogenetic relationships, evolution, and systematic revision of the septate gregarines (Apicomplexa: Eugregarinorida: Septatorina). *Comp. Parasitol.* **76**, 167–190 (2009).
- Clopton, R. E. & Clopton, D. T. Revision of *Coronoepimeritus* with description of *Coronoepimeritus otoensis* n. sp. infecting the admirable grasshopper *Syrbula admirabilis* (Orthoptera: Acrididae). *J. Parasitol.* **108**, 487–499 (2022).
- Miroliubova, T. S. et al. Phylogeny of stenophorid gregarines from millipedes (Diplopoda). *Invert. Zool.* **20**, 125–139 (2023).
- Simdyanov, T. G. et al. A new view on the morphology and phylogeny of eugregarines suggested by the evidence from the gregarine *Ancora sagittata* (Leuckart, 1860) Labbé, 1899 (Apicomplexa: Eugregarinorida). *PeerJ* **5**, e3354 (2017).
- Goodrich, H. P. *Nina*: A remarkable gregarine. *J. Cell Sci.* **2**, 107–126 (1938).
- Ormières, R. Grégarines parasites de Myriapodes Chilopodes: observations sur les genres *Echinomera* Labbé, 1899 et *Acutispora* Crawley, 1903. *Protistologica* **2**, 15–21 (1966).
- Ormières, R. & Marqués, A. Fixation a leurs hotes de quelques Dactylophoridae, eugregarines parasites de myriapodes chilopodes. *Etude ultrastructurale. Protistologica* **12**, 415–424 (1976).
- Léger, L. Recherches sur les Grégarines. in *Tablettes Zoologiques* vol. 3 1–183 (Typographie Oudin et Cie, Poitiers, 1892).
- Ormières, R., Marqués, A. & Puisségur, C. *Trichorhynchus pulcher* Schneider, 1882, eugregarine parasite de *Scutigera coleoptrata* L. Cycle, ultrastructure, systématique. *Protistologica* **13**, 407–417 (1977).
- Marqués, A., Ormières, R. & Puisségur, C. Observations en microscopie électronique a balayage de quelques stades de *Trichorhynchus pulcher* Schneider, 1882, eugregarine parasite de *Scutigera coleoptrata* L. (Myriapodes, Chilopodes). *Ann. Sci. Nat.* **20**, 27–36 (1978).

15. Tuzet, O. & Ormières, R. Sur quelques Grégarines de la région de Sète. *Ann. Parasitol. Hum. Comp.* **31**, 317–330 (1956).
16. Coleman, A. W. The significance of a coincidence between evolutionary landmarks found in mating affinity and a DNA sequence. *Protist* **151**, 1–9 (2000).
17. Coleman, A. W. Is there a molecular key to the level of “biological species” in eukaryotes? A DNA guide. *Mol. Phylogenet. Evol.* **50**, 197–203 (2009).
18. Müller, T., Philippi, N., Dandekar, T., Schultz, J. & Wolf, M. Distinguishing species. *RNA* **13**, 1469–1472 (2007).
19. Wolf, M., Chen, S., Song, J., Ankenbrand, M. & Müller, T. Compensatory base changes in ITS2 secondary structures correlate with the biological species concept despite intragenomic variability in ITS2 sequences—A proof of concept. *PLoS One* **8**, e66726 (2013).
20. Paskerova, G. G., Miroliubova, T. S., Valigurová, A., Aleoshin, V. V. & Simdyanov, T. G. Morphological and Phylogenetic Study of Protococcidians Sheds Light on the Evolution of Epicellular Parasitism in Sporozoa (Apicomplexa), with the Description of *Eleutherschizon planoratum* sp. nov. *Diversity* **15**, 863 (2023).
21. Schultz, J., Maisel, S., Gerlach, D., Müller, T. & Wolf, M. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. *RNA* **11**, 361–364 (2005).
22. Watson, M. E. Studies on gregarines including descriptions of twenty-one new species and a synopsis of the eugregarine records from the Myriapoda, Coleoptera, and Orthoptera of the world. *Ill. Biol. Monogr.* **2**, 1–258 (1916).
23. Geus, A. *Sporentierchen, Sporozoa. Die Gregarinida Der Land- Und Süßwasserbewohnenden Arthropoden Mitteleuropas.* vol. 57 (Gustav Fischer Verlag, Jena, 1969).
24. Hoshide, H. Studies on the cephaline Gregarines of Japan. II.3. Description of the members belonging to the families Didymophyidae, Actinocephalidae, Acanthosporidae, Stylocephalidae, Dactylophoridae. *Bull. Fac. Educ. Yamaguchi Univ.* **8**, 35–101 (1959).
25. Devetak, D., Mihelak, K. & Kos, I. Gregarines (Apicomplexa: Eugregarinida) of Chilopoda and Diplopoda in Slovenia. *Acta zool. bulg.* **71**, 121–128 (2019).
26. Misra, P. L. On the life-history of a new gregarine, *Grebneckiella pixellae*, sp. nov., from the centipede, *Scolopendra morsitans* Linn., with a note on the family Dactylophoridae Léger, 1892. *Rec. Zool. Surv. India* **44**, 324–337 (1946).
27. Schellack, C. Über die Entwicklung und Fortpflanzung von *Echinomera hispida* (A. Schn.). *Arch. Protistenkd.* **9**, 297–345 (1907).
28. Labbé, A. *Sporozoa* 1–80 (R. Friedländer und Sohn, 1899).
29. Clopton, R. E., States, C. & Clopton, D. T. *Gregarina lutescens* n. sp. Infecting the Harlequin Ladybird *Harmonia axyridis* (Coleoptera: Coccinellidae). *J. Parasitol.* **110**, 66–78 (2024).
30. Yazaki, E. et al. Signs of the plastid: Enzymes involved in plastid-localized metabolic pathways in a eugregarine species. *Parasitol. Int.* **83**, 102364 (2021).
31. Mathur, V. et al. Reconstruction of plastid proteomes of apicomplexans and close relatives reveals the major evolutionary outcomes of cryptic plastids. *Mol. Biol. Evol.* **40**, msad002 (2023).
32. Roskin, G. I. & Levinson, L. B. *Microscopic Techniques 1–467* (Sovetskaya Nauka, 1957).
33. Clopton, R. E. Standard nomenclature and metrics of plane shapes for use in gregarine taxonomy. *Comp. Parasitol.* **71**, 130–140 (2004).
34. Holterman, M. et al. Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown clades. *Mol. Biol. Evol.* **23**, 1792–1800 (2006).
35. Dierckx, N., Mardulyn, P. & Smits, G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **45**, e18 (2017).
36. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
37. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166 (2019).
38. Ronquist, F. et al. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
39. Miller, M. A., Pfeiffer, W. & Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. in *2010 Gateway Computing Environments Workshop (GCE)* 1–8 (2010). <https://doi.org/10.1109/GCE.2010.5676129>.
40. Minh, B. Q. et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
41. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**, 783–791 (1985).

Acknowledgements

The SEM study of *Trichorhynchus efeykini* sp. n. was conducted using the Joint Usage Center “Instrumental methods in ecology” at the IPEE RAS.

Author contributions

T.S.M. and A.I.K. performed sampling, LM, SEM, and histological studies of *Trichorhynchus efeykini*. D.T.D. organized fieldwork in Vietnam. T.G.S. performed sampling and LM study of *Grebneckiella gracilis*. A.I.K. performed sampling and LM study of *Echinomera hispida*. T.G.S. and A.I.K. performed SEM study of *G. gracilis* and *E. hispida*. T.S.M. prepared cDNA libraries. A.I.K. performed DNA extraction and PCR. K.V.M. and V.V.A. obtained rDNA sequences from transcriptomic data. T.S.M., K.V.M., V.V.A., and A.I.K. performed the ITS2 study. T.S.M. performed phylogenetic study. T.S.M. and A.I.K. wrote the manuscript. All authors worked on review and editing.

Funding

The study of *Trichorhynchus efeykini* sp. n. was funded by the Russian Science Foundation (project No. 24-24-00560). The work of author T.G.S. on *G. gracilis* was carried out as part of the Scientific Project of the State Order of the Government of Russian Federation to Lomonosov Moscow State University No. 121032300117-3.

Declarations

Competing interests

The author declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-83990-4>

[0.1038/s41598-024-83990-4](https://doi.org/10.1038/s41598-024-83990-4).

Correspondence and requests for materials should be addressed to T.S.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024