



Research



Cite this article: Sato S, Appeldorff C, Wangenstein OS, Garcés-Pastor S, Laumer CE, Herranz M, Giribet G, Renault D, Rask Møller P, Worsaae K. 2025 Phylogenomics of the rarest animals: a second species of Micrognathozoa identified by machine learning. *Proc. R. Soc. B* **292**: 20242867.

<https://doi.org/10.1098/rspb.2024.2867>

Received: 29 November 2024

Accepted: 20 January 2025

Subject Category:

Evolution

Subject Areas:

evolution, taxonomy and systematics, molecular biology

Keywords:

meiobenthos, Spiralia, phylotranscriptomics, species delimitation, confocal laser scanning microscopy, integrated taxonomy

Author for correspondence:

Katrine Worsaae

e-mail: kworsaae@bio.ku.dk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.7667300>.

Phylogenomics of the rarest animals: a second species of Micrognathozoa identified by machine learning

Shoyo Sato¹, Cecilie Appeldorff¹, Owen S. Wangenstein³, Sandra Garcés-Pastor^{3,4}, Christopher E. Laumer⁵, María Herranz^{1,6,7}, Gonzalo Giribet⁸, David Renault^{9,10}, Peter Rask Møller² and Katrine Worsaae¹

¹Marine Biological Section, Department of Biology, and ²Natural History Museum of Denmark, University of Copenhagen, Copenhagen 2100, Denmark

³Department of Evolutionary Biology Ecology and Environmental Sciences & Biodiversity Research Institute (IRBio), Universitat de Barcelona, Barcelona, Catalonia 08028, Spain

⁴Institute of Marine Sciences (ICM), CSIC, Barcelona, Catalonia 08003, Spain

⁵Department of Life Sciences, The Natural History Museum, London, UK

⁶Area of Biodiversity and Conservation, Superior School of Experimental Science and Technology (ESCET), Rey Juan Carlos University, Móstoles, Madrid, Spain

⁷Global Change Research Institute (IICG-URJC), Móstoles, Madrid 28933, Spain

⁸Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

⁹UMR CNRS 6553 ECOBIO (Ecosystèmes, biodiversité, évolution), Université Rennes, avenue du Général Leclerc, Rennes cedex 35042, France

¹⁰Institut Universitaire de France, Paris, France

id SS, 0000-0002-2970-8702; OSW, 0000-0001-5593-348X; SG-P, 0000-0001-5652-7264; CEL, 0000-0001-8097-8516; MH, 0000-0002-3020-6072; GG, 0000-0002-5467-8429; DR, 0000-0003-3644-1759; PRM, 0000-0002-0177-0977; KW, 0000-0003-0443-4298

The latest animal phylum to be discovered, Micrognathozoa, constitutes a rare group of limnic meiofauna. These microscopic ‘jaw animals’ are among the smallest metazoans yet possess highly complex jaw structures. The single species of Micrognathozoa, *Limnognathia maerski* Kristensen and Funch, 2000, was first described from Greenland, later reported from a remote Subantarctic island and more recently discovered in the Pyrenees on the European continent. Successful collections of these three known populations facilitated investigations of the intraphylum relationships and species limits within *Limnognathia* for the first time. Through detailed anatomical comparisons, we substantiate the lack of morphological differences between the three geographically disjunct populations. With transcriptomic data from single specimens, we conducted the first intraphylum phylogenetic analyses and extensively tested species hypotheses using standard approaches and novel machine learning methods. Analyses clearly delimited the Subantarctic population, here described as *Limnognathia desmeti* sp. nov., the second species of Micrognathozoa, but did not definitively split the Greenland and Pyrenees populations as separate species. Divergence dating analysis suggests the disjunct distribution of Micrognathozoa is not human mediated but the result of long-distance dispersal raising questions about their dispersal capabilities and potential undiscovered populations.

1. Background

Micrognathozoa (figure 1a) was first described in 2000 from a single species, *Limnognathia maerski* Kristensen and Funch, 2000, from cold springs on Disko Island, Greenland [1]. An affinity with Gnathostomulida and particularly

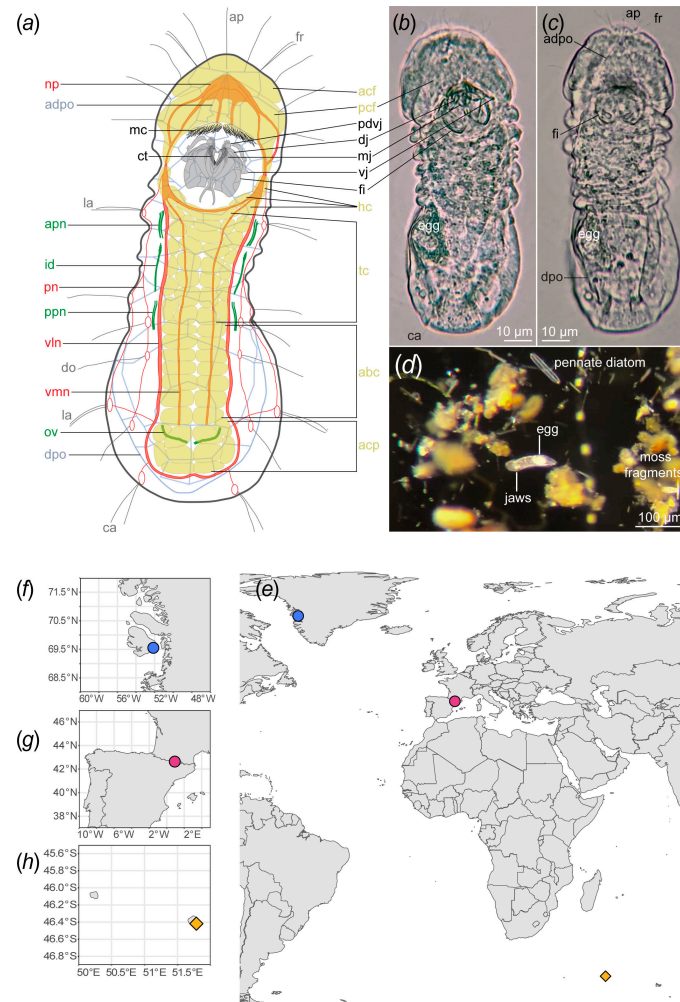


Figure 1. Distribution and habitus of Micrognathozoa. (a) Schematic of Micrognathozoa showing key morphological characteristics investigated in this study. (b) Ventral and (c) LM of *Limnognathia desmeti* sp. nov. (paratype, NHMD-001801226). (d) Dissecting scope image of *Limnognathia desmeti* sp. nov. among sampled substrate. (e) Map of the three populations of Micrognathozoa sampled for this study. Circles denote populations of *Limnognathia maerski* and diamonds denote *Limnognathia desmeti* sp. nov. (f) Isunngua Spring, Disko Island, Greenland. (g) Bassa Nera, Aiguèstortes i Estany de Sant Maurici NP, Catalunya. (h) Pointe du Bougainville, Île de la Possession, Crozet Archipelago. Colours in the map correspond to the colours used in all other figures. abc, abdominal ciliophores; acf, anterior ciliated field; acp, adhesive ciliary pad; adpo, apical dorsal plate outline; ap, apicalia; apn, anterior protonephridia; ca, caudalia; ct, ciliary tufts; dj, dorsal jaw; do, dorsalia; dpo, dorsal plate outline; egg, egg; fi, fibularium; hc, head ciliophores; id, intermediate duct; la, lateralia; mc, mouth ciliation; mj, main jaw; np, neuropil; ov, oviduct; pcf, posterior ciliated field; pdvj, pseudodigit of the ventral jaw; pn, peripheral nerve; ppn, posterior protonephridia; tc, trunk ciliophores; vln, ventro-lateral nerve; vj, ventral jaw; vmn, ventro-median nerve.

Syndermata (syn. Rotifera *sensu lato*), a grouping called Gnathifera, was soon hypothesized based on morphological similarities [1–6]. Molecular phylogenetic analyses testing the position of Micrognathozoa found that it did not nest within known phyla [7,8] and Micrognathozoa was elevated in status to become the most recently described and only exclusively freshwater animal phylum. Phylogenomic analyses of Spiralia later firmly placed Micrognathozoa as sister phylum to Syndermata [9–11], together called Gynognathifera [12], within the larger clade of Gnathifera. Despite their minuscule size with a maximum length of 150 µm, Micrognathozoa possess the most complex jaw structures in the animal kingdom [1–3]. No males of the species have been found and the phylum has been proposed to reproduce primarily or solely by parthenogenesis [1,8,13].

Soon after the description of the Greenlandic species, Micrognathozoa were reported from Île de la Possession, in the Subantarctic Crozet Archipelago (hereon Crozet) [2]. Despite their distance from Greenland, because of their seemingly identical morphology, the Crozet specimens were also assigned to *L. maerski* [2]. A few additional specimens were later reported from Wales and Berkshire, UK [13], and sequence matches in environmental DNA (eDNA) metabarcoding samples indicated micrognathozoan presence in continental Europe [14,15]. However, a European population was never localized or collected until the recent discovery of Micrognathozoa in the Pyrenees [16]. Currently, a single species is described in the phylum Micrognathozoa. Only data from pooled Greenlandic specimens have been included in molecular phylogenetic analysis (e.g. [7,10,11,17]). Neither the intraphylum relationships among populations nor potential cryptic speciation have been investigated. Through extensive sampling efforts over the last 10 years, we obtained multiple individuals from each of the three known populations of *Limnognathia*: Greenland, Pyrenees and Crozet (figure 1e–h). Those samples are here used in the first phylogenetic analysis of the internal relationships within Micrognathozoa using phylotranscriptomics. The hypothesis of cryptic speciation is tested with advanced microscopy and novel machine learning species delimitation analyses.

We approach integrative species delimitation from species discovery to validation using a range of methods and data types [18]. Species discovery methods, agnostic to species hypotheses, do not require *a priori* labelling of individuals to species [18]. Morphological species discovery is nearly impossible in cryptic, morphologically conserved groups like soft-bodied meiofauna [19,20], resulting in underestimates of true diversity. Common molecular species delimitation methods oversplit and identify highly divergent populations as species, particularly in systems that do not meet model assumptions, lack informed priors and have limited dispersal [21–24].

Unsupervised machine learning (UML) clustering methods are agnostic to population parameters and species divergences. These parameters are difficult to estimate in meiofauna where baseline knowledge is poor compared with well-studied groups. UML methods were not specifically developed for the problem of species delimitation but can reveal underlying structures in the data that could correspond to species [25,26].

Species validation methods explicitly test *a priori* species hypotheses [18]. Common methods that rely on the multi-species coalescent are limited to divergence-only models, which might not reflect the underlying speciation processes [27]. Supervised machine learning (SML) has allowed testing of models using broader assumptions and non-divergence processes [27], and has even inferred unknown species boundaries using knowledge from biologically similar systems [28–30].

Standard integrative methods can produce contradictory results by both over- and under-splitting species numbers [28]. Due to its conserved morphology, high habitat specificity, geographic isolation, lack of precise priors and putative parthenogenesis, Micrognathozoa is an ideal case to test the novel methods of both agnostic (UML) and informed (SML) machine learning.

2. Methods

Detailed methodology can be found in the electronic supplementary material.

(a) Specimen collection

Specimens were collected by examining water squeezed and filtered from aquatic moss samples under a dissecting scope and isolating them with pipettes. Specimens were anaesthetized with MgCl_2 and fixed for morphological study or stored in RNAlater or flash frozen for molecular study.

(b) Morphological study

Light microscopy (LM) images and videos were obtained using a LabCam (iDu optics) mounted iPhone 11. Immunohistochemistry and confocal laser scanning microscopy (CLSM) followed previously described methodology [6] (electronic supplementary material, tables S1 and S2) using an Olympus FV-1000. Z-stacks were analysed with Imaris (Oxford Instrument 2022). Comparisons of the jaws are based on previously published scanning electron microscopy (SEM) images [2,3].

(c) Sequencing

Single-individual transcriptome libraries were generated using the Clontech SMART-Seq HT and Nextera XT DNA kits. Sequencing and demultiplexing was done on an Illumina NovaSeq 6000. Outgroup sequences were obtained from NCBI (electronic supplementary material, table S3 for taxon sampling). A 684 bp reference Sanger cytochrome *c* oxidase subunit I (COI) was sequenced from the cDNA using universal primers (electronic supplementary material, table S4) [31], purified with the E.Z.N.A. Cycle Pure Kit, sequenced at MacroGen Europe and assembled in Geneious [32].

(d) Transcriptome assembly and orthology inference

All transcriptomes (electronic supplementary material, table S3) were assembled and processed following published pipelines [33–39]. Assembly quality was assessed with BUSCO and the Metazoa Orthodb v10 [40,41]. Orthology inference of all transcriptomes followed the PhyloPyPruner (<https://pypi.org/project/phylopypruner/>) pipeline [42–48] to create 50 and 70% occupancy matrices (matrices M1 and M2, respectively).

A 100% occupancy single-nucleotide polymorphism (SNP) (matrix M3) and 18S rRNA datasets were created following GATK ‘best practices’ [49–52]. Unlinked SNP datasets were created with `RANDSNPs_FROM_VCF.PL`. Resulting VCF files were filtered at 50% occupancy (matrix M4) and 100% occupancy (matrices M5–M9), converted to phylip with `VCF2PHYLIP.PY` [53] and processed using phrynomics [54]. COI sequences were extracted with MitoGeneExtractor [55] using the generated Sanger reference and aligned with MAFFT [44]. Poorly aligned bases were trimmed (alignment = 612 bp).

A 100% occupancy matrix of Metazoa_odb v10 BUSCO loci (matrix M10) was extracted from *Limnognathia* assemblies with TOAST [56], aligned with MAFFT and processed with Gblocks [57,58] to obtain single-copy nucleotide orthologues. Information of all matrices used in this study can be found in electronic supplementary material, table S5.

(e) Phylogenetic analyses

Maximum likelihood phylogenetic analyses were conducted in IQ-TREE [59–63]. Bayesian analyses were run in ExaBayes [64]. Log files were combined with LogCombiner [65] and convergence was checked in Tracer [66].

Divergence dating analysis was conducted in BEAST [67] using the trimmed COI and 18S rRNA alignments. Substitution models were determined with ModelFinder in IQ-TREE. Universal COI (1.76%) [68] and 18S rRNA (0.02%) [69] substitution rates were used. Convergence was checked in Tracer. A maximum clade credibility tree was produced with TreeAnnotator [67].

(f) Species delimitation

PTP and bPTP [70] were run using the maximum likelihood tree of matrix M1 as input. GMYC [71] was run using the ultrametric BEAST tree from divergence time estimation.

Data in matrix M4 were clustered with principal component analysis (PCA), random forest (RF) and t-distributed stochastic neighbour embedding (tSNE) [25,72–81]. Variational autoencoder (VAE) analysis [82] was run using matrices M3 and M5–M9. Phylip files were one-hot encoded using a custom python script ONEHOTENCODE.PY.

Matrix M4 was analysed with SNAPPER [83]. The Yule model lambda parameter was estimated with Pyule [84]. Two xml files were created with BEAUti [67] corresponding to two and three species models. Marginal likelihood estimates were calculated with PathSampleAnalyser, and models were selected with Bayes factor (BF).

DelimitR [27] was run using an MSFS file that was created from matrix M4 with easySFS [85]. CLADES analysis [29] followed Derkarabetian *et al.* [28] with matrix M10 using the ‘general’ model provided by Pei *et al.* [29] and a custom model created from a rotifer transcriptome dataset (electronic supplementary material, figure S1 and Table S6) using Libsvm [86]. Admixture analysis [87] was conducted on matrix M4.

(g) Molecular species description

Pairwise COI distances between all specimens were computed in MEGA [88] under the K2P model [89]. The 18S rRNA and COI alignments were visualized in Geneious [32] to identify fixed molecular differences between species.

3. Results and discussion

(a) Morphological observations and comparison

We found no significant differences using advanced microscopic techniques between Northern Hemisphere (Greenland and Pyrenees) and Southern Hemisphere (Crozet) individuals (electronic supplementary material, table S7).

Specimens from all populations did not differ in total body length and had three main body divisions (figure 1a–d). No males were found in any of the explored localities. The three previous detailed studies on the complex jaw morphology (figure 1a–c) revealed no differences between Greenlandic and Crozet *Limnognathia* except for a minor, non-diagnostic difference in the number of ‘teeth’ of the pseudodigits [*sensu* 1, 3] or pseuduncus [*sensu* 2] (pdvj, figure 1a). However, this seemingly reflects intraspecific variation or preservation differences since both four [1] and five [3] ‘teeth’ have been noted in both populations [2].

The dorsal and lateral epidermal cells of micrognathozoans lack a cuticle and are instead supported by an intracellular matrix (plates), first observed using LM. All dorsal plates of Greenlandic specimens were originally noted to be composed of 3–4 cells except for the single giant cell of the apical plate [1]. However, in the apical plate of Crozet specimens, De Smet distinguished four epidermal cells [2]. Later studies of Greenlandic specimens found a more exact number, shape and configuration of epidermal plates with CLSM [5,6] and that plate number seemingly equals cell number. We confirmed these results through reanalysis of these CLSM scans and further investigation of Crozet and Pyrenees specimens using immunohistochemistry and CLSM. Consistently across multiple protocols, unspecific accumulation (not staining) of background dye along cell borders and junctions allowed us to distinguish ~60 lateral and dorsal epidermal cells that are identical across individuals from all populations (figure 1a). The ‘apical plate’ always comprises four epidermal plates/cells (1–4 in figure 2a,d) with 2–4 detectable nuclei. The outside cell borders (adpo, figures 1ac; 2a,d) are clearly delimited by LM and immunostaining while the internal boundaries between the four cells are primarily detectable with increased intensity and magnification in CLSM (figure 2A and D).

Tubulin staining and CLSM revealed no differences in ciliation among Greenland, Pyrenees and Crozet specimens. The ventral ciliophores of Pyrenees and Crozet individuals (acf, pcf, hc, tc, abc, acp in figure 1a) exactly match the updated arrangement and number of Greenlandic specimens [6], which differs markedly from the original description using LM [1]. The form and number of sensory collar receptors (ap, fr, la, do, ca in figures 1a–c and 2a,b,g) corroborate previous descriptions [6] and is here clearly shown to originate between the epidermal plates (dob, figure 2a). The presence of the pharyngeal ciliary tuft [6] is confirmed in this study for both Greenland and Crozet specimens. Internal ciliary structures (i.e. oviducts, protonephridia) are also identical among individuals from the three populations down to the number and position of the terminal cells (figure 2e).

The nervous system of Crozet specimens, including the peripheral nerves (figure 1a) is identical to that described previously for Greenland specimens [6]. The presence of a pharyngeal ganglion [6], with possible homology to the buccal ganglion or group of somata in rotifers, gnathostomulids and chaetognaths [12,90], is confirmed in Crozet specimens.

Due to the consistent morphological findings across specimens from the three populations regarding jaws, ciliation, epidermis and nervous system, those traits uncovered since the original species description of *L. maerski* are here amended to the genus diagnosis.

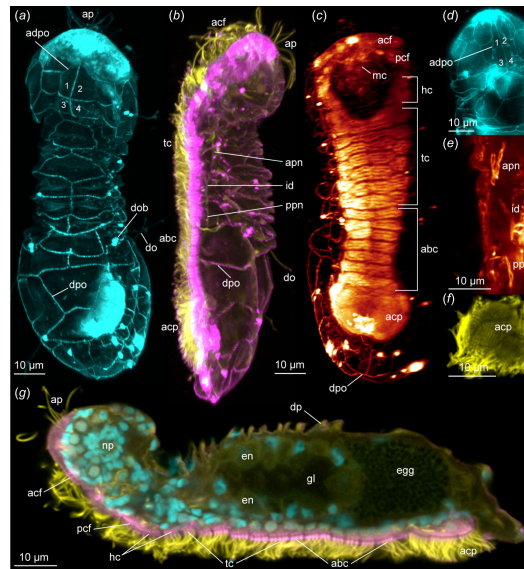


Figure 2. CLSM of *Limnognathia*. All images maximum intensity z-projection. Anterior end up on (a)–(f), and left on (g). (a) Dorsal *Limnognathia desmeti* sp. nov. (paratype, NHMD-001801226) whole body showing epidermal plates (cyan = anti-serotonin-like immunostaining background signal (anti-5 HT)). (b) Left side *Limnognathia desmeti* sp. nov. (holotype, NHMD-001801220) whole body showing cilia and epidermal plates (yellow = acetylated α -tubulin-like immunoreactivity (α -tub-LIR), magenta = anti-5 HT background signal). (c) Ventral *Limnognathia desmeti* sp. nov. (holotype) whole body showing ciliated cells (glow = α -tub-LIR). (d) Dorsal anterior *Limnognathia maerski* Greenland, detail of apical plate (cyan = anti-synapsin1-like immunostaining background signal (ENZO Life Sciences, ADI-VAS-SV061-E)). (e) Lateral thoracic section *Limnognathia desmeti* sp. nov. (paratype, NHMD-001801223) detail of nephridial structures (glow = α -tub-LIR). (f) Vento-posterior abdomen of *Limnognathia desmeti* sp. nov. (paratype, NHMD-001801223) detail of adhesive ciliary pad (yellow = α -tub-LIR). (g) Mid-sagittal body of *Limnognathia desmeti* sp. nov. (holotype) whole body showing ciliation and internal structures (yellow = α -tub-LIR, cyan = DAPI, magenta = anti-5 HT background signal). dob, dorsalia cell body; dp, dorsal plate; en, endodermis; gl, gut lumen; all other abbreviations are identical to figure 1. Cells of the apical plate are numbered 1–4.

(b) Phylogenetics

All phylogenetic analyses recovered identical topologies except for a small difference in the relationships among Greenlandic specimens between topologies of M1 and M2 (figure 3a, electronic supplementary material, figures S2–S5). All nodes in all analyses had full support (bootstrap support (BS) = 100%, posterior probability (PP) = 1.00) except one node in the Greenland population of the maximum likelihood analysis of M2 (BS = 93%). We recover a monophyletic Micrognathozoa sister group to Syndermata along with monophyly of the three populations. The two Northern Hemisphere populations form a clade with a deep divergence between them and the Crozet population. The topology of all outgroup taxa reflects our current understanding of spiralian relationships (e.g. [9,10]), including Gnathifera (e.g. [11]) and the nested placement of Acanthocephala within Rotifera (e.g. [91,92]).

(c) Divergence dating

We dated the split between Northern Hemisphere and Crozet populations to ~4.6 mya, 95% CI (2.38–6.65 mya) and the split between Greenland and Pyrenees populations to ~0.686 mya, 95% CI (0.29–1.09 mya) (electronic supplementary material, figure S6). Universal substitution rates were used in the absence of specific rates for Micrognathozoa or other Gnathifera. Due to their small size and soft bodies, definitive fossil calibrations are not available for this clade. Despite the caveats of using universal rates, the magnitude of the divergences is too old to be explained by human introduction but too recent to be explained by vicariance. The disjunct distribution presumably arose by long-distance dispersal, likely involving a dormant egg [1,93].

Bipolar distributions could arise by a stepping-stone dispersal scenario [94] or a one-shot jump between poles [95,96]. Regardless, many Antarctic species are allied to South America [97–99] with similar divergence time estimates to Micrognathozoa [100–102]. Floristic connectivity in the Southern Hemisphere is more closely linked to wind connectivity than geographic proximity suggesting the predominant westerlies are the main driver of dispersal in the region [103] particularly in bryophytes [104,105], including genera that harbour Micrognathozoa [99]. Micrognathozoa are difficult to find due to their size, seasonality and patchy distributions. More populations are likely to be discovered in the future possibly mediated by eDNA sequencing [15,16]. Thus, we do not rule out pole-to-pole dispersal, mountain-top hopping or wind dispersal, e.g. via South America along with their necessary substrate (i.e. mosses).

(d) Species discovery

Molecular species discovery methods conflicted (figure 3a, electronic supplementary material, figures S7–S9). GMYC and bPTP both oversplit to varying degrees (seven species and eight species, respectively). bPTP results were consistent across runs, seeds and chain lengths. PTP joined the two Northern Hemisphere populations and delimited two species.

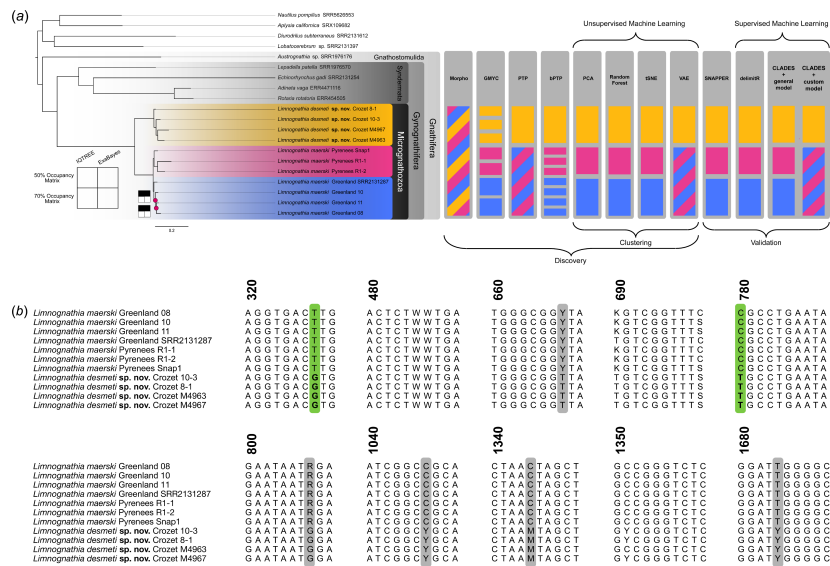


Figure 3. Phylotranscriptomic and species delimitation analyses of Micrognathozoa and single nucleotide polymorphisms in 18S rRNA. (a) Summary of phylogenetic and species delimitation analysis. Topology obtained from the maximum likelihood and Bayesian analysis of the 50% occupancy matrix. All nodes received full support in both analyses (BS = 100%, PP = 1). Differences in topology among the four analyses (50 versus 70% occupancy and maximum likelihood versus Bayesian) are summarized by the stability plots. Colours correspond to figure 1f–h. Results of species delimitation analyses are plotted to the right of the phylogeny. Blocks correspond to the number of recovered species. (b) Variation in 18S rRNA among *Limnognathia*. Fixed molecular diagnostic characters highlighted in green and other variable characters differing between the two species are highlighted in grey. Characters for *Limnognathia desmeti* sp. nov. are bolded.

UML clustering methods were largely congruent (figure 3a, electronic supplementary material, figures S10–S12) with three clusters corresponding to the three populations. RF+Gap Statistic found *L. maerski* Pyrenees Snap1 in its own cluster and RF+Hierarchical Clustering found nine clusters. Hierarchical Clustering underperformed in previous analyses and recovered more clusters than other methods [25]. VAE was congruent across runs and supported two species corresponding to Northern Hemisphere and Crozet.

(e) Species validation

Species discovery analyses mostly recovered two or three species of *Limnognathia*. Additionally, *Limnognathia* specimens from each locality likely belong to the same population because they were collected in very close proximity (e.g. from the same moss patches in a small body of water). Informed by these results, we attempted to validate either two or three species hypotheses.

SNAPPER results were consistent across step and MCMC generation configurations and preferred the three species model over the two species model (BF > 10) (figure 3a).

DelimitR implements RF classification by training the algorithm on data simulated under user-specified models to explicitly test species hypotheses and various population parameters including gene flow [27]. The out-of-bag error rates were small for the five models created by delimitR except those corresponding to model 3 and model 5 (electronic supplementary material, table S8). DelimitR selected the three species model without gene flow albeit with some error associated with model 5 (figure 3a, electronic supplementary material, figures S13a and Table S9). Admixture analysis confirmed the lack of gene flow among populations (electronic supplementary material, figure S13b).

The monogonont rotifer *Brachionus plicatilis* species complex shares many biological characteristics with *Limnognathia* including parthenogenesis (in part) [106,107] and widespread distribution [108,109]. Additionally, Syndermata is widely accepted as the sister phylum of Micrognathozoa [9–12]. In contrast to Micrognathozoa, *Brachionus*, as a model system, has been well studied, with its species supported by integrative taxonomy [110–113]. Additionally, the availability of transcriptomic data on this species complex makes it a suitable system to help us infer species limits in *Limnognathia*.

The two CLADES analyses produced conflicting results. The maximum likelihood classification based on the general model assigned the Greenland and Pyrenees populations as different species with moderate support (probability 0.713) (figure 3a). The general model has been shown to oversplit in some understudied taxa with complex speciation processes [25,28]. Classification using the custom model joined Greenland and Pyrenees with strong support (probability 0.8237) but split the Crozet population and the two Northern Hemisphere populations with moderate support (probability > 0.70) (figure 3a). We prefer the most conservative estimate across discovery and validation (PTP, VAE, custom CLADES) that clearly delimit only the Crozet population as a separate species. The Greenland and Pyrenees populations are likely undergoing incipient speciation.

(f) COI divergence

Saturation tests showed that COI was not saturated in *Limnognathia* (electronic supplementary material, figure S14). Intrapopulation pairwise distances of the trimmed alignment were 0% in all three populations (electronic supplementary material, table S10). Delimiting species with COI barcoding is standard [114] although the use of fixed thresholds is contentious [115–

117]. The distances between the Greenlandic and Pyrenees individuals (2.51%) are slightly higher than some well-studied taxa [118] and the common 2% divergence threshold for distinguishing species [119] but match values in species which are known to be overestimated by fixed thresholds [111,120–122]. The interpopulation divergences between Crozet and the other two populations (17.86%) are typical of congeneric species. Additionally, the difference between intra- and interpopulation divergence aligns with the $10 \times$ threshold [123] for population versus species divergence among congeners [19,20].

(g) Cryptic speciation in Micrognathozoa

Anatomical redundancy has been linked to morphological disparity, often leading to species diversity. Loss and gain of redundancy have led to evolutionary constraint and release from functional constraint, respectively [124–127]. With miniaturization, redundancy is often reduced [127,128]. Micrognathozoa, whether an example of extreme miniaturization or ancestrally small size, lack redundancy in certain structures such as muscles where each muscle may be formed by one cell [93]. The morphological conservatism between the two species of Micrognathozoa could stem from lack of redundancy shoehorning the phylum into morphological stasis.

Micrognathozoa's presumed parthenogenetic reproduction may be linked to its morphological stasis and cryptic speciation. Asexual lineages with geographical isolation are expected to diverge and form distinct species [129,130], but their phenotypic differences may not be obviously detectable [131]. Cryptic species have been found in many asexual lineages [110,120,132–134]. In fact, cryptic speciation is more common and morphology more conserved in parthenogenetic rotifers compared with sexually reproducing relatives [129,135]. Micrognathozoans have high habitat specificity, often only found in certain moss clumps even within the same pools of water [16]. Extreme isolation from infrequent long-distance dispersal of micrognathozoans with parthenogenetic reproduction could produce divergent species with identical morphologies.

4. Conclusion

Through extensive morphological analyses using LM, SEM and CLSM [1–3], this study revealed no distinct differences between Northern Hemisphere and Crozet specimens. The deep phylogenetic divergence between these two groups, their extreme geographical separation, differences in COI and 18S rRNA sequences, and the results of extensive species delimitation analyses clearly suggest that there are at least two species of *Limnognathia*. The age of the divergences between Northern and Southern Hemisphere populations indicate long-distance dispersal potentially through high-elevation stepping stones, e.g. along the American Cordillera or directly from pole to pole. We highlight the ongoing search for more micrognathozoan populations globally in high latitudes and high elevations, particularly in South America, to fully understand the biogeographic history and dispersal capabilities of Micrognathozoa.

The novelty of this highly divergent lineage of Micrognathozoa on a Subantarctic island with broad implications for biogeography and speciation processes highlights the need for describing and naming this new species. Due to its indistinguishable morphology, we describe the new species using molecular data [136]. Highly conserved ribosomal RNA sequences have been used in the species descriptions of other understudied and morphologically conserved groups (e.g. [19,137,138]) and are subsequently used here to establish fixed diagnostic characters.

5. Taxonomy

Phylum **Micrognathozoa** Kristensen and Funch, 2000

Family **Limnognathiidae** Kristensen and Funch, 2000

Genus ***Limnognathia*** Kristensen and Funch, 2000 figs. 1 and 2

urn:lsid:zoobank.org:act:B91157F7-BFB1-40B5-AC02-B658EE76F027

Limnognathia Kristensen and Funch, 2000:

Type species: Limnognathia maerski Kristensen and Funch, 2000

Amended diagnosis

Acoelomate, bilaterally symmetrical, unsegmented animals with intracellular plates in the non-ciliated dorsal and lateral epidermis. Epidermis is not syncytial. Adults 100–150 μm in length, juveniles $\sim 85 \mu\text{m}$. Body with three divisions: head, accordion-like thorax and abdomen. Transient dorsal anus on the abdomen. Females with one pair of ovaries, carrying a single large egg ($40 \times 60 \mu\text{m}$) at a time. No males known, likely parthenogenetic.

Unique cuticular pharyngeal apparatus with multiple sclerotized elements organized into four main sets—a pair of large fibularia (including basal plates), central main jaws (including cauda and symphysis), ventral jaws (pseudophalangia and lamellae), dorsal jaws—and associated accessory sclerites. Ventral mouth opening with two anterior ciliary tufts. Paired pharyngeal ciliary tufts in the mouth cavity between the main jaws.

Ventral epidermis with unique arrangement of ciliophores. Head with anterior ciliary field, posterior ciliary field and three pairs of head ciliophores lateral to the mouth. Thorax and anterior abdomen with 15 midventral rows of 2–4 ciliophores forming a locomotory band. Posterior abdomen with cluster of 10 ciliophores forming an adhesive ciliary pad.

Sensory collar receptors formed by 1–3 cilia surrounded by microvilli arranged as minimum three apicalia, two pairs of frontalia, three to five pairs of lateralia, three pairs of dorsalia and two pairs of caudalia. Additional sensoria include two

antero-lateral ‘brushes’ of shorter cilia (apical ciliary tufts) and a pair of long bendable cilia on the head (flagellar head structure). On the head, a pair of hyaline vesicles associated with a dense layer of microvilli possibly form phaosomal eyes.

Internal ciliary structures include two L-shaped oviducts and an anterior pair of protonephridia with four monociliated terminal cells, connected via intermediate ducts to a posterior pair of protonephridia with two terminal cells.

The central nervous system comprises a dorsal brain and four longitudinal ventral neurite bodies. The two main ventro-lateral and two ventro-median bundles are connected via subpharyngeal and posterior commissures. Pharyngeal ‘ganglion’ with paired ciliary receptors, potentially innervates the jaw musculature. Peripheral nerves extend ventrolaterally and dorsolaterally from the circumesophageal connectives. Body musculature consists of 7 pairs of longitudinal and 13 pairs of oblique dorso-ventral muscles.

Limnognathia maerski Kristensen and Funch, 2000

urn:lsid:zoobank.org:act:F5DF3F6E-49F8-468D-95D9-A9A4342CBFD0

Limnognathia maerski Kristensen and Funch, 2000: 4–11, figs. 1 and 2

Material examined

Molecular vouchers. Greenland: Disko Island: four specimens collected in Isunngua Spring 69.716667, 51.933333, Aug.2013 (SRR29877518–SRR29877520); Spain: Parc Nacional d’Aiguestortes i Estany de Sant Maurici: three specimens collected in Bassa Nera 42.638183, 0.924217 Jun.2022 (SRR29877515–SRR29877517).

Amended diagnosis

L. maerski is distinguished from *Limnognathia desmeti* sp. nov. by two fixed molecular diagnostic synapomorphies 18S rRNA (figure 3B). Coordinates refer to the position of the Sanger reference.

18S: 327(T), 780(G)

Remarks

Although reciprocally monophyletic, we currently cannot distinguish the Pyrenees population from Bassa Nera as a separate species from *L. maerski* based on the diagnostic characters here provided, which are requisite for erecting a new species. These specimens share with *L. maerski* from Greenland the same molecular synapomorphies in 18S rRNA and are thus here tentatively classified as the same species. However, due to the results of most species delimitation analyses, the two populations are likely undergoing incipient speciation. Specimens from both Northern Hemisphere populations were frequently seen swimming with a characteristic slow, rotational upward-moving motion. Additional characters include two variable positions in 18S rRNA—(668(Y), 807(R))—that are homozygous in *L. desmeti* sp. nov. and 88 fixed molecular synapomorphies in COI that were consistent among all individuals sampled (COI: see electronic supplementary material, table S11 and figure S15). We tentatively include these as characters in remarks with the caveat of small sample size and acknowledge more variability might be uncovered in the future.

Limnognathia desmeti Worsaae & Møller, sp. nov.

(figures 1b–d, 2a–c, e–g)

urn:lsid:zoobank.org:act:7A9BCA8B-010B-44F0-9AE9-45044E037AF0

Material examined

Holotype. Crozet Archipelago: Île de la Possession: Pointe du Bougainville: NHMD-001801220, 1 specimen, –46.447919, 51.850842, 11.Dec.2019.

Paratypes. Crozet Archipelago: Île de la Possession: Pointe du Bougainville: NHMD-001801224, NHMD-001801225, NHMD-001801227, three specimens as permanent whole mounts on slides, –46.447919, 51.850842, 10.Dec.2019, NHMD-001801221, NHMD-001801222, NHMD-001801223, NHMD-001801226, four specimens as permanent whole mounts on slides, –46.447919, 51.850842, 11.Dec.2019

Molecular vouchers. Crozet Archipelago: Île de la Possession: four specimens collected at Pointe du Bougainville –46.446002, 51.849662, December 2019 (SRR29877511–SRR29877514).

Diagnosis

L. desmeti sp. nov. is distinguished from *L. maerski* by two fixed molecular diagnostic synapomorphies 18S rRNA (figure 3b). Coordinates refer to the position of the Sanger reference.

18S: 327(G), 780(T)

Type locality

Crozet Archipelago, Île de la Possession, Pointe du Bougainville, –46.447919, 51.850842, 130 m (electronic supplementary material, figure S16).

Etymology

This species is named in honour of Dr Willem H. De Smet, University of Antwerp, who first discovered *Limnognathia* on the Crozet Archipelago.

Remarks

This species is found in mosses from small freshwater creeks and ponds. Specimens were always observed attached to or gliding on the substrate and never swimming. Additional characters include three variable positions in 18S rRNA—1046(Y), 1344(M), 1684(Y)—that are homozygous in *L. maerski* and 88 fixed molecular synapomorphies in COI following the justification in the description for *L. maerski* (COI: see electronic supplementary material, table S11 and Figure 15)

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. New sequences are deposited in the Sequence Read Archive (BioProject PRJNA1137119, Accession Numbers SRR29877511–SRR29877520). All matrices, alignments, CLADES files, custom scripts and tree files can be found on Harvard Dataverse [139]. All type specimens have been deposited in the Natural History Museum of Denmark. Supplementary material is available online [140].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. S.S.: conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing—original draft; C.A.: investigation, writing—review and editing; O.S.W.: resources, writing—review and editing; S.G.-P.: resources, writing—review and editing; C.E.L.: investigation, writing—review and editing; M.H.: investigation, writing—review and editing; G.G.: funding acquisition, resources, writing—review and editing; D.R.: funding acquisition, writing—review and editing; P.R.M.: investigation, resources, writing—review and editing; K.W.: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, visualization, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. The Villum foundation (grant no. 1 02 544 to KW) funded the collection of Greenlandic specimens. The French Polar Institute Paul-Emile Victor (Project 136-SUBANTECO to DR) funded the collection of Crozet specimens. Funding from the Putnam Expedition Grant, Museum of Comparative Zoology to GG and the Department of Biology facilitated the collection of Pyrenees specimens. The project was further supported by research and infrastructure grants to KW (Villum Foundation Grant no. 50 428 and Carlsberg Foundation Grant Nos. CF15–0946 and CF20–0458). SGP was supported by the Beatriz de Pinós Programme (BP–2021–00131) and a fellowship from “la Caixa” Foundation (ID 100010434, fellowship code LCF/BQ/PI24/12040011). OSW is a member of the research group SGR2021–01271 funded by the Generalitat de Catalunya.

Acknowledgements. We thank Nicolas Bekkouche and Reinhardt M. Kristensen for help with sampling in Greenland in 2013 and for years of inspiring discussion on Micrognathozoa. We thank Olivia Ingrid Ben Lundgren for her assistance with mapping the ciliary morphology of the Crozet Limnognathia. We are indebted to the staff at the Arctic Station and Alfred-Faure base (Yann Le Meur, Romuald Bellec, Mathilde Defourneaux) who made collecting in Greenland and Crozet possible. We thank the Faculty of Science and the Department of Biology at the University of Copenhagen for computational resources on the Danish National Supercomputer for Life Sciences Computerome 2.0 on which all bioinformatic analyses were conducted. Three anonymous reviewers provided comments that greatly improved the manuscript.

References

- Kristensen RM, Funch P. 2000 Micrognathozoa: a new class with complicated jaws like those of Rotifera and Gnathostomulida. *J. Morphol.* **246**, 1–49. (doi:10.1002/1097-4687(200010)246:13.0.co;2-d)
- De Smet WH. 2002 A new record of *Limnognathia maerski* Kristensen & Funch, 2000 (Micrognathozoa) from the subantarctic Crozet Islands, with redescription of the trophi. *J. Zool.* **258**, 381–393. (doi:10.1017/s095283690200153x)
- Sørensen MV. 2003 Further structures in the jaw apparatus of *Limnognathia maerski* (Micrognathozoa), with notes on the phylogeny of the Gnathifera. *J. Morphol.* **255**, 131–145. (doi:10.1002/jmor.10038)
- Sørensen MV. 2000 An SEM study of the jaws of *Haplognathia rosea* and *Rastrognathia macrostoma* (Gnathostomulida), with a preliminary comparison with the rotiferan trophi. *Acta Zool.* **81**, 200. (doi:10.1046/j.1463-6395.2000.00032.x)
- Bekkouche N, Kristensen RM, Hejnol A, Sørensen MV, Worsaae K. 2014 Detailed reconstruction of the musculature in *Limnognathia maerski* (Micrognathozoa) and comparison with other Gnathifera. *Front. Zool.* **11**, 71. (doi:10.1186/s12983-014-0071-z)
- Bekkouche N, Worsaae K. 2016 Nervous system and ciliary structures of Micrognathozoa (Gnathifera): evolutionary insight from an early branch in Spiralia. *R. Soc. Open Sci.* **3**, 160289. (doi:10.1098/rsos.160289)
- Giribet G, Sørensen MV, Funch P, Kristensen RM, Sterrer W. 2004 Investigations into the phylogenetic position of Micrognathozoa using four molecular loci. *Cladistics* **20**, 1–13. (doi:10.1111/j.1096-0031.2004.00004.x)
- Giribet G, Edgecombe GD. 2020 *The Invertebrate Tree of Life*. Princeton, NJ: Princeton University Press. (doi:10.1515/9780691197067)
- Laumer CE *et al.* 2015 Spiralian phylogeny informs the evolution of microscopic lineages. *Curr. Biol.* **25**, 2000–2006. (doi:10.1016/j.cub.2015.06.068)
- Laumer CE *et al.* 2019 Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proc. R. Soc. B* **286**, 20190831. (doi:10.1098/rspb.2019.0831)
- Marlétaz F, Peijnenburg KTCA, Goto T, Satoh N, Rokhsar DS. 2019 A new spiralian phylogeny places the enigmatic arrow worms among gnathiferans. *Curr. Biol.* **29**, 312–318. (doi:10.1016/j.cub.2018.11.042)
- Bekkouche N, Gąsiorowski L. 2022 Careful amendment of morphological data sets improves phylogenetic frameworks: re-evaluating placement of the fossil *Amiskwia sagittiformis*. *J. Syst. Palaeontol.* **20**, 1–14. (doi:10.1080/14772019.2022.2109217)
- Worsaae K, Kristensen RM. 2016 Phylum Micrognathozoa: the micrognathozoans. In *Invertebrates* (eds RC Brusca, W Moore, SM Shuster), pp. 626–632, 3rd edn. Sunderland, MA: Sinauer Associates.
- Garcés-Pastor S, Wangenstein OS, Pérez-Haase A, Pélachs A, Pérez-Obiol R, Cañellas-Boltà N, Mariani S, Vegas-Vilarrúbia T. 2019 DNA metabarcoding reveals modern and past eukaryotic communities in a high-mountain peat bog system. *J. Paleolimnol.* **62**, 425–441. (doi:10.1007/s10933-019-00097-x)
- Cordier T, Reo E, Čermáková K, Höss S, Pawloski J, Weigand A. 2020 Metabarcoding of sediment communities. In *Joint Danube survey 4 scientific report: a shared analysis of the Danube River* (eds I Liška, F Wagner, S M., K Deutsch, J Slobodnik, M Paunović), pp. 157–164. Vienna, Austria: International Commission for the Protection of the Danube River.
- Giribet G, Wangenstein OS, Garcés-Pastor S, Møller PR, Worsaae K. 2023 Using eDNA to find Micrognathozoa. *Curr. Biol.* **33**, R756–R757. (doi:10.1016/j.cub.2023.04.075)
- Zverkov OA *et al.* 2019 Dicyemida and Orthonectida: two stories of body plan simplification. *Front. Genet.* **10**, 443. (doi:10.3389/fgene.2019.00443)
- Carstens BC, Pelletier TA, Reid NM, Satler JD. 2013 How to fail at species delimitation. *Mol. Ecol.* **22**, 4369–4383. (doi:10.1111/mec.12413)
- Defourneaux É, Herranz M, Armenteros M, Sørensen MV, Norenburg JL, Park T, Worsaae K. 2024 Circumtropical distribution and cryptic species of the meiofaunal enteropneust *Meioglossus* (Harrimaniidae, Hemichordata). *Sci. Rep.* **14**, 9296. (doi:10.1038/s41598-024-57591-0)
- Worsaae K, Mikkelsen MD, Martínez A. 2019 Description of six new species of *Mesonerilla* (Nerillidae, Annelida) and an emended description of *M. intermedia* Wilke, 1953, from marine and cave environments. *Mar. Biodivers.* **49**, 2141–2165. (doi:10.1007/s12526-019-00984-6)
- Leaché AD, Zhu T, Rannala B, Yang Z. 2019 The spectre of too many species. *Syst. Biol.* **68**, 15. (doi:10.1093/sysbio/syy051)
- Sukumaran J, Knowles LL. 2017 Multispecies coalescent delimits structure, not species. *Proc. Natl Acad. Sci. USA* **114**, 1607–1612. (doi:10.1073/pnas.1607921114)
- Chambers EA, Hillis DM. 2020 The multispecies coalescent over-splits species in the case of geographically widespread taxa. *Syst. Biol.* **69**, 184–193. (doi:10.1093/sysbio/syz042)
- Mason NA, Fletcher NK, Gill BA, Funk WC, Zamudio KR. 2020 Coalescent-based species delimitation is sensitive to geographic sampling and isolation by distance. *Syst. Biodivers.* **18**, 269–280. (doi:10.1080/14772000.2020.1730475)

25. Derkarabetian S, Castillo S, Koo PK, Ovchinnikov S, Hedin M. 2019 A demonstration of unsupervised machine learning in species delimitation. *Mol. Phylogenetics Evol.* **139**, 106562. (doi:10.1016/j.ympev.2019.106562)
26. Alexander Pyron R. 2023 Unsupervised machine learning for species delimitation, integrative taxonomy, and biodiversity conservation. *Mol. Phylogenetics Evol.* **189**, 107939. (doi:10.1016/j.ympev.2023.107939)
27. Smith ML, Carstens BC. 2020 Process-based species delimitation leads to identification of more biologically relevant species*. *Evolution* **74**, 216–229. (doi:10.1111/evo.13878)
28. Derkarabetian S, Starrett J, Hedin M. 2022 Using natural history to guide supervised machine learning for cryptic species delimitation with genetic data. *Front. Zool.* **19**, 8. (doi:10.1186/s12983-022-00453-0)
29. Pei J, Chu C, Li X, Lu B, Wu Y. 2018 CLADES: a classification-based machine learning method for species delimitation from population genetic data. *Mol. Ecol. Resour.* **18**, 1144–1156. (doi:10.1111/1755-0998.12887)
30. Heine HLA, Derkarabetian S, Morisawa R, Fu PA, Moyes NHW, Boyer SL. 2024 Machine learning approaches delimit cryptic taxa in a previously intractable species complex. *Mol. Phylogenetics Evol.* **195**, 108061. (doi:10.1016/j.ympev.2024.108061)
31. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994 DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **3**, 294–299.
32. Kearse M *et al.* 2012 Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649. (doi:10.1093/bioinformatics/bts199)
33. Song L, Florea L. 2015 Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. *GigaScience* **4**, 48. (doi:10.1186/s13742-015-0089-y)
34. Krueger F. 2021 Trim Galore! See https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
35. Langmead B, Salzberg SL. 2012 Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359. (doi:10.1038/nmeth.1923)
36. Grabherr MG *et al.* 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652. (doi:10.1038/nbt.1883)
37. Li W, Godzik A. 2006 CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659. (doi:10.1093/bioinformatics/btl158)
38. Cunha TJ, Giribet G. 2019 A congruent topology for deep gastropod relationships. *Proc. R. Soc. B* **286**, 20182776. (doi:10.1098/rspb.2018.2776)
39. HaasBJ. TransDecoder. See <https://github.com/TransDecoder/TransDecoder>.
40. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015 BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212. (doi:10.1093/bioinformatics/btv351)
41. Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM. 2019 OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Res.* **47**, D807–D811. (doi:10.1093/nar/gky1053)
42. Emms DM, Kelly S. 2015 OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157. (doi:10.1186/s13059-015-0721-2)
43. Takebayashi N. 2015 uniqHaplo. See <http://raven.wrrb.uaf.edu/~ntakebay/teaching/programming/perl-scripts/uniqHaplo.pl>.
44. Katoh K, Standley DM. 2013 MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780. (doi:10.1093/molbev/mst010)
45. Di Franco A, Poujol R, Baurain D, Philippe H. 2019 Evaluating the usefulness of alignment filtering methods to reduce the impact of errors on evolutionary inferences. *BMC Evol. Biol.* **19**, 21. (doi:10.1186/s12862-019-1350-2)
46. Criscuolo A, Gribaldo S. 2010 BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, 210. (doi:10.1186/1471-2148-10-210)
47. Krug PJ, Caplins SA, Algosio K, Thomas K, Valdés ÁA, Wade R, Wong NLWS, Eernisse DJ, Kocot KM. 2022 Phylogenomic resolution of the root of Panpulmonata, a hyperdiverse radiation of gastropods: new insight into the evolution of air breathing. *Proc. R. Soc. B* **289**, 20211855. (doi:10.1098/rspb.2021.1855)
48. Price MN, Dehal PS, Arkin AP. 2010 FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490. (doi:10.1371/journal.pone.0009490)
49. Van der Auwera GA *et al.* 2013 From FastQ data to high-confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **43**, 11. (doi:10.1002/0471250953.bi1110s43)
50. Li H, Durbin R. 2009 Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760. (doi:10.1093/bioinformatics/btp324)
51. Danecek P *et al.* 2021 Twelve years of SAMtools and BCFtools. *GigaScience* **10**. (doi:10.1093/gigascience/giab008)
52. Broad Institute. 2019 Picard Toolkit. See <https://broadinstitute.github.io/picard/>.
53. Ortiz EM. 2019 vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis. (doi:10.5281/zenodo.2540861)
54. Banbury B, May MR. Phrynomics. See <https://github.com/bbanbury/phrynomics>.
55. Brasseur MV, Astrin JJ, Geiger MF, Mayer C. 2023 MitoGeneExtractor: efficient extraction of mitochondrial genes from next-generation sequencing libraries. *Methods Ecol. Evol.* **14**, 1017–1024. (doi:10.1111/2041-210x.14075)
56. Wcisł DJ, Howard JT III, Yoder JA, Dornburg A. 2020 Transcriptome Ortholog Alignment Sequence Tools (TOAST) for phylogenomic dataset assembly. *BMC Evol. Biol.* **20**, 41. (doi:10.1186/s12862-020-01603-w)
57. Castresana J. 2000 Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552. (doi:10.1093/oxfordjournals.molbev.a026334)
58. Talavera G, Castresana J. 2007 Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577. (doi:10.1080/10635150701472164)
59. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015 IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274. (doi:10.1093/molbev/msu300)
60. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018 UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 017. (doi:10.1093/molbev/msx281)
61. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. 2017 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589. (doi:10.1038/nmeth.4285)
62. Lanfear R, Calcott B, Ho SYW, Guindon S. 2012 PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* **29**, 1695–1701. (doi:10.1093/molbev/mss020)
63. Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A. 2014 Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol. Biol.* **14**. (doi:10.1186/1471-2148-14-82)

64. Aberer AJ, Kobert K, Stamatakis A. 2014 ExaBayes: massively parallel Bayesian tree inference for the whole-genome era. *Mol. Biol. Evol.* **31**, 2553–2556. (doi:10.1093/molbev/msu236)
65. Drummond AJ, Rambaut A. 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214. (doi:10.1186/1471-2148-7-214)
66. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018 Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**, 901–904. (doi:10.1093/sysbio/syy032)
67. Bouckaert R *et al.* 2019 BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **15**, e1006650. (doi:10.1371/journal.pcbi.1006650)
68. Wilke T, Schultheiß R, Albrecht C. 2009 As time goes by: a simple fool's guide to molecular clock approaches in invertebrates. *Am. Malacol. Bull.* **27**, 25–45. (doi:10.4003/006.027.0203)
69. Ochman H, Wilson AC. 1987 Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol.* **26**, 74–86. (doi:10.1007/bf02111283)
70. Zhang J, Kapli P, Pavlidis P, Stamatakis A. 2013 A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* **29**, 2869–2876. (doi:10.1093/bioinformatics/btt499)
71. Fujisawa T, Barraclough TG. 2013 Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Syst. Biol.* **62**, 707–724. (doi:10.1093/sysbio/syt033)
72. Purcell S *et al.* 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575. (doi:10.1086/519795)
73. Jombart T. 2008 adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405. (doi:10.1093/bioinformatics/btn129)
74. Jombart T, Ahmed I. 2011 adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071. (doi:10.1093/bioinformatics/btr521)
75. Liaw A, Wiener M. 2002 Classification and regression by randomForest. *R News* **2**, 18–22.
76. Donaldson J. 2016 tsne: t-distributed stochastic neighbor embedding for R (t-SNE). See <https://CRAN.R-project.org/package=tsne>.
77. Pedregosa F, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V. 2011 Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2835–2830.
78. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. 2018 cluster: cluster analysis basics and extensions. See <https://CRAN.R-project.org/package=cluster>.
79. Kassambara A, Mundt F. 2017 factoextra: extract and visualize the results of multivariate data analysis. See <https://CRAN.R-project.org/package=factoextra>.
80. Scrucra L, Fop M, Murphy TB, Raftery AE. 2016 mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *R J.* **8**, 289–317. (doi:10.32614/rj-2016-021)
81. Coombes KR, Wang M. 2018 PCDimension: finding the number of significant principal components. See <https://CRAN.R-project.org/package=PCDimension>.
82. Kingma DP, Welling M. 2013 Auto-encoding variational Bayes. In Proceedings of the International Conference on Learning Representations (ICLR). *arXiv* (doi:10.48550/arXiv.1312.6114)
83. Stoltz M, Baeumer B, Bouckaert R, Fox C, Hiscott G, Bryant D. 2021 Bayesian inference of species trees using diffusion models. *Syst. Biol.* **70**, 145–161. (doi:10.1093/sysbio/syaa051)
84. Oaks J. 2013 pyule. See <https://github.com/joaks1/pyule>.
85. Overcast I. 2023 easySFS. See <https://github.com/isaacovercast/easySFS>.
86. Chang CC, Lin CJ. 2011 Libsvm. *ACM Trans. Intell. Syst. Technol.* **2**, 1–27. (doi:10.1145/1961189.1961199)
87. Alexander DH, Novembre J, Lange K. 2009 Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664. (doi:10.1101/gr.094052.109)
88. Tamura K, Stecher G, Kumar S. 2021 MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027. (doi:10.1093/molbev/msab120)
89. Kimura M. 1980 A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120. (doi:10.1007/bf01731581)
90. Gąsiorowski L, Bekkouche N, Worsaae K. 2017 Morphology and evolution of the nervous system in Gnathostomulida (Gnathifera, Spiralia). *Org. Divers. Evol.* **17**, 447–475. (doi:10.1007/s13127-017-0324-8)
91. Garey JR, Near TJ, Nonnemacher MR, Nadler SA. 1996 Molecular evidence for Acanthocephala as a subtaxon of Rotifera. *J. Mol. Evol.* **43**, 287–292. (doi:10.1007/pl00006088)
92. Wey-Fabrizius AR, Herlyn H, Rieger B, Rosenkranz D, Witek A, Welch DBM, Ebersberger I, Hankeln T. 2014 Transcriptome data reveal syndermatan relationships and suggest the evolution of endoparasitism in Acanthocephala via an epizoid stage. *PLoS One* **9**, e88618. (doi:10.1371/journal.pone.0088618)
93. Sørensen MV, Kristensen RM. 2015 Micrognathozoa. In *Handbook of Zoology Gastrotricha and Gnathifera* (ed. A Schmidt-Rhaesa), pp. 197–216. Boston, MA: De Gruyter. (doi:10.1515/9783110274271.197)
94. Villaverde T, Escudero M, Martín-Bravo S, Jiménez-Mejías P, Sanmartín I, Vargas P, Luceño M. 2017 Bipolar distributions in vascular plants: a review. *Am. J. Bot.* **104**, 1680–1694. (doi:10.3732/ajb.1700159)
95. Biersma EM, Jackson JA, Hyvönen J, Koskinen S, Linse K, Griffiths H, Convey P. 2017 Global biogeographic patterns in bipolar moss species. *R. Soc. Open Sci.* **4**, 170147. (doi:10.1098/rsos.170147)
96. Zawierucha K *et al.* 2023 Cryophilic Tardigrada have disjunct and bipolar distribution and establish long-term stable, low-density demes. *Polar Biol.* **46**, 1011–1027. (doi:10.1007/s00300-023-03170-4)
97. Guidetti R, McInnes SJ, Cesari M, Rebecchi L, Rota-Stabelli O. 2017 Evolutionary scenarios for the origin of an Antarctic tardigrade species based on molecular clock analyses and biogeographic data. *Contrib. Zool.* **86**, 97–110. (doi:10.1163/18759866-08602001)
98. Allegrucci G, Carchini G, Todisco V, Convey P, Sbordoni V. 2006 A molecular phylogeny of Antarctic Chironomidae and its implications for biogeographical history. *Polar Biol.* **29**, 320–326. (doi:10.1007/s00300-005-0056-7)
99. Saługa M, Ochrya R, Żarnowiec J, Ronikier M. 2018 Do Antarctic populations represent local or widespread phylogenetic and ecological lineages? Complicated fate of bipolar moss concepts with *Drepanocladus longifolius* as a case study. *Org. Divers. Evol.* **18**, 263–278. (doi:10.1007/s13127-018-0372-8)
100. Maturana CS *et al.* 2024 Ancient diversification in extreme environments: exploring the historical biogeography of the Antarctic winged midge *Parochlus steinenii* (Diptera: Chironomidae). *Front. Ecol. Evol.* **12**. (doi:10.3389/fevo.2024.1393376)
101. Maturana CS, Biersma EM, Díaz A, González-Wevar C, Contador T, Convey P, Jackson JA, Poulin E. 2022 Survivors and colonizers: contrasting biogeographic histories reconciled in the Antarctic freshwater copepod *Boeckella poppei*. *Front. Ecol. Evol.* **10**. (doi:10.3389/fevo.2022.1012852)
102. Maturana CS *et al.* 2021 Historical biogeography of the Gondwanan freshwater genus *Boeckella* (Crustacea): timing and modes of speciation in the Southern Hemisphere. *Divers. Distrib.* **27**, 2330–2343. (doi:10.1111/ddi.13405)
103. Muñoz J, Felicísimo AM, Cabezas F, Burgaz AR, Martínez I. 2004 Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science* **304**, 1144–1147. (doi:10.1126/science.1095210)

104. Kato K, Arikawa T, Imura S, Kanda H. 2013 Molecular identification and phylogeny of an aquatic moss species in Antarctic lakes. *Polar Biol.* **36**, 1557–1568. (doi:10.1007/s00300-013-1373-x)
105. Biersma EM, Jackson JA, Bracegirdle TJ, Griffiths H, Linse K, Convey P. 2018 Low genetic variation between South American and Antarctic populations of the bank-forming moss *Chorisodontium aciphyllum* (Dicranaceae). *Polar Biol.* **41**, 599–610. (doi:10.1007/s00300-017-2221-1)
106. Serra M, King CE. 1999 Optimal rates of bisexual reproduction in cyclical parthenogens with density-dependent growth. *J. Evol. Biol.* **12**, 263–271. (doi:10.1046/j.1420-9101.1999.00026.x)
107. Stelzer CP, Snell TW. 2003 Induction of sexual reproduction in *Brachionus plicatilis* (Monogononta, Rotifera) by a density-dependent chemical cue. *Limnol. Oceanogr.* **48**, 939–943. (doi:10.4319/lo.2003.48.2.0939)
108. Segers H. 2008 Global diversity of rotifers (Rotifera) in freshwater. *Hydrobiologia* **595**, 49–59. (doi:10.1007/s10750-007-9003-7)
109. Dumont HJ. 1983 Biogeography of rotifers. *Hydrobiologia* **104**, 19–30.
110. Fontaneto D, Giordani I, Melone G, Serra M. 2007 Disentangling the morphological stasis in two rotifer species of the *Brachionus plicatilis* species complex. *Hydrobiologia* **583**, 297–307. (doi:10.1007/s10750-007-0573-1)
111. Mills S *et al.* 2017 Fifteen species in one: deciphering the *Brachionus plicatilis* species complex (Rotifera, Monogononta) through DNA taxonomy. *Hydrobiologia* **796**, 39–58. (doi:10.1007/s10750-016-2725-7)
112. Gabaldón C, Fontaneto D, Carmona MJ, Montero-Pau J, Serra M. 2017 Ecological differentiation in cryptic rotifer species: what we can learn from the *Brachionus plicatilis* complex. *Hydrobiologia* **796**, 7–18. (doi:10.1007/s10750-016-2723-9)
113. Walczyńska A, Serra M. 2014 Inter- and intraspecific relationships between performance and temperature in a cryptic species complex of the rotifer *Brachionus plicatilis*. *Hydrobiologia* **734**, 17–26. (doi:10.1007/s10750-014-1859-8)
114. Hebert PD, Cywinska A, Ball SL, deWaard JR. 2003 Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B* **270**, 313–321. (doi:10.1098/rspb.2002.2218)
115. Rubinfeld D. 2006 Utility of mitochondrial DNA barcodes in species conservation. *Conserv. Biol.* **20**, 1026–1033. (doi:10.1111/j.1523-1739.2006.00372.x)
116. Hickerson MJ, Meyer CP, Moritz C. 2006 DNA barcoding will often fail to discover new animal species over broad parameter space. *Syst. Biol.* **55**, 729–739. (doi:10.1080/10635150600969898)
117. Meier R, Shiyang K, Vaidya G, Ng PKL. 2006 DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Syst. Biol.* **55**, 715–728. (doi:10.1080/10635150600969864)
118. Ward RD. 2009 DNA barcode divergence among species and genera of birds and fishes. *Mol. Ecol. Resour.* **9**, 1077–1085. (doi:10.1111/j.1755-0998.2009.02541.x)
119. Ratnasingham S, Hebert PDN. 2013 A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. *PLoS ONE* **8**, e66213. (doi:10.1371/journal.pone.0066213)
120. Fontaneto D, Iakovenko N, Eyres I, Kaya M, Wyman M, Barraclough TG. 2011 Cryptic diversity in the genus *Adineta* Hudson & Gosse, 1886 (Rotifera: Bdelloidea: Adinetidae): a DNA taxonomy approach. *Hydrobiologia* **662**, 27–33. (doi:10.1007/s10750-010-0481-7)
121. Zhang HG, Lv MH, Yi WB, Zhu WB, Bu WJ. 2017 Species diversity can be overestimated by a fixed empirical threshold: insights from DNA barcoding of the genus *Cletus* (Hemiptera: Coreidae) and the meta-analysis of COI data from previous phylogeographical studies. *Mol. Ecol. Resour.* **17**, 314–323. (doi:10.1111/1755-0998.12571)
122. García-Morales AE, Elías-Gutiérrez M. 2013 DNA barcoding of freshwater Rotifera in Mexico: evidence of cryptic speciation in common rotifers. *Mol. Ecol. Resour.* **13**, 1097–1107. (doi:10.1111/1755-0998.12080)
123. Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM. 2004 Identification of birds through DNA barcodes. *PLoS Biol.* **2**, e312. (doi:10.1371/journal.pbio.0020312)
124. Schaefer SA, Lauder GV. 1996 Testing historical hypotheses of morphological change: biomechanical decoupling in loricarioid catfishes. *Evolution* **50**, 1661–1675. (doi:10.1111/j.1558-5646.1996.tb03938.x)
125. Friel JP, Wainwright PC. 1998 Evolution of motor patterns in tetraodontiform fishes: does muscle duplication lead to functional diversification? *Brain Behav. Evol.* **52**, 159–170. (doi:10.1159/000006560)
126. Frédéric B, Olivier D, Litsios G, Alfaro ME, Parmentier E. 2014 Trait decoupling promotes evolutionary diversification of the trophic and acoustic system of damselfishes. *Proc. R. Soc. B* **281**, 20141047. (doi:10.1098/rspb.2014.1047)
127. Boudinot BE, Beutel RG, Gorb SN, Polilov AA. 2021 Functional diversity of attachment and grooming leg structures is retained in all but the smallest insects. *J. Zool.* **313**, 99–113. (doi:10.1111/jzo.12840)
128. Diakova AV, Makarova AA, Polilov AA. 2018 Between extreme simplification and ideal optimization: antennal sensilla morphology of miniaturized *Megaphragma* wasps (Hymenoptera: Trichogrammatidae). *PeerJ* **6**, e6005. (doi:10.7717/peerj.6005)
129. Fontaneto D, Barraclough TG. 2015 Do species exist in asexuals? Theory and evidence from bdelloid rotifers. *Integr. Comp. Biol.* **55**, 253–263. (doi:10.1093/icb/icv024)
130. Barraclough TG, Birky CW, Burt A. 2003 Diversification in sexual and asexual organisms. *Evolution* **57**, 2166. (doi:10.1554/02-339)
131. William Birky C Jr, Barraclough TG. 2009 Asexual speciation. In *Lost sex* (eds I Schön, K Martens, P Dijk), pp. 201–216. Dordrecht, The Netherlands: Springer. (doi:10.1007/978-90-481-2770-2_10)
132. Fontaneto D, Boschetti C, Ricci C. 2008 Cryptic diversification in ancient asexuals: evidence from the bdelloid rotifer *Philodina flaviceps*. *J. Evol. Biol.* **21**, 580–587. (doi:10.1111/j.1420-9101.2007.01472.x)
133. Heethoff M, Domes K, Laumann M, Maraun M, Norton RA, Scheu S. 2007 High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *J. Evol. Biol.* **20**, 392–402. (doi:10.1111/j.1420-9101.2006.01183.x)
134. von Saltzwedel H, Scheu S, Schaefer I. 2017 Genetic structure and distribution of *Parisotoma notabilis* (Collembola) in Europe: cryptic diversity, split of lineages and colonization patterns. *PLoS One* **12**, e0170909. (doi:10.1371/journal.pone.0170909)
135. Fontaneto D, Kaya M, Herniou EA, Barraclough TG. 2009 Extreme levels of hidden diversity in microscopic animals (Rotifera) revealed by DNA taxonomy. *Mol. Phylogenetics Evol.* **53**, 182–189. (doi:10.1016/j.ympev.2009.04.011)
136. Cook LG, Edwards RD, Crisp MD, Hardy NB. 2010 Need morphology always be required for new species descriptions? *Invertebr. Syst.* **24**, 322. (doi:10.1071/is10011)
137. Edgecombe GD, Giribet G. 2008 A New Zealand species of the trans-Tasman centipede order Craterostigmomorpha (Arthropoda: Chilopoda) corroborated by molecular evidence. *Invertebr. Syst.* **22**, 1. (doi:10.1071/is07036)
138. Sato S, Buckman-Young RS, Harvey MS, Giribet G. 2018 Cryptic speciation in a biodiversity hotspot: multilocus molecular data reveal new velvet worm species from Western Australia (Onychophora: Peripatopsidae: *Kumbadjena*). *Invertebr. Syst.* **32**, 1249. (doi:10.1071/is18024)
139. Sato S. Data for: Phylogenomics of the rarest animals: a second species of Micrognathozoa identified by machine learning. Harvard Dataverse (doi:10.7910/DVN/UHC3SR)
140. Sato S, Appeldoerff C, Wangenstein OS, Garcés-Pastor S, Laumer C, Herranz M *et al.* 2025 Supplementary material from: Phylogenomics of the rarest animals: a second species of Micrognathozoa identified by machine learning. Figshare (doi:10.6084/m9.figshare.c.7667300)