A fossil-calibrated phylogenomic analysis of *Daphnia* and the Daphniidae

Luca Cornetti[1], Peter D. Fields[1], Kay Van Damme[2,3] and Dieter Ebert[1]

[1] *University of Basel, Department of Environmental Sciences, Zoology, Vesalgasse 1, 4051 Basel, Switzerland*

[2] *Senckenberg Research Institute and Natural History Museum Frankfurt, Senckenberganlage 25, 60325 Frankfurt am Main, Germany*

[3] *Tvärminne Zoological Station, University of Helsinki, J.A. Palménin tie 260, Hanko, Finland*

Corresponding author: L. Cornetti, University of Basel, Department of Environmental Sciences, Zoology, Vesalgasse 1, 4051 Basel, Switzerland.
E-mail: luca.cornetti@unibas.ch

Citation: Luca Cornetti, Peter D. Fields, Kay Van Damme, Dieter Ebert, A fossil-calibrated phylogenomic analysis of Daphnia and the Daphniidae, Molecular Phylogenetics and Evolution, Volume 137, 2019, Pages 250-262, ISSN 1055-7903, https://doi.org/10.1016/j.ympev.2019.05.018.
http://www.sciencedirect.com/science/article/pii/S1055790318306018

Abstract

In the post-genomic era, much of phylogenetic analyses still relies on mitochondrial DNA, either alone or in combination with few nuclear genes. Although this approach often makes it possible to construct well-supported trees, it is limited because mtDNA describes the history of a single locus, and nuclear phylogenies based on a few loci may be biased, leading to inaccurate tree topologies and biased estimations of species divergence time. In this study, we perform a phylogenomic analysis of the Daphniidae family (Crustacea: Branchiopoda: Anomopoda) including some of the most frequently studied model organisms (*Daphnia magna* and *D. pulex*) whose phylogenetic relationships have been based primarily on an assessment of a few mtDNA genes. Using high-throughput sequencing, we were able to assemble 38 whole mitochondrial genomes and draft nuclear genomes for 18 species, including at least one species for each known genus of the family Daphniidae. Here we present phylogenies based on 636 nuclear single-copy genes shared among all sampled taxa and based on whole mtDNA genomes. The phylogenies we obtained were highly supported and showed some discrepancies between nuclear and mtDNA based trees at deeper nodes. We also identified a new candidate sister lineage of *Daphnia magna*. Our time-calibrated genomic trees, which we constructed using both fossil records and substitution rates, yielded very different estimates of branching event times compared to those based on mtDNA. By providing multi-locus, fossil-calibrated trees of the Daphniidae, our study contributes to an improved phylogenetic framework for ecological and evolutionary studies that use water fleas as a model system.

Keywords: substitution rates; fossil records; tree discrepancy; gene tree vs species tree; *Daphnia magna*; *Daphnia pulex*.

1. Introduction

For several decades now, molecular data has allowed researchers to resolve phylogenetic relationships in diverse organisms. However, phylogenies for most taxa are based on a limited number of genes that are usually part of the mitochondrial genome (mtDNA). Although mtDNA has the advantage of being a fast-evolving haploid molecule that can help reconstruct well-supported trees, it has substantial limitations, since it only describes the history of a single locus (Galtier et al., 2009). Indeed, discrepancies between individual gene phylogenies and the underlying species tree have been observed (Rubinoff and Holland, 2005). This problem is partially resolved by using nuclear genetic markers; however, selecting only a handful of independent markers from the nuclear genome can mislead researchers about phylogenetic relationships between taxa, and the topologies derived from those genes have often been inconsistent (Salichos and Rokas, 2013). These discrepancies may be caused by biological processes, such as gene duplication, incomplete lineage sorting and horizontal gene transfer via introgression (Maddison, 1997). Also, nuclear genes typically have a limited number of variable characters, which reduces their value for phylogenetic reconstructions. Given the recent rise of high-throughput sequencing techniques, the use of multiple independent loci is now becoming a prerequisite for obtaining robust phylogenies, increasing not only the accuracy of the tree topologies but also the accuracy of species divergence times estimated using molecular dating methods (Zhu et al., 2015).

Species trees can be time-calibrated with fossil records or dated biogeographical events, providing a temporal framework of taxa diversification (Rieux and Balloux, 2016). For those groups of organisms for whom reliable calibration points (i.e., fossil data) are scarce or nonexistent, a common alternative is to calibrate the molecular clock using substitution rates from species that are closely related to the focal group of organisms. This option can also be problematic, however, as the rate of molecular evolution has been shown to vary not only among lineages, but also within individual genomes, and over evolutionary time (Bromham, 2009; Bromham et al., 2018). In addition, using a limited number of markers or, for example, only mitochondrial genes, may yield inaccurate estimates of species divergence time (Arbogast et al., 2002). A well resolved time-calibrated phylogeny is not only a requirement to delimit the taxa and understand the relationships between them, but also provides the necessary baseline for ecological and evolutionary studies.

Certain species of the freshwater crustacean genus *Daphnia* O.F. Müller, 1785 (Branchiopoda, Cladocera, Anomopoda, Daphniidae) have become important models in ecology and evolution (Lampert, 2011; Stollewerk, 2010), epidemiology (Ebert, 2005), toxicology (Shaw et al., 2008), and, more recently, genomics (Colbourne et al., 2011). However, although these are some of the most frequently studied invertebrates, our understanding of the phylogenetic relationships among *Daphnia* species still relies mostly on mtDNA genes (cytochrome c oxidase subunit I; COI, 12S and 16S; Adamowicz et al., 2009; Marková et al., 2007; Popova et al., 2016; Schwenk et al., 2000), and divergence times have been estimated with a molecular clock based primarily on one mtDNA gene (reviewed in Van Damme and Kotov,

2016). The shortcomings of this approach became apparent when, in evaluating the divergence time between *D. magna* and *D. pulex* using eight nuclear genes, Haag et al., (2009) found very different time estimates compared to those based on both fossil records (Kotov and Taylor, 2011) and the mtDNA molecular clock (Colbourne and Hebert, 1996). The nature of such discrepancies may be due to multiple reasons, including the application of unspecific substitution rates, the use of different markers (nuclear vs mtDNA) and the lack of reliable fossil calibrations (Van Damme and Kotov, 2016).

The cladoceran fossil record has recently been revised, so that it now allows minimal time estimates for a few major nodes in the phylogeny of the Daphniidae (Van Damme and Kotov, 2016). Here, we use whole genome sequencing to shed light on the phylogenetic relationships among the Daphniidae in a temporal framework. We performed a phylogenetic study including 18 species from all five genera of the Daphniidae, as well as other anomopod families as outgroups. We focused primarily on the genus *Daphnia* and, in particular, on species phylogenetically close to *D. magna,* a keystone aquatic organism, with the aim of identifying its proper outgroup for comparative genomic studies. To date, both *D. exilis* and *D. similis* have been regarded as sister lineages to *D. magna* (Adamowicz et al., 2009; Colbourne and Hebert, 1996; Orsini et al., 2013), although Popova et al., (2016) has cautioned against assuming sister status on partially resolved phylogenies (Popova et al., 2016). To ascertain sister status definitively, studies must include more taxa and also increase the resolution of the analysis at deeper phylogenetic levels by analyzing multiple genes from the mitochondrial and nuclear genomes. For species within our focal group of taxa in the subgenus *Ctenodaphnia* (including *D. magna* and *D. similis* group), we included samples from wide-spread populations since some of these species have a very wide geographic distribution (Fields et al., 2018; Popova and Kotov, 2013) and it is not always clear how closely they are related.

Although *D. magna* and *D. pulex* are model species for the study of ecology and evolutionary processes, building on centuries of research (Lampert, 2011), they represent only a small part of Daphniidae diversity. In this study, we thus include other Daphniidae to provide a broader range of insights on important questions in evolution, ecology, and environmental sciences. For example, the *D. longispina* species complex has proven to be an excellent system for studying hybridization (e.g. Alric et al., 2016); *D. lumholtzi* has been widely studied as a successful invasive species in North America (i.e. Engel and Tollrian, 2009), and *D. similis* has been proposed as an alternative organism to *D. magna* in ecotoxicological tests (Rodgher et al., 2010). We also include *Daphnia* species whose remarkable inducible defenses have made them useful in phenotypic plasticity studies (i.e. *D. barbata* (Herzog et al., 2016); *D. atkinsoni* (Petrusek et al., 2009); *D. carinata* (Barry, 2000)). To date, the phylogenetic relationships among these species have been assessed by milestone studies in the field, based, however, on a few mitochondrial genes (Adamowicz et al., 2009, 2004; Crease et al., 2012; Petrusek et al., 2008; Popova et al., 2016; Taylor et al., 1996). We extend these studies, here, to the whole genome level. Furthermore, we incorporate all five Daphniidae genera, including the genus *Megafenestra*, which

had not been included in earlier studies (i.e. deWaard et al., 2006; Richter et al., 2007). This genus is important to include, since at least two genera of each subfamily *sensu* Dumont and Pensaert (Dumont and Pensaert, 1983) (Daphniinae: *Daphnia, Ceriodaphnia, Simocephalus*; Scapholeberinae: *Megafenestra, Scapholeberis*) are necessary to clarify the relationships in the family. Additionally, the proper identification of sister clades as choices of outgroups is crucial for comparative genomic studies (Rota-Stabelli and Telford, 2008) and is therefore a further aim of our study. Former phylogenetic studies at the systematic level of order to classes that included Daphniidae may suggest that any non-*Daphnia* genus could be the sister clade to *Daphnia* (Abele and Spears, 2000; deWaard et al., 2006; Richter et al., 2007; Stenderup et al., 2006; Swain and Taylor, 2003; Van Damme et al., 2007).

For the 18 species included here, we newly sequenced and assembled the entire mtDNA and nuclear genomes and used a multi-locus species tree to build a robust molecular phylogeny. We time calibrated these phylogenies using: (i) all the fossil records available for Anomopoda and (ii) mtDNA and nuclear substitution rates obtained from mutation-accumulation studies in *D. pulex* (Keith et al., 2016; Xu et al., 2012). We then compared time calibrated phylogenies obtained with different approaches and markers and discussed their similarities and discrepancies. By providing multi-locus, fossil-calibrated trees of Daphniidae, our study provides a robust phylogenetic framework for ecological and evolutionary studies that involve water fleas of the genus *Daphnia*.

2. Materials and Methods

2.1 Samples and Genomic DNA Extraction

Cladocera are cyclical parthenogens and can be maintained as stable genotypes (clones) under lab conditions in an asexual mode of reproduction. These clonal cultures produce sufficient material for genome sequencing. We analyzed the mitochondrial and nuclear genomes of 40 clones of the family Daphniidae and, as outgroups, one clone of *Moina brachiata* (family Moinidae) and one of *Bosmina* cf. *longispina maritima* (*B. coregoni*-group) (family Bosminidae). Both families of the Aradopoda (Daphniidae and Moinidae) are represented and, within the Daphniidae, taxa from all genera of the two subfamilies are included: the Daphniinae (*Ceriodaphnia, Daphnia, Simocephalus*) and the Scapholeberinae (*Megafenestra, Scapholeberis*).

Within *Daphnia,* we analyzed 14 different taxa belonging to the two subgenera, *Daphnia* (including *D. pulex* group and *D. longispina* group) and *Ctenodaphnia* (Table 1). Since one focus of the study is to provide a comparative phylogenomic frame for *D. magna,* being the most studied cladoceran, the taxon coverage in *Ctenodaphnia* (ten species) is higher than for the subgenus *Daphnia* (four species). While the dataset here represents a fraction of the diversity in the genus, ca. 90 valid taxa and many more with unresolved taxonomy (Kotov, 2015), it covers representatives of major species groups and includes three of the four major branches in *Ctenodaphnia* retrieved by Adamowicz et al. (2009). *D. ephemeralis*, which forms a separate lineage in the latter study, was not included here.

We also included *Daphnia* clones whose genomic resources were already available in sequence databases: two clones of the *D. pulex* complex belonging to different lineages, TCO (Colbourne et al., 2011) and PA42 (Ye et al., 2017), one *D. magna* clone (XINB3, Daphnia Genome Consortium) that was analyzed for both the mitochondrial and nuclear genomes, and one clone of the *D. obtusa* complex (Tucker et al., 2013) for which we were able to assemble only the mitochondrial genome. The cladoceran clones sequenced for this study were obtained from parthenogenetic females collected in the field, or were hatched from resting eggs collected in the field. Isofemale lines were produced by keeping individual females in isolation and allowing them to reproduce only asexually.

To reduce bacterial DNA, all animals were kept for three days in a solution of Ampicillin, Streptomycin and Tetracycline (Sigma) at a concentration of 50 mg/L each, and transferred daily into a fresh antibiotic solution before their DNA were extracted. To reduce gut content, the animals were also not fed during this three-day treatment, instead receiving 5 mg of superfine beads of the gel filtration resin Sephadex ® G-25 (Sigma-Aldrich) twice a day in their medium. For smaller species (less than 1 mm body size), bentonite clay (Bentonite MED) at a concentration of 50mg/L was added daily to the medium. Sephadex and bentonite clay cause gut evacuation when ingested by the animals. We extracted DNA from 30-100 animals of each clone. Genomic DNA was extracted using the QIAGEN Gentra Puregene Tissue Kit, including the RNaseA (100 mg/ml; Sigma) digestion step. Whole-genome Illumina paired-end sequencing (read length 125bp) was performed by the Genomics Facility service platform at the Department of Biosystem Science and Engineering (D-BSSE, ETH) in Basel, Switzerland, on an IlluminaHiSeq 2500.

2.2 Mitochondrial Genome Assemblies and Mitochondrial Datasets

After removing Illumina adapters from the raw reads using Trimmomatic version 0.35 (Bolger et al., 2014), the mitochondrial genomes were assembled. For each clone, a subset of two million randomly selected reads were used as input for the MITObim package (Hahn et al., 2013). MITObim employs a baiting and iteration mapping approach, implemented in the MIRAbait module of the MIRA assembler (Chevreux et al., 1999). The mitochondrial genome derived from the *D. magna* XINB3 individual genome (V2.4; Daphnia Genome Consortium) was used as a reference. This procedure was repeated four times for each clone, using a different subset of two million reads each time. To assess the consistency of the mitochondrial genome assemblies, the four individual sequences were aligned using MUSCLE v3.8.31 (Edgar, 2004) and visually checked for discrepancies. In the rare instances where discrepancies were found, the haplotype supported by the highest number of sequences was considered for further analyses.

We annotated the mitochondrial genomes independently using the MITOS web server (Bernt et al., 2013), which allowed us to identify the thirteen protein coding genes and the two structural rRNA genes. These genes were individually aligned with MUSCLE v3.8.31 (Edgar, 2004) and concatenated into a data matrix using the software Sequence Matrix (Vaidya et al., 2011) and keeping the information about the gene partitioning (i.e. the start and end positions for each

protein coding gene in the big data matrix). We prepared and separately analyzed three different datasets that consisted of (i) the concatenation of the thirteen protein coding genes (PCGs) and the two structural rRNA genes, (ii) the concatenation of only the thirteen protein coding genes, and (iii) the amino acid sequences of the thirteen protein coding genes. Because amino acid sequences are sometimes preferred in recovering deep phylogenetic relationships(Simmons et al., 2002), we included this dataset, even though several studies have shown that nucleotide sequences outperformed amino acid sequences also for the resolution of deep nodes (Simmons, 2017; Simmons et al., 2004).

To compare our data with previously published phylogenies, we aligned our nucleotide sequences first with those used by Adamowicz et al., (2009), analyzing the concatenation of 16S, 12S and COI mtDNA genes, but focusing only on taxa without missing data. We also aligned them with the nucleotide sequences used by Popova et al., (2016), but analyzed 12S and COI independently, since in the original article, different taxa were sequenced for these two mtDNA genes.

2.3 Nuclear Genome Assemblies

MaSuRCA (Zimin et al., 2013) was used to assemble the nuclear genome of each cladoceran clone. This method relies on the computational efficiency of the de Brujin graph methods combined with the flexibility of overlap-based assembly strategies. The Illumina paired-end reads were used as input for MaSuRCA and were assembled into super-reads. The assembly procedure was repeated three times for each clone, using the default settings but varying the kmer size (GRAPH_KMER_SIZE= 55, 65 and 75 were tested). The assembly statistics (number of scaffolds, N50, maximum scaffold length and total assembly length) were evaluated with ABySS 2.0.2 (Jackman et al., 2017). The resulting assembly containing the lowest number of scaffolds and expected genome length was considered the optimal assembly for our purposes and used for downstream analyses.

2.4 Ortholog Identification, Alignment and Nuclear Datasets

The nuclear genome assemblies were assessed for biological completeness using BUSCOv3 (Waterhouse et al., 2017). A total of 1,066 single-copy arthropod genes were searched against each individual assembly. The single-copy genes identified by the BUSCO approach as “complete” (i.e. without any in-frame stop-codon) were used to define ortholog groups across the cladoceran genomes and to build phylogenetic trees.

To obtain the high-confidence sequence alignments required for accurate phylogenetic analysis, especially when divergence time among taxa is relatively old (Kumar and Filipski, 2007), we computed the alignments with the software TranslatorX (Abascal et al., 2010). For each protein coding gene, the alignment, performed with MUSCLE v3.8.31 (Edgar, 2004), was guided by the corresponding deduced open reading frame. Ambiguously aligned positions were removed using Gblocks v.0.91b (Castresana, 2000), so that there were no gaps in the final alignments. Single gene alignments were concatenated with Sequence Matrix

(Vaidya et al., 2011). For the nuclear genome, we prepared and separately analyzed three different datasets consisting of (i) the concatenation of orthologs, (ii) the concatenation of the four-fold degenerate sites of the orthologs, and (iii) the amino acid sequences of the orthologs. As mentioned above, the analysis of nucleotide and amino acid datasets allowed us to compare their performances, especially in the resolution of deep phylogenetic relationships. The rationale behind analyzing the four-fold degenerate sites is that such sites are nucleotide positions where all changes are synonymous and assumed to be neutrally evolving; as such, they are well-suited for phylogenetic reconstruction and for estimating species divergence time (Edwards, 2009).

2.5 Maximum-likelihood Phylogenetic Inference and Tree Comparison

Maximum-likelihood phylogenies for all the mitochondrial and nuclear datasets were obtained using the software RAxML v.8.1.20 (Stamatakis, 2014). The best ML trees were inferred as follows: 1) assuming a General Time Reversal (GTR) model of sequence evolution with a gamma-distribution model of rate heterogeneity for nucleotide sequences, and 2) using the automatic selection of protein sequence evolution (PROTGAMMAAUTO) for amino acid sequences, always taking into account gene partitioning. One hundred pseudo-replicates were generated by applying a bootstrap approach to test the reliability of the best trees. Using the same setting as above, we inferred ML trees for each independent mitochondrial and nuclear gene. In addition to the concatenated analysis, we also estimated species trees with ASTRAL-III v5.6.3 (Zhang et al., 2018), which uses quartet frequencies found in gene trees and has shown to be accurate also in the presence of incomplete lineage sorting (e.g. Davidson et al., 2015). We ran ASTRAL using default settings and individual nuclear gene trees (obtained from both nucleotide and amino acid sequences) as input.

We used Densitree 2.2.5 to visualize phylogenetic discrepancies between single gene trees of the nuclear genome (Bouckaert, 2010). We made the best nucleotide ML tree of each gene ultrametric, using the function *chronos* in the R package APE (Paradis et al., 2004) and then plotted them . We also quantitatively evaluated the discrepancies between the single mitochondrial gene trees, that were reconstructed excluding *D.* cf. *obtusa*, which was missing in the nuclear matrix, and the best ML nuclear phylogeny using Ktreedist (Soria-Carrasco et al., 2007), which measures the differences in the relative branch length and topology between phylogenetic trees.

2.6 Bayesian Estimate of Species Divergence Time

We applied BEAST 2.4.5 (Bouckaert et al., 2014) to estimate the ages and confidence intervals of branching events using two independent approaches. In the first approach, we included in the phylogeny all relevant and available fossil records for Anomopoda as calibration priors. Our calibrations were based on the following argumentation: The oldest unambiguous Cladocera fossils are from Mesozoic times, in the Early Jurassic (174 -201Mya; Kotov, 2007; Van Damme and Kotov, 2016), and the first fossils of the order Anomopoda date to the end of the Jurassic

(Jurassic/Cretaceous boundary; 145Mya; Kotov and Taylor, 2011). It is likely, therefore, that the Anomopoda ancestor and the divergence of the two suborders, Aradopoda and Radopoda (Dumont and Silva-Briano, 1998; Kotov, 2013), occurred well before the Jurassic age, and even before the Mesozoic era, though there is no fossil evidence (Van Damme and Kotov, 2016).

For our analysis, we used a uniform prior on the root describing the most recent common ancestor (MRCA) of all Anomopoda—in our case the split between the Radopoda, represented by the Bosminidae and the Aradopoda, which include Daphniidae and Moinidae—and placed it at the most recent divergence time suggested by fossil evidence—the Late Jurassic (145-163.5 Mya). Although the split Bosminidae/Aradopoda is likely to be older, we consider the Late Jurassic prior on the root the most conservative and reliable prior substantiated by fossil evidence in describing the MRCA of all Anomopoda. We used four additional Log-normal priors derived from fossil evidence as reported in Van Damme and Kotov (2016): one describing the Moinidae-Daphniidae split (at least 145 Mya, coded in BEAUti as M=2, S=1.25, offset=145, which translates into median 152 Mya, 95%CI 146-203 Mya); the second describing the *Simocephalus-Daphnia* split (at least 145 Mya); a third describing the *Ceriodaphnia-Daphnia* split (at least 118 Mya, coded in BEAUti as M=2, S=1.25, offset=118, which translates into median 125 Mya, 95 %CI 119-176 Mya), and finally, the subgenus *Ctenodaphnia- Daphnia* split (at least 145 Mya). We specified constraints of monophyly for all fossil calibrated nodes; this ensured, for example, that the Moinidae-Daphniidae split must have occurred earlier than the *Ctenodaphnia-Daphnia* split.

As the origin of the Anomopoda likely pre-dates the Late Jurassic, and several extant families may have been established even before the Mesozoic (Van Damme and Kotov, 2016), we explored the timing of the nodes by estimating the MRCA of the Anomopoda at two earlier hypothetical dates: one in the Early Jurassic, the appearance of first Cladocera fossils, 174-201 Mya (Kotov, 2007), and another in the Permian, a Paleozoic estimate, when Cladoceromorpha fossils of the order Cyclestherida first appeared, about 252-299 Mya (Raymond, 1946; Sun et al., 2016). In BEAUti, we set trees and site models as linked for the mtDNA concatenated phylogeny based on the results of site JModelTest 2 (Darriba et al., 2012). All partitions, in fact, resulted in the same substitution model (GTR + GAMMA + Invariant sites). We specified unlinked clock models, allowing partition-specific estimates of substitution rates using a strict clock.

For the nuclear analyses, an unpartitioned alignment of four-fold degenerate sites was used with a "GTR + GAMMA + Invariant sites" substitution model. Computational constraints did not allow us to analyze a partitioned dataset of hundreds of orthologs in BEAST. The concatenation of four-fold degenerate sites of the identified orthologs was used to estimate a genome-wide substitution rate, while the fossil priors were used to calibrate the molecular clock. The BEAST analyses were run with a MCMC chain length of 10,000,000, after discarding the first 10 % of iterations as burn-in, and parameter sampling every 1,000 generations. We examined the log files with Tracer v1.6 (Rambaut et al., 2014) to evaluate the convergence of the analysis and to ensure that the effective sample size (ESS) of the

parameters was greater than 200.

In our second approach, we estimated sequence divergence time using a strict molecular clock based on the neutral substitution rates derived from mutation-accumulation experiments in sexual lines of *D. pulex.* Substitution rates of 2.0×10^{-8} per nucleotide per generation (Xu et al., 2012) and 4.33×10^{-9} (Keith et al., 2016) per nucleotide per generation were used for the mitochondrial and nuclear phylogeny, respectively. As *Daphnia* undergoes sexual reproduction about once per year (Lampert, 2011), we used this rate to test time-based estimates (Mya).

3. Results

3.1 Mitochondrial Phylogenies

We found no inconsistencies between the four individually assembled genome sequences for each clone. The sequences of all genes (PCGs and rRNA) could be unambiguously aligned (ENA study number ERP109988, project: PRJEB27855, accession numbers: LS991483-LS991524). The best ML trees obtained with the three different mtDNA datasets had identical topologies (Fig. 1). The dataset consisting of all PCGs and the two structural rRNA genes was the most highly supported (an average bootstrap value > 92), although all three phylogenies showed high bootstrap values. The mtDNA amino acid dataset showed the lowest topological support (average bootstrap value of 86.7), with bootstrap values similar to the nucleotide datasets for deep nodes. Our results confirmed the subgenera *Daphnia* and *Ctenodaphnia* as significantly supported monophyletic groups. *D. longispina,* the only representative in our study of this widely studied species group, appeared to be a sister taxon to the *D. pulex*/*D. obtusa* groups.

Within the subgenus *Ctenodaphnia,* we observed that the Australian *D.* cf. *carinata* was sister lineage to the *D. similis/D. sinensis/D. lumholtzi* clade. In such clade, the three species constitute clearly distinct groups with little differentiation after an initial branching off by *D. lumholtzi*. In *D. magna,* we observed the split between Western Eurasia and East Asian (and North American) lineages that Fields et al. (2018) and Bekker et al., (2018) previously reported. The mtDNA phylogeny also showed some unexpected results. We found that *D.* cf. "*similis*" appeared to be a sister taxon to *D. magna*. *D.* cf. "*similis*" does not cluster with other species of the *D. similis* group. In fact, this unnamed taxon from Canada likely belongs to the *D. exilis* group (Adamowicz et al., 2009; Popova et al., 2016). We found that *D. hispanica* clusters with *D. atkinsoni* (Fig. 1) and does not appear close to *D. barbata*. Indeed, the bootstrap values for *D. hispanica* appearing near *D. barbata* in a previous mtDNA phylogeny had low support (Adamowicz et al., 2009).

The same pattern was confirmed when we analyzed the concatenation of the mtDNA genes 16S, 12S and COI of the taxa in our study compared with the taxa in Adamowicz et al., (2009) (Supplementary Fig. S1). Our best 16S/12S/COI ML tree was poorly supported, especially for the majority of internal nodes (Supplementary Fig. S1). We also compared our sequences with those published by Popova et al., (2016), which confirmed that our populations of "true" *D. similis* form a monophyletic clade when analyzed simultaneously with populations from Israel, Russia and several European locations (bootstrap values of 100 and 97, for 12S and

COI respectively, Figures S2 and S3). These populations contain several clones from the vicinity of the type locality of *D. similis* in Israel. Clones of *D.* cf "*similis"* from the Nearctic analyzed here are genetically very similar, or in some cases identical, to the *D.* cf. "*similis*" clones analyzed for COI gene in (Popova et al., 2016)(Supplementary Fig. S3). Since we never observed monophyly with *D. similis*, and found a sequence dissimilarity of about 12 and 18 % for 12S and COI, respectively, against *D. similis*, our mtDNA data indicate that *D.* cf. "*similis*" represents a different species from *D. similis*. This confirms earlier observations; in fact, the New World *D.* cf. "*similis*" belongs to the *D. exilis*-group instead (Adamowicz et al., 2009; Popova et al., 2016).

3.2 Nuclear Phylogenies

We obtained a draft nuclear genome assembly for each of the 38 cladoceran clones sequenced for this study. The quality of the genome assemblies varied substantially as assessed by scaffold number constituting the assembly (Table S1). However, the genome assessment performed with BUSCO retrieved the great majority of complete single-copy arthropod genes (on average about 94 %, Supplementary Fig. S4). These genes were used to build a matrix for genome-wide nuclear phylogenetic analyses. After removing ambiguously aligned positions and allowing no gaps in the final alignments, we obtained a set of 636 orthologs, with only 4 % missing data. The average gene length was about 800 bp (range 138 to 3366 bp, ENA study number ERP109988, project: PRJEB27855, accession numbers: LR000001-LR025064).

The best ML trees obtained from the three concatenated nuclear datasets had almost identical topologies. The two nucleotide trees were identical (Fig. 2), while the amino acid tree showed some variations in the relationships of non-*Daphnia* taxa. More specifically, *Scapholeberis* formed a separate branch (nucleotide trees) or grouped together with *Megafenestra* and *Simocephalus* (amino acid tree; Supplementary Fig. S5). All three phylogenies were highly supported, showing bootstrap values of 100 in almost all internal nodes. A bootstrap value of 90 was observed in the amino acid tree at the branching of non-*Daphnia* species (Supplementary Fig. S5). The ASTRAL species trees obtained with nucleotide and amino acid sequences were identical and in complete agreement with the ML topology resulting from the concatenated amino acid sequences (Supplementary Fig. S5). Local posterior probabilities were high across the ASTRAL species trees, with only one internal node with posterior probability less than 1.0. Such node describes the phylogenetic relationships among non-*Daphnia* genera (Figure 2, Supplementary Fig. S5). Throughout most of the tree, the nuclear (nucleotide) topology resembled the mtDNA phylogeny, showing dissimilarities only in the position of the non-*Daphnia* taxa. The mtDNA tree showed both *Simocephalus* and *Ceriodaphnia* as the sister taxa to *Daphnia*, whereas the nuclear tree showed only *Ceriodaphnia* as a *Daphnia* sister clade. Also the nuclear phylogeny showed *Scapholeberis* (which groups with *Megafenestra*, the other representative of the subfamily Scapholeberinae in the mtDNA tree), as the first Daphniidae branching off and as sister to all other Daphniidae. Additionally, in the nuclear trees, *Megafenestra* and *Simocephalus* always cluster together (Fig. 2, Supplementary Fig. S5).

3.3 Tree Comparison

We performed two types of tree comparisons for the mitochondrial and nuclear genomes. First, we graphically compared the topologies obtained from each single nuclear gene. Second, we quantitatively assessed the discordance between individual mtDNA gene trees and the species tree, considering the topology of the ML nuclear phylogeny as being the most reliable in describing the relationships within the Daphniidae.

The DensiTree plot shows that, overall, there is a marked discrepancy among gene tree topologies at the nuclear level (Fig. 3). In fact, after drawing all the single gene trees, some densely colored areas become apparent, especially in correspondence with external nodes where many trees agree on the topology and branch length. There were also, however, places in the plot where webs of lines are visible, for example in the relationships among taxa of the subgenus *Ctenodaphnia*, indicating low levels of concordance among gene trees (Fig. 3).

We also observed a remarkable discordance between individual mtDNA gene trees and the species tree, with K-scores varying from 0.40 for 12S to 0.67 for atp8. Interestingly, COI, a gene included in several phylogenetic, taxonomical and molecular studies concerning *Daphnia* (Adamowicz et al., 2009; Petrusek et al., 2008; Popova et al., 2016), shows one of the highest K-scores (0.60, Table S2), suggesting a strong discrepancy between the COI gene tree and the species tree based on nuclear genes.

3.4 Divergence Time Estimation

We performed divergence time estimation for the mtDNA and nuclear phylogenies independently, using both fossil record information and substitution rates. In all cases, we observed high ESSs (>200) for the parameters. Our assessment with Tracer also showed that all the analyses had converged. To directly compare the resulting time-calibrated phylogenies, we selected six relevant nodes: the MRCA of *Daphnia-Ctenodaphnia*; the MRCA of the *D. longispina - D. pulex* group; the MRCA of *D. magna* and the *D. similis* group (*D. similis-sinensis-lumholtzi*); the MRCA of *D. magna*/*D.* cf "*similis*"; the MRCA of *D. similis*/*D. sinensis;* and the MRCA *D. magna* (Europe)/*D. magna* (Asia) (Table 2). Table 2 summarizes the divergence times observed in our analyses. We base the following comparisons on estimates obtained using a uniform prior that places the MRCA of all Anomopoda in the Late Jurassic period (145-163.5 Mya), the most conservative estimate supported by fossil data. The latter resulted in the youngest age in comparison to when the MRCA of all anomopods was hypothetically dated, at least in the Early Jurassic or in the Permian, though there is no fossil evidence (see Table 2).

The time-dated mtDNA datasets with fossil records and substitution rates led to very different estimates of branching events. For example, the MRCA of the sub-genera *Daphnia-Ctenodaphnia* was dated 145.2 Mya (95%HPD: 145.0-145.5) when fossil priors were included; however, the same node was dated 26.9 Mya (26.1-27.8) using a fixed substitution rate (Table 2). This latter estimate obtained with substitution rates must be considered as unfounded, however, as the fossil record confirms that the two subgenera, *Daphnia-Ctenodaphnia*, already co-existed at least

145 Mya (Kotov and Taylor, 2011). In line with these marked discrepancies, the estimated substitution rates obtained for single mtDNA genes ranged between 2.36 x 10^{-9} (12S) and 6.18 x 10^{-9} (atp8). This was, on average (4.12 x 10^{-9}, Table S3), about one order of magnitude slower that what was documented (2.0 x 10^{-8}) for *D. pulex* (Xu et al., 2012). When we used the two earlier, hypothetical priors on the root, we observed a similar pattern (Early Jurassic: between 2.14 x 10^{-9} to 5.68 x 10^{9}, average: 3.76 x 10^{-9}, Table S4; Permian: between 1.74 x 10^{-9} to 4.63 x 10^{-9}, average: 3.07 x 10^{-9}, Table S5).

The marked variation in time estimates for branching events that we observed in the mtDNA analyses are not as evident in terms of discrepancy in time estimates between methods as they are for the genome-wide nuclear phylogenies. Here, the MRCA of the sub-genera *Daphnia-Ctenodaphnia* was dated 145 Mya (95 %HPD: 145.0-145.1) when fossil priors were included, and 102.3 Mya (100.5-104.2) using a fixed substitution rate—still a substantial underestimation (Table 2). The genome-wide substitution rate when fossil priors were included was found to be similar (5.88 x 10^{-9}; 95 % HPD: 5.78 x 10^{-9} – 5.98 x 10^{-9}) to what was documented for *D. pulex* (4.33 x 10^{-9}). As before, when we used the two earlier priors on the anomopod root, we observed a consistent pattern in the nuclear substitution rate estimations (i.e. the older the prior, the slower the mutation rate; Early Jurassic: 5.01 x 10^{-9}, 95 % HPD: 4.92 x 10^{-9} – 5.10 x 10^{-9}; Permian: 3.93 x 10^{-9}, 95 % HPD: 3.86 x 10^{-9} – 4.00 x 10^{-9}).

4. Discussion

High-throughput/next-generation sequencing approaches now enable us to assemble entire genomes and extract genes to explore genome-wide relationships between taxa. In this paper, we have introduced calibrated molecular phylogeny for the family Daphniidae based on entire mitochondrial genomes and 636 nuclear genes. This is the first time a phylogenomic approach has been applied in the Branchiopoda. Although genomic methods have begun to be applied in other microcrustacean groups (i.e. Copepoda, Eyun 2017), major zooplankton groups are still lagging behind in these modern approaches, which are more commonly used for the reconstruction of phylogenies in diverse taxa.

4.1 Phylogenetic Relationships between non-Daphnia Genera

Whereas the mtDNA and nuclear topologies showed complete agreement about the position of species and species groups within the *Daphnia* genus, the phylogenetic relationships between non-*Daphnia* taxa in the family were not as clearly resolved. These differing topologies might, however, reveal important findings for the evolutionary history of the family Daphniidae. The positions of *Scapholeberis,* as a clade furthest from *Daphnia,* and of *Ceriodaphnia,* as (or in) a sister clade to *Daphnia*, remained consistent in all analyses. The position of both these genera relative to *Daphnia* was unresolved in previous phylogenies (not focused on Daphniidae) that combined mitochondrial and nuclear genes, where both of these genera showed similar positioning (e.g., deWaard et al., 2006, Richter et al., 2007).

The position of *Scapholeberis* and its strong divergence in the Daphniidae is

intriguing and runs counter to classical assumptions about this genus: It is generally assumed that this genus' morphological adaptation to an unusual hyponeustonic lifestyle living below the surface film of the water (such as a specially adapted infolded ventral rim of the valves carapace margin), is an advanced, derived state in the family, forming a synapomorphy with *Megafenestra* (Dumont and Pensaert, 1983; Fryer, 1991). However, the phylogeny presented here leads us to propose the exact opposite of this "classical" view (see Daphniidae phylogeny in (Dumont and Pensaert, 1983): Fig. XXII). Indeed, we suggest that *Scapholeberis* may be one of the earliest offshoots in the Daphniidae tree. The morphologically very similar *Megafenestra* (e.g. Alonso, 1996), which was originally placed within the genus *Scapholeberis* by taxonomists (and here incorporated for the first time in a wider molecular phylogenetic context), shows unexpectedly large molecular differences from *Scapholeberis*. This confirms our suggestion that the only two hyponeustonic specialist genera in the Daphniidae likely diverged early in the evolutionary history of the family, to such an extent that they do not cluster together in the nuclear gene tree.

From the nuclear fossil-calibrated phylogeny it seems that the separation of the *Scapholeberis* lineage happened well before the divergence between the genera *Simocephalus* or *Ceriodaphnia* from *Daphnia* (Fig. S8). The remarkable morphologies of *Scapholeberis* and *Megafenestra*, such as the special valve rim and associated rectangular body shape, are considered of secondary origin in the family (Fryer, 1991). The two genera show a different degree of adaptation, with *Scapholeberis* being the more specialized (Dumont & Pensaert, 1993). Due to the large genetic distance observed here, we cannot exclude a potential independent evolution of both from a general daphniid stock to the hyponeustonic lifestyle, despite their external similarities.

Because of the discrepancy between the mitochondrial and nuclear gene trees, the exact positions of *Simocephalus* and *Megafenestra* remain unclear. In the nuclear tree (Fig. 2; S5), the position of *Simocephalus*, forming a clade with *Megafenestra*, even disrupts the classification of the two subfamilies, which would suggest a paraphyly in the Daphniinae as well as in the Scapholeberinae. In the mtDNA tree, however, *Simocephalus* clusters with *Ceriodaphnia*, and the subfamilies appear monophyletic. *Simocephalus* has a lifestyle that differs from closely related taxa: It is neither constantly free-swimming (like *Daphnia* and *Ceriodaphnia*), nor hyponeustonic (like *Megafenestra* and *Scapholeberis*). Rather, these animals attach themselves dorsally to surfaces most of the time, using the second antennae which included specially adapted setae; *Simocephalus* behaves more like a sedentary animal exploring different niches from *Daphnia* (Fryer, 1991; Orlova-Bienkowskaja, 2001).

Finally, because *Ceriodaphnia* has a consistent phylogenetic position in all analyses (with or without *Simocephalus*), we hereby consider it the most reliable sister lineage to the genus *Daphnia*. Both these genera diversified mainly as pelagic free-swimming taxa (Fryer, 1991). The placement of *Ceriodaphnia* as a direct sister clade to *Daphnia* has useful implications for ecology and ecotoxicology, as taxa of both genera are widely used as experimental organisms. Future increased taxon

sampling of non-*Daphnia* lineages for Daphniidae phylogenomics will help to provide a better understanding of their evolution.

4.2 The Subgenus Ctenodaphnia

The mtDNA and nuclear topologies showed complete agreement for the position of taxa within the genus *Daphnia*. Only a portion of the diversity of species and species groups in this diverse genus is represented here, yet some observations can be made. The African endemics *D. barbata* and *D. dolicocephala* appear together near the root of the subgenus *Ctenodaphnia*, followed by a second branch that includes *D. atkinsoni* and the Iberian endemic *D. hispanica*. Adamowicz et al. (2009) described a similar position for *D. barbata,* although with low support in their tree; the South African *Daphnia dolicocephala* has been included here in a larger molecular phylogeny for the first time. This species is part of a group of African endemic daphniids in need of revision (Van Damme et al., 2013). Preliminary data of limb morphologies in *Ctenodaphnia* suggest that the lineages *D. atkinsoni, D. hispanica, D. dolicocephala* and *D. barbata* have some similar features (limbs of "*atkinsoni*-type"), whereas the limbs of *D. magna, D. similis, D. lumholtzi* and *D. carinata* have different sets of morphological characters (Alonso, 1985; Glagolev and Alonso, 1990; and references therein). The present phylogenomic analysis does not contradict such larger divisions in *Ctenodaphnia*. The close relationship of the Iberian endemic *D. hispanica* to *D. atkinsoni* also corroborates the thoracic limb characters; indeed, *D. hispanica*, whose limbs are morphologically closely related to those of *D. chevreuxi* (not included here), harbors several morphological features only found in non-*Daphnia* genera, which suggests a relatively basal position in the subgenus (Glagolev and Alonso, 1990). The comparison of our phylogenomic analysis with the scarce available data on limb features suggests that limb morphologies may contain a powerful phylogenetic signal for the deeper systematics in *Daphnia*, as larger clades match general relationships suggested by the preliminary morphological data (e.g. Glagolev and Alonso, 1990).

Our analysis seems to suggest that all *Ctenodaphnia* clades in basal position are entirely made up of Old World taxa; this is however a result of a sampling bias. Our analysis lacks important lineages such as *D. ephemeralis*, a critical Nearctic taxon at the basis of the *Ctenodaphnia* tree in previous phylogenies (Adamowicz et al., 2009); also, the Australian endemic *D. pusilla*-group was not sampled and only one Australian species from the large *D. carinata* group was included in this study (i.e. *D.* cf. *carinata*). However, the Old World species, in particular the endemic *Ctenodaphnia* African species, merit further analysis.

4.3 The Sister Clade of D. magna is not D. similis

In the clade leading to *D. magna,* the Northern Nearctic *D.* cf. "*similis"* appears as the closest sister species to *D. magna* among the taxa included here. This taxon is not directly related to other *D. similis*-like species (*D. similis, D. sinensis, D. lumholtzi*), and is instead likely part of the *D. exilis* group (Popova et al., 2016). In their study, Popova et al. (2016) cautioned about using *D. similis* s.str. prematurely as a

comparative genomic and ecotoxicological model to *D. magna*. Our phylogenomic analysis supports this caution and shows that better candidates may be found among the *D. exilis* group, which is restricted to the New World (Adamowicz et al., 2004). We consider this group as a more suitable sister lineage to *D. magna* for now. The external morphological similarities of *D.* cf. *"similis*" and *D. similis* seem to be a result of convergence.

The age of divergence in this clade (between *D.* cf. "*similis*" and *D. magna*) is at least 10.6 Mya, according to the conservative estimate using the fossil-calibrated nuclear phylogeny. The split between *D. magna* and *D.* cf. "*similis*" is among the relatively most recent splits in our *Ctenodaphnia* tree, taking place around the same time as the *D. similis* and *D. lumholtzi* split and more recently than the separation of the latter species from *D. atkinsoni* and *D. barbata,* which happened much earlier, at least 54.4 Mya according to the conservative fossil-calibrated nuclear gene tree. Nuclear data from *D. magna* suggest that the already reported split between Western Eurasia and Eastern Asia populations (Bekker et al., 2018; Fields et al., 2015; Fields et al., 2018) might have occurred at least 1.2 Mya.

The *D. similis*-group, which is represented in our phylogenomic analysis by two Old World taxa—*D. similis* s.str. and the recently separated *D. sinensis*—shows that these species delimits are well supported with little intraspecific divergence. They group reliably with *D. lumholtzi,* consistent with previous studies and morphology (Adamowicz et al., 2009; Popova et al., 2016). We reject the suggestion of Popova et al., (2016) that *D. barbata* is a *D. similis*-like taxon; in all our analysis, it is far from *D. similis*.

However, our results may be affected by sampling bias. The Australian *D. carinata*-like stock may well contain close sister lineages to *D. magna,* yet they are not studied in more detail here. *D.* cf. *carinata* is sister-group in our analysis to *D. similis*/*D. sinensis*/*D. lumholtzi*, while it appears as a sister lineage to *D. exilis* in Adamowicz et al., (2009). *D.* cf. *carinata* is the sole representative here of a large group of Australian endemics including *D. jollyi, D. cephalata, D. longicephala* and others, several of which are well known for extreme cyclomorphosis (e.g. Hebert, 1978).

4.4 Gene Tree Discordance

Because mitochondrial genomes typically do not undergo recombination, mitochondrial genes are often assumed to reconstruct the same topology (Rokas et al., 2003). However, discrepancies between phylogenies based upon different mitochondrial regions may occur due to the way such regions accumulate substitutions (e.g. Meiklejohn et al., 2014; Zhang et al., 2013). Consequently, if researchers have to select one or two mtDNA markers to sequence, their choice of mitochondrial genes greatly influences the reliability of the resulting topology. This discrepancy becomes especially relevant when the aim is to elucidate phylogenetic relationships between species within a old genus, like *Daphnia* where mtDNA is limited in dealing with deep divergence times due to saturation effects (Rubinoff and Holland, 2005). In most previous phylogenetic studies on *Daphnia*, the genes COI, 12S and 16S were selected to represent the mitogenomes as a whole. Our results suggest that 12S and 16S should indeed be the first choices because they

showed the lowest K-scores. However, other mitochondrial genes are preferable to COI in phylogenetic studies, as this gene had one of the highest K-scores. Nevertheless, given the abundance of Daphniidae COI sequences in reference databases, this gene may serve as a useful genetic marker for DNA barcoding (but see Thielsch et al., 2017)

We observed gene tree discordance among the nuclear genes. Because nuclear genes typically have a limited number of variable characters required to provide a phylogenetically informative signal and build an accurate phylogeny, multigene phylogenies are generally preferred. The markers we used for our phylogenetic analysis were selected from a set of genes known to be highly conserved across a broad range of taxa. Such highly conserved genes are the preferred markers for phylogenetic reconstruction, especially when deep nodes are of interest (e.g. see Nosenko et al., 2013; Zeng et al., 2014; Zhang et al., 2012). Another explanation for nuclear gene discordance is introgression. However, we have no evidence that introgression contributed to nuclear gene tree discordance in our study. Among the species included here, hybridization has been well studied in the *D. pulex* (Xu et al., 2015) and in the *D. longispina* complexes (Schwenk and Spaak, 1997). However, sister species of *D. longispina* were not included in our study. Little is known about hybridization in other *Daphnia* species complexes, although it is known to occur in the *D. carinata* and *D. obtusa* complexes as well (Schwenk and Spaak, 1997), though both are represented here by one species, respectively. The approach used in our study of concatenating multiple unlinked genes should reduce, or even overcome, discordant gene topologies, since the discrepancies between single nuclear gene trees and species tree are not expected to occur in the same way for the majority of the genes (Wiens et al., 2010). Incomplete lineage sorting is another alternative explanation for gene tree discordance (Maddison, 1997). Within the genus *Daphnia*, the species trees obtained using the multi-species coalescent method implemented in ASTRAL were highly supported and identical to the concatenated nuclear ML trees, suggesting that the *Daphnia* phylogeny observed here is likely not biased by incomplete lineage sorting. The discrepancy in the phylogenetic relationships among non-*Daphnia* genera between the nucleotide RAxML species trees and the ASTRAL species trees, also highlighted by the consensus tree of DensiTree, suggests that further sampling of non-*Daphnia* genera may be necessary to provide a better understanding of their relationships.

4.5 Divergence Time Estimates

Previous studies have used the molecular clock approach based on a few markers to estimate the divergence time among taxa within the family Daphniidae. However, the age estimate of branching events differs widely in these studies. Our study was not able to fully resolve the known discrepancies regarding the timing of branching events, even though we performed multiple analyses based on different criteria and genetic markers.

Our mtDNA substitution rate-calibrated phylogenies substantially underestimated the time of the *Daphnia*/*Ctenodaphnia* split. The same issue was observed for other branching events older than about 20 Mya and is typical of the

mtDNA "saturation effect," described by DeSalle et al. (1987). The saturation effect can cause homoplasy, making it especially problematic to compare, for example, *Drosophila* species that diverged > 20 Mya (Barrio et al., 1992). On the other hand, we observed relatively consistent time estimates for nodes dated < 15 Mya among mitochondrial and nuclear substitution-rate-based calibration and nuclear-fossil-based calibration. The fact that these multiple independent analyses recover similar time estimates for more recent branching events suggests that the age of such nodes may be useful as minimal estimates. When the mtDNA tree is calibrated with fossil records, the estimates of divergence time are considerably older than when a calibration based on the substitution rate was applied —at least 100 Mya for the divergence between the *D. longispina* and *D. pulex* groups for example, congruent with previous calculations based on mtDNA (16S; Schwenk et al., 2000; Taylor et al., 1996). Although these mtDNA estimates are theoretically possible, however, the discrepancies between the fossil- and substitution rate-calibrated mtDNA phylogenies and the fact that mtDNA describes the history of only one single locus must be taken into account.

Unfortunately, we lack reliable fossils to determine the first appearance of *Daphnia* species groups, which would strengthen the support for minimal ages of divergence (Van Damme and Kotov, 2016). We know from the fossil record that *D. "pulex"* and *D. "magna"*-like ephippia (resting eggs) were found in the Eocene palaeolake Messel (ca. 47 Mya; Lutz, 1991) and that well-preserved *D.* "*magna-similis*"-like ephippia and parthenogenetic females were found in Cenozoic German paleolakes (ca. 24 Mya and 17-15 Mya, respectively; Kotov and Wappler, 2015). However, the diagnostic resolution of these ephippia's morphological features does not allow a full identification at the species level. Although the fossil-calibrated nuclear gene time estimates potentially underestimate divergence times (e.g., likely in the timing of the *D. magna-D. similis* split), the mtDNA estimates may potentially overestimate them (e.g., likely in the separation of *D. magna* populations between Europe and Asia; Table 2). None of these time-calibrated estimates directly contradicts the fossil record, as there are no reliable records for species groups, yet some estimates may be less plausible. Indeed, any divergence estimates for *Daphnia* should not be taken as definitive, considering how the results vary depending on the method used. Even fossil calibrations are tentative because of the scarcity of nodes that can be fossil-calibrated. Realistically, only the minimal time estimates based on molecular clocks can be evaluated carefully, and these should always be assessed against an updated fossil record. More paleontological data is necessary to increase the resolution and facilitate the interpretation of such clocks. Even biogeographical data (distribution patterns) is of little help in estimating ages in the genus (Popova and Kotov, 2013).

The nuclear genes proved to be more consistent than the mtDNA when time estimates of fossil- and substitution rate-calibrated phylogenies were compared. Our results contradict what has been suggested by Haag et al. (2009), who dated the split of the subgenera *Daphnia*/*Ctenodaphnia* at 7.6 Mya, much more recently than what has been suggested by fossil records (at least 145 Mya) and the nuclear substitution rate clock shown in our study (at least 100 Mya). This discrepancy may

be because, in order to calculate divergence times, Haag et al. (2009) used (i) the mutation rates of *Drosophila melanogaster* and *Caenorhabditis elegans*, since the mutation rate in *Daphnia* was not yet known at their time of publication; (ii) only eight nuclear genes and (iii) a mathematical formula that included assumptions about biological parameters such as the effective population size of *Daphnia* and the number of sexual and asexual generations per year, which are difficult to estimate confidently (Haag et al. (2009) assumed them to be 10). The fact that we obtained relatively similar estimates using two independent approaches—substitution rate (which requires an assumption on the number of generations per year), and fossil data—likely suggests that our choice of using one sexual generation per year (following Lampert, 2011) has some merit.

Using different prior maximum age constraints in a Bayesian analysis can change the time estimates of branching events in a phylogeny (Cracraft et al., 2015; Warnock et al., 2012). Given that no fossils are available to calibrate the MRCA of all Anomopoda accurately, we used three different plausible priors at the root of our tree. When we shifted the prior on the MRCA of all Anomopoda from the Late Jurassic (most recent and most conservative; 145-163.5 Mya) to the Early Jurassic (174-201 Mya) and Permian (252-299 Mya), we found that branching events were estimated to have happened much earlier. It is complicated to speculate on which divergence times are the most realistic because of the lack of fossil evidence older than the Late Jurassic. However, it is likely that the Anomopoda ancestor lived even before the Mesozoic (Van Damme and Kotov, 2016). Nevertheless, when our study estimates the minimum divergence times between taxa using most conservative and reliably dated Late Jurassic prior, we see divergence of most *Daphnia* species as taking place at least in the Cenozoic (since ca. 66 Mya).

5. Conclusions

Our phylogenomic study of the family Daphniidae indicates that topologies obtained from mtDNA and nuclear genomes are similar for younger nodes, but diverge at some of the deep nodes. This is consistent with the suggestion that saturation, causing a loss of signal, occurs in the mtDNA, which represents a single locus (Rubinoff and Holland, 2005). We also observed a major discrepancy in the temporal estimation of branching events for the mtDNA between fossil- and substitution rate-calibrated trees. Our analysis uncovered a new sister taxon to *D. magna, D.* cf. "*similis*" (*D. exilis* group) from Western Canada, which can be considered an appropriate outgroup in future phylogenetic and comparative genomic investigations. Also, we established that *Ceriodaphnia* is the most reliable sister genus to *Daphnia,* and that the morphologically similar *Scapholeberis* and *Megafenestra* show surprisingly deep genetic divergence. We believe that our study provides a solid phylogenetic baseline for future studies involving species of the family Daphniidae and illustrates the power, but also some limitations, of whole genome sequence data for phylogenomic analyses.

Acknowledgements

We thank Jürgen Hottinger, Urs Stiefel and Michelle Krebs for their help in the lab.

We thank the following people for help in providing samples for this study: Frida Ben-Ami, Maciej Bartosiewicz, Jürgen Hottinger, Lukas Schärer, Yan Galimov, Ian Gardiner, France Dufresne, Jason Andras, Christoph Haag, Tom Little, Joachim Mergeay, Adam Petrusek, Alexey Kotov, Maria José Caramujo R. de Carvalho, Jim Elser, John Havel. We also thank the Editor and two anonymous reviewers whose suggestions helped improve this manuscript. We are grateful to Darren Obbard, Tom Little and Seanna McTaggart (University of Edinburgh), supported by the NERC grant NE/J010790/1, who contributed to the sequencing of some *Daphnia* clones. We thank the Genomics Facility of Basel (DBSSE), in particular Christian Beisel and Elodie Burcklen, for next generation sequencing service. Part of the analyses for this study was performed using the sciCORE (http://scicore.unibas.ch) scientific computing core facility at the University of Basel. This work was supported by the Swiss National Science Foundation (grant number: 310030B_166677) and the University of Basel.

References

Abascal, F., Zardoya, R., Telford, M.J., 2010. TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 38, 7–13. https://doi.org/10.1093/nar/gkq291

Abele, L.G., Spears, T., 2000. Branchiopod Monophyly and Interordinal Phylogeny Inferred from 18S Ribosomal DNA. J. Crustac. Biol. 20, 1–24. https://doi.org/10.1163/20021975-99990012

Adamowicz, S.J., Hebert, P.D.N., Marinone, M.C., 2004. Species diversity and endemism in the *Daphnia* of Argentina: a genetic invasion. Zool. J. Linn. Soc. 140, 171–205. https://doi.org/10.1111/j.1096-3642.2003.00089.x

Adamowicz, S.J., Petrusek, A., Colbourne, J.K., Hebert, P.D.N., Witt, J.D.S., 2009. The scale of divergence: A phylogenetic appraisal of intercontinental allopatric speciation in a passively dispersed freshwater zooplankton genus. Mol. Phylogenet. Evol. 50, 423–436. https://doi.org/10.1016/j.ympev.2008.11.026

Alonso, M., 1996. Crustacea, Branchiopoda. En: Fauna Ibérica, vol.7. Ramos, M.A. et al. (Eds.). Museo Nacional de Ciencias Naturales. CSIC. Madrid. 486 pp

Alonso, M., 1985. *Daphnia* (*Ctenodaphnia*) *Mediterranea*: A new species of hyperhaline waters, long confused with *D.* (*C.*) *Dolichocephala* Sars, 1895. Hydrobiologia 128, 217–228. https://doi.org/10.1007/BF00006817

Alric, B., Möst, M., Domaizon, I., Pignol, C., Spaak, P., Perga, M.E., 2016. Local human pressures influence gene flow in a hybridizing *Daphnia* species complex. J. Evol. Biol. 29, 720–735. https://doi.org/10.1111/jeb.12820

Arbogast, B.S., Edwards, S. V., Wakeley, J., Beerli, P., Slowinski, J.B., 2002. Estimating Divergence Times from Molecular Data on Phylogenetic and Population Genetic Timescales. Annu. Rev. Ecol. Syst. 33, 707–740. https://doi.org/10.1146/annurev.ecolsys.33.010802.150500

Barrio, E., Latorre, A., Moya, A., Ayala, F.J., 1992. Phylogenetic reconstruction of the *Drosophila obscura* group, on the basis of mitochondrial DNA. Mol. Biol. Evol. 9, 621–35.

Barry, M.J., 2000. Inducible defences in *Daphnia* : responses to two closely related predator species. Oecologia 124, 396–401. https://doi.org/10.1007/s004420000420

Bekker, E.I., Karabanov, D.P., Galimov, Y.R., Haag, C.R., Neretina, T. V, Kotov, A.A., 2018. Phylogeography of *Daphnia magna* Straus (Crustacea : Cladocera) in Northern Eurasia : Evidence for a deep longitudinal split between mitochondrial lineages. PLoS One 13, e0194045.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., Stadler, P.F., 2013. MITOS: Improved de novo metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69, 313–319. https://doi.org/10.1016/j.ympev.2012.08.023

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A Software Platform for Bayesian

Evolutionary Analysis. PLoS Comput. Biol. 10, 1–6. https://doi.org/10.1371/journal.pcbi.1003537
Bouckaert, R.R., 2010. DensiTree: Making sense of sets of phylogenetic trees. Bioinformatics 26, 1372–1373. https://doi.org/10.1093/bioinformatics/btq110
Bromham, L., 2009. Why do species vary in their rate of molecular evolution? Biol. Lett. 5, 401–404. https://doi.org/10.1098/rsbl.2009.0136
Bromham, L., Duchêne, S., Hua, X., Ritchie, A.M., Duchêne, D.A., Ho, S.Y.W., 2018. Bayesian molecular dating: opening up the black box. Biol. Rev. 93, 1165–1191. https://doi.org/10.1111/brv.12390
Castresana, J., 2000. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. Mol. Biol. Evol. 17, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334
Chevreux, B., Wetter, T., Suhai, S., 1999. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. Comput. Sci. Biol. Proc. Ger. Conf. Bioinforma. 45–56. https://doi.org/10.1.1.23/7465
Colbourne, J.K., Hebert, P.D.N., 1996. The Systematics of North American *Daphnia* (Crustacea: Anomopoda): A Molecular Phylogenetic Approach. Philos. Trans. R. Soc. B Biol. Sci. 351, 349–360. https://doi.org/10.1098/rstb.1996.0028
Colbourne, J.K., Pfrender, M.E., Gilbert, D., Thomas, W.K., Tucker, A., Oakley, T.H., Tokishita, S., Aerts, A., Arnold, G.J., Basu, M.K., Bauer, D.J., Caceres, C.E., Carmel, L., Casola, C., Choi, J.-H., Detter, J.C., Dong, Q., Dusheyko, S., Eads, B.D., Froehlich, T., Geiler-Samerotte, K.A., Gerlach, D., Hatcher, P., Jogdeo, S., Krijgsveld, J., Kriventseva, E. V, Kueltz, D., Laforsch, C., Lindquist, E., Lopez, J., Manak, J.R., Muller, J., Pangilinan, J., Patwardhan, R.P., Pitluck, S., Pritham, E.J., Rechtsteiner, A., Rho, M., Rogozin, I.B., Sakarya, O., Salamov, A., Schaack, S., Shapiro, H., Shiga, Y., Skalitzky, C., Smith, Z., Souvorov, A., Sung, W., Tang, Z., Tsuchiya, D., Tu, H., Vos, H., Wang, M., Wolf, Y.I., Yamagata, H., Yamada, T., Ye, Y., Shaw, J.R., Andrews, J., Crease, T.J., Tang, H., Lucas, S.M., Robertson, H.M., Bork, P., Koonin, E. V, Zdobnov, E.M., Grigoriev, I. V, Lynch, M., Boore, J.L., 2011. The Ecoresponsive Genome of *Daphnia pulex*. Science. 331, 555–561. https://doi.org/10.1126/science.1197761
Cracraft, J., Houde, P., Ho, S.Y.W., Mindell, D.P., Fjeldså, J., Lindow, B., Edwards, S. V., Rahbek, C., Mirarab, S., Warnow, T., Gilbert, M.T.P., Zhang, G., Braun, E.L., Jarvis, E.D., 2015. Response to Comment on "Whole-genome analyses resolve early branches in the tree of life of modern birds." Science. 349, 1460–b. https://doi.org/10.1126/science.aab1578
Crease, T.J., Omilian, A.R., Costanzo, K.S., Taylor, D.J., 2012. Transcontinental Phylogeography of the *Daphnia pulex* Species Complex. PLoS One 7, e46620. https://doi.org/10.1371/journal.pone.0046620
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772. https://doi.org/10.1038/nmeth.2109
Davidson, R., Vachaspati, P., Mirarab, S., Warnow, T., 2015. Phylogenomic species tree estimation in the presence of incomplete lineage sorting and horizontal

gene transfer. BMC Genomics 16, 1–12. https://doi.org/10.1186/1471-2164-16-S10-S1
DeSalle, R., Freedman, T., Prager, E.M., Wilson, A., 1987. Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. J. Mol. Evol. 26, 157–64.
deWaard, J.R., Sacherova, V., Cristescu, M.E.A., Remigio, E.A., Crease, T.J., Hebert, P.D.N., 2006. Probing the relationships of the branchiopod crustaceans. Mol. Phylogenet. Evol. 39, 491–502. https://doi.org/10.1016/j.ympev.2005.11.003
Dumont, H.J., Pensaert, J., 1983. A revision of the Scapholeberinae (Crustacea: Cladocera). Hydrobiologia 100, 3–45. https://doi.org/10.1007/BF00027420
Dumont, H.J., Silva-Briano, M., 1998. A reclassification of the anomopod families Macrothricidae and Chydoridae, with the creation of a new suborder, the Radopoda (Crustacea: Branchiopoda). Hydrobiologia 384, 119–149. https://doi.org/10.1023/A:1003259630312
Ebert, D., 2005. Ecology, Epidemiology and Evolution of Parasitism in *Daphnia* [Internet].Bethesda (MD): National Library of Medicine (US), National Center for BiotechnologyInformation. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Books
Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340
Edwards, S. V., 2009. Natural selection and phylogenetic analysis. Proc. Natl. Acad. Sci. 106, 8799–8800. https://doi.org/10.1073/pnas.0904103106
Engel, K., Tollrian, R., 2009. Inducible defences as key adaptations for the successful invasion of *Daphnia lumholtzi* in North America? Proc. R. Soc. B Biol. Sci. 276, 1865–1873. https://doi.org/10.1098/rspb.2008.1861
Eyun, S. Il, 2017. Phylogenomic analysis of Copepoda (Arthropoda, Crustacea) reveals unexpected similarities with earlier proposed morphological phylogenies. BMC Evol. Biol. 17, 1–12. https://doi.org/10.1186/s12862-017-0883-5
Fields, P.D., Obbard, D.J., Mctaggart, S.J., Galimov, Y., Tom, J., Ebert, D., 2018. Mitogenome phylogeographic analysis of a planktonic crustacean. Mol. Phylogenet. Evol. https://doi.org/10.1016/j.ympev.2018.06.028
Fields, P.D., Reisser, C., Dukic, M., Haag, C.R., Ebert, D., 2015. Genes mirror geography in *Daphnia magna*. Mol. Ecol. 24, 4521–4536. https://doi.org/10.1111/mec.13324
Fryer, G., 1991. Functional morphology and the adaptive radiation of the Daphniidae (Branchiopoda: Anomopoda). Philos. Trans. R. Soc. London. Ser. B Biol. Sci. 331, 1-99.
Galtier, N., Nabholz, B., Glemin, S., Hurst, G.D.D., 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 18, 4541–4550. https://doi.org/10.1111/j.1365-294X.2009.04380.x
Glagolev, S.M., Alonso, M., 1990. *Daphnia* (*Ctenodaphnia*) *hispanica* sp. nov., a new daphnid (Cladocera) from Spain. Hydrobiologia 194, 149–162. https://doi.org/10.1007/BF00028416

Haag, C.R., McTaggart, S.J., Didier, A., Little, T.J., Charlesworth, D., 2009. Nucleotide polymorphism and within-gene recombination in *Daphnia magna* and *D. pulex*, two cyclical parthenogens. Genetics 182, 313–323. https://doi.org/10.1534/genetics.109.101147

Hahn, C., Bachmann, L., Chevreux, B., 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads - A baiting and iterative mapping approach. Nucleic Acids Res. 41, e129. https://doi.org/10.1093/nar/gkt371

Hebert, P.D.N., 1978. The adaptive significance of cyclomorphosis in *Daphnia*: more possibilities. Freshw. Biol. 8, 313–320. https://doi.org/10.1111/j.1365-2427.1978.tb01452.x

Herzog, Q., Rabus, M., Ribeiro, B.W., Laforsch, C., 2016. Inducible defenses with a "twist": *Daphnia barbata* abandons bilateral symmetry in response to an ancient predator. PLoS One 11, e0148556. https://doi.org/10.1371/journal.pone.0148556

Jackman, S.D., Yeo, S., Coombe, L., Warren, R.L., 2017. ABySS 2.0: Resource-Efficient Assembly of Large Genomes using a Bloom Filter. Genome Res. 27, 768–777. https://doi.org/10.1101/gr.214346.116.Freely

Keith, N., Tucker, A.E., Jackson, C.E., Sung, W., Lledó, J.I.L., Schrider, D.R., Schaack, S., Dudycha, J.L., Ackerman, M., Younge, A.J., Shaw, J.R., Lynch, M., 2016. High mutational rates of large-scale duplication and deletion in *Daphnia pulex*. Genome Res. 26, 60–69. https://doi.org/10.1101/gr.191338.115

Kotov, A., 2013. Morphology and Phylogeny of the Anomopoda (Crustacea: Cladocera). KMW, Moskow.

Kotov, A.A., 2015. A critical review of the current taxonomy of the genus *Daphnia* O. F. Müller, 1785 (Anomopoda, Cladocera). Zootaxa 3911, 184–200. https://doi.org/10.11646/zootaxa.3911.2.2

Kotov, A.A., 2007. Jurassic Cladocera (Crustacea, Branchiopoda) with a description of an extinct Mesozoic order. J. Nat. Hist. 41, 13–37. https://doi.org/10.1080/00222930601164445

Kotov, A.A., Taylor, D.J., 2011. Mesozoic fossils (>145 Mya) suggest the antiquity of the subgenera of *Daphnia* and their coevolution with chaoborid predators. BMC Evol. Biol. 11, 129. https://doi.org/10.1186/1471-2148-11-129

Kotov, A.A., Wappler, T., 2015. Findings of *Daphnia* (*Ctenodaphnia*) Dybowski et Grochowski (Branchiopoda: Cladocera) in Cenozoic volcanogenic lakes in Germany, with discussion of their indicator value. Palaeontol. Electron. 1–9. https://doi.org/10.26879/542

Kumar, S., Filipski, A., 2007. Multiple sequence alignment : In pursuit of homologous DNA positions Multiple sequence alignment : In pursuit of homologous DNA positions. Genome Res. 17, 127–135. https://doi.org/10.1101/gr.5232407

Lampert, W., 2011. Daphnia: development of a model organism in ecology and evolution. Excellence in Ecology, Volume 21. International Ecology Institute Publishers, Oldendorf/Luhe, ISSN 0932-2205. 250 pp.

Lutz, H., 1991. Autochthone aquatische Arthropoda aus dem Mittel-Eozän der Fundstätte Messel (Insecta: Heteroptera; Coleoptera; cf. Diptera-Nematocera;

1010 Crustacea: Cladocera). Cour. Forschungsinstitut Senckenb. 139, 119–125.
1011 Maddison, W.P., 1997. Gene trees in species trees. Syst Biol 46, 523–536.
1012 https://doi.org/10.1093/sysbio/46.3.523
1013 Marková, S., Dufresne, F., Rees, D.J., Černý, M., Kotlík, P., 2007. Cryptic
1014 intercontinental colonization in water fleas *Daphnia pulicaria* inferred from
1015 phylogenetic analysis of mitochondrial DNA variation. Mol. Phylogenet. Evol. 44,
1016 42–52. https://doi.org/10.1016/j.ympev.2006.12.025
1017 Meiklejohn, K.A., Danielson, M.J., Faircloth, B.C., Glenn, T.C., Braun, E.L., Kimball, R.T.,
1018 2014. Incongruence among different mitochondrial regions: A case study using
1019 complete mitogenomes. Mol. Phylogenet. Evol. 78, 314–323.
1020 https://doi.org/10.1016/j.ympev.2014.06.003
1021 Nosenko, T., Schreiber, F., Adamska, M., Adamski, M., Eitel, M., Hammel, J.,
1022 Maldonado, M., Müller, W.E.G., Nickel, M., Schierwater, B., Vacelet, J., Wiens, M.,
1023 Wörheide, G., 2013. Deep metazoan phylogeny: When different genes tell
1024 different stories. Mol. Phylogenet. Evol. 67, 223–233.
1025 https://doi.org/10.1016/j.ympev.2013.01.010
1026 Orlova-Bienkowskaja, M.Y., 2001. Cladocera: Anomopoda: Daphniidae, genus
1027 *Simocephalus*. Guides to the identification of the microinvertebrates of the
1028 continental waters of the World. Vol. 17. Backhuys, Leyden, 130 pp.
1029 Orsini, L., Mergeay, J., Vanoverbeke, J., De Meester, L., 2013. The role of selection in
1030 driving landscape genomic structure of the waterflea *Daphnia magna*. Mol. Ecol.
1031 22, 583–601. https://doi.org/10.1111/mec.12117
1032 Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of phylogenetics and
1033 evolution in R language. Bioinformatics 20, 289–290.
1034 https://doi.org/10.1093/bioinformatics/btg412
1035 Petrusek, A., Hobæk, A., Nilssen, J.P., Skage, M., Černý, M., Brede, N., Schwenk, K.,
1036 2008. A taxonomic reappraisal of the European *Daphnia longispina* complex
1037 (Crustacea, Cladocera, Anomopoda). Zool. Scr. 37, 507–519.
1038 https://doi.org/10.1111/j.1463-6409.2008.00336.x
1039 Petrusek, A., Tollrian, R., Schwenk, K., Haas, A., Laforsch, C., 2009. A “crown of thorns”
1040 is an inducible defense that protects *Daphnia* against an ancient predator. Proc.
1041 Natl. Acad. Sci. 106, 2248–2252. https://doi.org/10.1073/pnas.0808075106
1042 Popova, E.Y., Kotov, A.A., 2013. Latitudinal patterns in the diversity of two
1043 subgenera of the genus *Daphnia* O.F. Müller (Crustacea: Cladocera: Daphniidae).
1044 Zootaxa 3736, 159–174. https://doi.org/10.11646/zootaxa.3736.2.4
1045 Popova, E. V., Petrusek, A., Kořínek, V., Mergeay, J., Bekker, E.I., Karabanov, D.P.,
1046 Galimov, Y.R., Neretina, T. V., Taylor, D.J., Kotov, A.A., 2016. Revision of the old
1047 world *Daphnia* (*Ctenodaphnia*) *similis* group (Cladocera: Daphniidae). Zootaxa
1048 4161, 1–40. https://doi.org/10.11646/zootaxa.4161.1.1
1049 Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1.6. Available
1050 from: http://tree.bio.ed.ac.uk/software/tracer/.
1051 Raymond, P.E., 1946. The genera of fossil Conchostracan order of bivalved Crustacea.
1052 Bull. Museum Comp. Zool. 96, 215–307.
1053 Richter, S., Olesen, J., Wheeler, W.C., 2007. Phylogeny of Branchiopoda (Crustacea)
1054 based on a combined analysis of morphological data and six molecular loci.

Cladistics 23, 301–336.
Rieux, A., Balloux, F., 2016. Inferences from tip-calibrated phylogenies: A review and a practical guide. Mol. Ecol. 25, 1911–1924. https://doi.org/10.1111/mec.13586
Rodgher, S., Espíndola, E.L.G., Lombardi, A.T., 2010. Suitability of *Daphnia similis* as an alternative organism in ecotoxicological tests: Implications for metal toxicity. Ecotoxicology 19, 1027–1033. https://doi.org/10.1007/s10646-010-0484-1
Rokas, A., Ladoukakis, E., Zouros, E., 2003. Animal mitochondrial DNA recombination revisited. Trends Ecol. Evol. 18, 411–417. https://doi.org/10.1016/S0169-5347(03)00125-3
Rota-Stabelli, O., Telford, M.J., 2008. A multi criterion approach for the selection of optimal outgroups in phylogeny: Recovering some support for Mandibulata over Myriochelata using mitogenomics. Mol. Phylogenet. Evol. 48, 103–111. https://doi.org/10.1016/j.ympev.2008.03.033
Rubinoff, D., Holland, B.S., 2005. Between Two Extremes: Mitochondrial DNA is neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference. Syst. Biol. 54, 952–961. https://doi.org/10.1080/10635150500234674
Salichos, L., Rokas, A., 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature 497, 327–331. https://doi.org/10.1038/nature12130
Schwenk, K., Posada, D., Hebert, P.D.N., 2000. Molecular systematics of European Hyalodaphnia: the role of contemporary hybridization in ancient species. Proc. R. Soc. B Biol. Sci. 267, 1833–1842. https://doi.org/10.1098/rspb.2000.1218
Schwenk, K., Spaak, P., 1997. Ecology and genetics of interspecific hybridization in *Daphnia*, In B. Streit, T. Staedler, & C. M. Lively (Eds.), *Ecology and Evolution of Freshwater Animals* (pp. 199-229). Basel: Birkhäuser Verlag.
Shaw, J.R., Pfrender, M.E., Eads, B.D., Klaper, R., Callaghan, A., Sibly, R.M., Colson, I., Jansen, B., Gilbert, D., Colbourne, J.K., 2008. *Daphnia* as an emerging model for toxicological genomics. Adv. Exp. Biol. 2, 5–7. https://doi.org/10.1016/S1872-2423(08)00005-7
Simmons, M.P., 2017. Relative benefits of amino-acid, codon, degeneracy, DNA, and purine-pyrimidine character coding for phylogenetic analyses of exons. J. Syst. Evol. 55, 85–109. https://doi.org/10.1111/jse.12233
Simmons, M.P., Carr, T.G., O'Neill, K., 2004. Relative character-state space, amount of potential phylogenetic information, and heterogeneity of nucleotide and amino acid characters. Mol. Phylogenet. Evol. 32, 913–926. https://doi.org/10.1016/j.ympev.2004.04.011
Simmons, M.P., Ochoterena, H., Freudenstein, J. V., 2002. Amino acid vs. nucleotide characters: Challenging preconceived notions. Mol. Phylogenet. Evol. 24, 78–90. https://doi.org/10.1016/S1055-7903(02)00202-6
Soria-Carrasco, V., Talavera, G., Igea, J., Castresana, J., 2007. The K tree score: Quantification of differences in the relative branch length and topology of phylogenetic trees. Bioinformatics 23, 2954–2956. https://doi.org/10.1093/bioinformatics/btm466
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-

1100 analysis of large phylogenies. Bioinformatics 30, 1312–3.
1101 https://doi.org/10.1093/bioinformatics/btu033
1102 Stenderup, J.T., Olesen, J., Glenner, H., 2006. Molecular phylogeny of the
1103 Branchiopoda (Crustacea)-Multiple approaches suggest a "diplostracan"
1104 ancestry of the Notostraca. Mol. Phylogenet. Evol. 41, 182–194.
1105 https://doi.org/10.1016/j.ympev.2006.06.006
1106 Stollewerk, A., 2010. The water flea *Daphnia*--a "new" model system for ecology and
1107 evolution? J. Biol. 9, 21. https://doi.org/10.1186/jbiol212
1108 Sun, X.Y., Xia, X., Yang, Q., 2016. Dating the origin of the major lineages of
1109 Branchiopoda. Palaeoworld 25, 303–317.
1110 https://doi.org/10.1016/j.palwor.2015.02.003
1111 Swain, T.D., Taylor, D.J., 2003. Structural rRNA characters support monophyly of
1112 raptorial limbs and paraphyly of limb specialization in water fleas. Proc. R. Soc.
1113 B Biol. Sci. 270, 887–896. https://doi.org/10.1098/rspb.2002.2297
1114 Taylor, D.J., Hebert, P.D.N., Colbourne, J.K., 1996. Phylogenetics and evolution of the
1115 *Daphnia longispina* group (Crustaceae) based on 12S r DNA sequence and
1116 allozyme variation. Mol. Phylogenet. Evol. 5, 495–510.
1117 Thielsch, A., Knell, A., Mohammadyari, A., Petrusek, A., Schwenk, K., 2017. Divergent
1118 clades or cryptic species? Mito-nuclear discordance in a *Daphnia* species
1119 complex. BMC Evol. Biol. 17, 1–9. https://doi.org/10.1186/s12862-017-1070-4
1120 Tucker, A.E., Ackerman, M.S., Eads, B.D., Xu, S., Lynch, M., 2013. Population-genomic
1121 insights into the evolutionary origin and fate of obligately asexual *Daphnia*
1122 *pulex*. Proc. Natl. Acad. Sci. U. S. A. 110, 15740–5.
1123 https://doi.org/10.1073/pnas.1313388110
1124 Vaidya, G., Lohman, D.J., Meier, R., 2011. SequenceMatrix: Concatenation software
1125 for the fast assembly of multi-gene datasets with character set and codon
1126 information. Cladistics 27, 171–180. https://doi.org/10.1111/j.1096-
1127 0031.2010.00329.x
1128 Van Damme, K., Bekker, E.I., Kotov, A.A., 2013. Endemism in the Cladocera
1129 (Crustacea: Branchiopoda) of Southern Africa. J. Limnol. 72, 440–463.
1130 https://doi.org/10.4081/jlimnol.2013.e36
1131 Van Damme, K., Kotov, A.A., 2016. The fossil record of the Cladocera (Crustacea:
1132 Branchiopoda): Evidence and hypotheses. Earth-Science Rev. 163, 162–189.
1133 https://doi.org/10.1016/j.earscirev.2016.10.009
1134 Van Damme, K., Shiel, R.J., Dumont, H.J., 2007. *Notothrix halsei* gen. n., sp. n.,
1135 representative of a new family of freshwater cladocerans (Branchiopoda,
1136 Anomopoda) from SW Australia, with a discussion of ancestral traits and a
1137 preliminary molecular phylogeny of the order. Zool. Scr. 36, 465–487.
1138 https://doi.org/10.1111/j.1463-6409.2007.00292.x
1139 Warnock, R.C.M., Yang, Z., Donoghue, P.C.J., 2012. Exploring uncertainty in the
1140 calibration of the molecular clock. Biol. Lett. 8, 156–159.
1141 https://doi.org/10.1098/rsbl.2011.0710
1142 Waterhouse, R.M., Seppey, M., Simão, F.A., Manni, M., Ioannidis, P., Klioutchnikov, G.,
1143 Kriventseva, E. V, Zdobnov, E.M., 2017. BUSCO Applications from Quality
1144 Assessments to Gene Prediction and Phylogenomics. Mol. Biol. Evol. 1–6.

https://doi.org/10.1093/molbev/msx319
Wiens, J.J., Kuczynski, C.A., Stephens, P.R., 2010. Discordant mitochondrial and nuclear gene phylogenies in emydid turtles: Implications for speciation and conservation. Biol. J. Linn. Soc. 99, 445–461. https://doi.org/10.1111/j.1095-8312.2009.01342.x
Xu, S., Schaack, S., Seyfert, A., Choi, E., Lynch, M., Cristescu, M.E., 2012. High mutation rates in the mitochondrial genomes of *Daphnia pulex*. Mol. Biol. Evol. 29, 763–769. https://doi.org/10.1093/molbev/msr243
Xu, S., Spitze, K., Ackerman, M.S., Ye, Z., Bright, L., Keith, N., Jackson, C.E., Shaw, J.R., Lynch, M., 2015. Hybridization and the origin of contagious asexuality in *Daphnia pulex*. Mol. Biol. Evol. 32, 3215–3225. https://doi.org/10.1093/molbev/msv190
Ye, Z., Xu, S., Spitze, K., Asselman, J., Jiang, X., Ackerman, M.S., Lopez, J., Harker, B., Raborn, R.T., Thomas, W.K., Ramsdell, J., Pfrender, M.E., Lynch, M., 2017. A New Reference Genome Assembly for the Microcrustacean *Daphnia pulex* 7, 1405–1416. https://doi.org/10.1534/g3.116.038638/-/DC1.1
Zeng, L., Zhang, Q., Sun, R., Kong, H., Zhang, N., Ma, H., 2014. Resolution of deep angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times. Nat. Commun. 5, 1–12. https://doi.org/10.1038/ncomms5956
Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19, 15–30. https://doi.org/10.1186/s12859-018-2129-y
Zhang, N., Zeng, L., Shan, H., Ma, H., 2012. Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms. New Phytol. 195, 923–937. https://doi.org/10.1111/j.1469-8137.2012.04212.x
Zhang, P., Liang, D., Mao, R.L., Hillis, D.M., Wake, D.B., Cannatella, D.C., 2013. Efficient sequencing of anuran mtDNAs and a mitogenomic exploration of the phylogeny and evolution of frogs. Mol. Biol. Evol. 30, 1899–1915. https://doi.org/10.1093/molbev/mst091
Zhu, T., Reis, M. Dos, Yang, Z., 2015. Characterization of the uncertainty of divergence time estimation under relaxed molecular clock models using multiple loci. Syst. Biol. 64, 267–280. https://doi.org/10.1093/sysbio/syu109
Zimin, A. V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S.L., Yorke, J.A., 2013. The MaSuRCA genome assembler. Bioinformatics 29, 2669–2677. https://doi.org/10.1093/bioinformatics/btt476

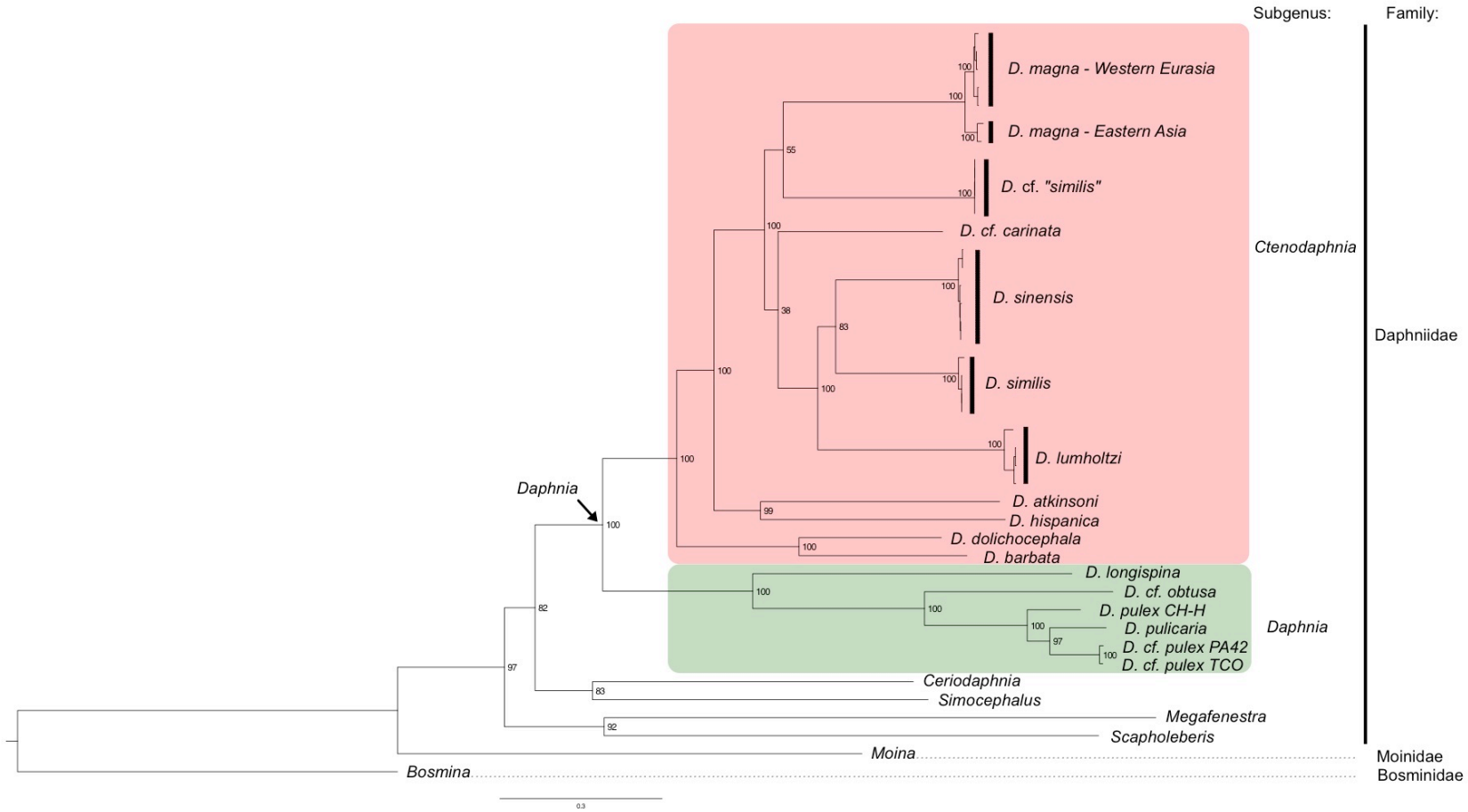
Figure Captions

Figure 1. MtDNA Maximum Likelihood tree. This tree is based on the 13 PCGs and rRNA genes. Bootstrap values are reported beside internal nodes.

Figure 2. Genome-wide nuclear Maximum Likelihood phylogeny. This tree is based on 636 nuclear genes. Bootstrap values from the RAxML analysis are reported above the nodes and local posterior probability values from the ASTRAL analysis are reported below the nodes.

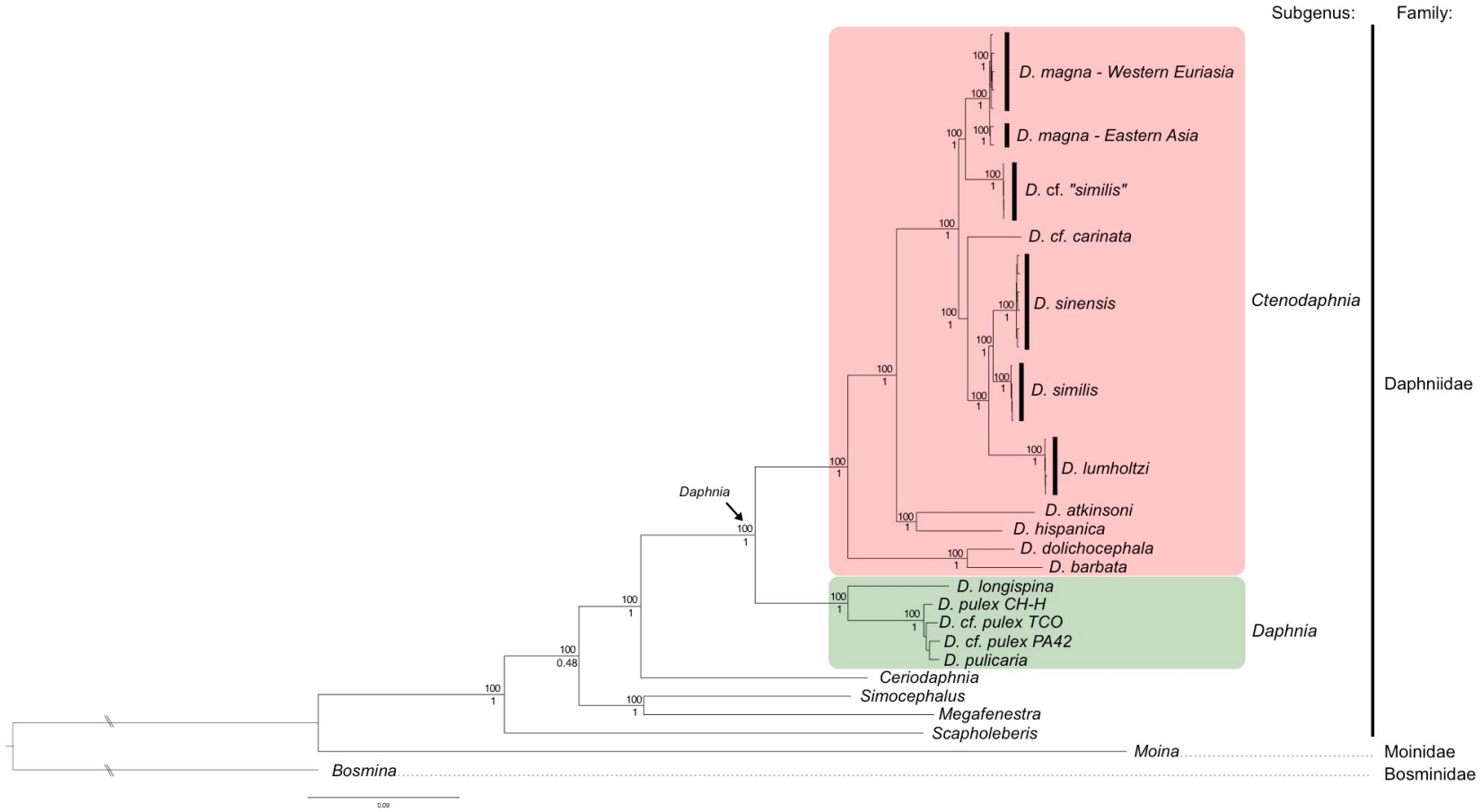
Figure 3. DensiTree of the nuclear genes. In this figure only one representative per species was included in order to facilitate the visualization. *D*. cf. *carinata* is not included here due to its relevant number of missing genes. The topology of each nuclear tree is drawn in green and the blue consensus tree has been obtained using the 'root canal' function of DensiTree.

1200 **Figure 1**



1201
1202

1203 **Figure 2**



1204
1205

Figure 3

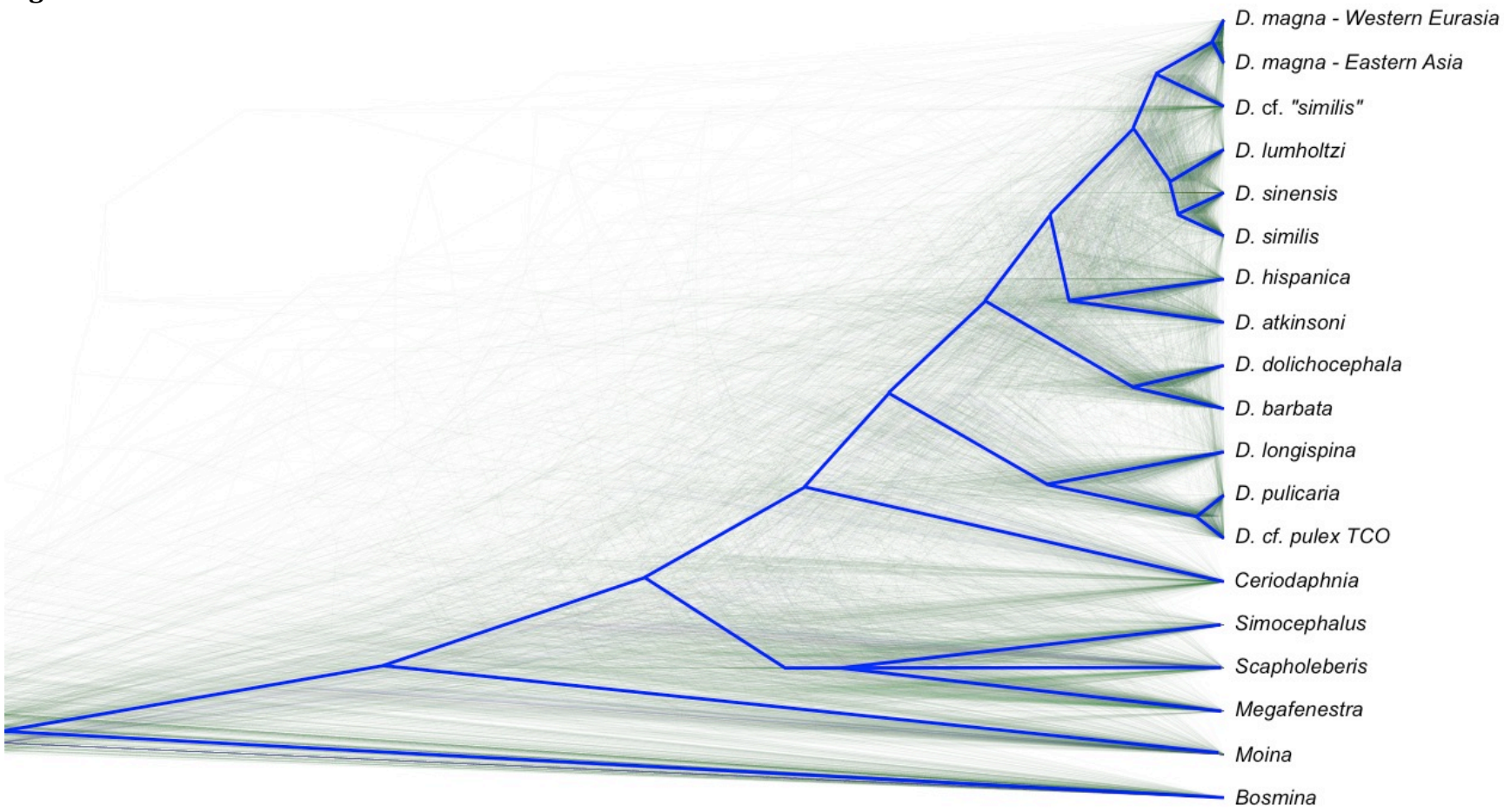

D. magna - Western Eurasia
D. magna - Eastern Asia
D. cf. "similis"
D. lumholtzi
D. sinensis
D. similis
D. hispanica
D. atkinsoni
D. dolichocephala
D. barbata
D. longispina
D. pulicaria
D. cf. pulex TCO
Ceriodaphnia
Simocephalus
Scapholeberis
Megafenestra
Moina
Bosmina

Table 1. List and details of cladoceran clones included in this study

Family	Genus	Subgenus	Species (Species group)	Country	Clone name	Latitude	Longitude
Bosminidae	*Bosmina*	Eubosmina	*Bosmina* cf. *longispina maritima* (*B. coregoni* group)	Finland	Bosmina_cf.longispina_maritima_FI-BAL1-1	59.845123	23.249092
Moinidae	*Moina*		*Moina brachiata* (*M. brachiata* group)	Germany	Moina_sp_DE-FRO-2-1	48.214947	11.613311
Daphniidae	*Megafenestra*		*Megafenestra aurita*	Switzerland	Megafenestra_aurita_CH-H-2	47.557769	8.862608
Daphniidae	*Scapholeberis*		*Scapholeberis mucronata* (*S. mucronata* group)	Belgium	Scapholeberis_mucronata_BE-ASS	51.04907	3.666028
Daphniidae	*Simocephalus*		*Simocephalus* cf. *serrulatus* (*S. serrulatus* group)	Oman	Simocephalus_cf_serrulatus_OM-SAIQ-clone2	23.079561	57.670617
Daphniidae	*Ceriodaphnia*		*Ceriodaphnia* cf. *reticulata* (*C. reticulata* group)	Oman	Ceriodaphnia_sp_OM-SAIQ-clone2	23.079561	57.670617
Daphniidae	*Daphnia*	Ctenodaphnia	*D.* cf. *carinata* (*D. carinata* group)	Australia	D.cf."carinata"_group_AU-BEG-1	-38.268732	144.536689
Daphniidae	*Daphnia*	Ctenodaphnia	*D.* cf. "*similis*" (*D. exilis* group)	Canada	D.cf."similis"_CA-CBC-31	49.567824	-115.725945
Daphniidae	*Daphnia*	Ctenodaphnia	*D.* cf. "*similis*" (*D. exilis* group)	Canada	D.cf."similis"_CA-CBC-34	49.567824	-115.725945
Daphniidae	*Daphnia*	Ctenodaphnia	*D.* cf. "*similis*" (*D. exilis* group)	Canada	D.cf."similis"_CA-CBC-37	49.567824	-115.725945
Daphniidae	*Daphnia*	Ctenodaphnia	*D.* cf. "*similis*" (*D. exilis* group)	Canada	D.cf."similis"_CA-CBC-38	49.567824	-115.725945
Daphniidae	*Daphnia*	Ctenodaphnia	*D. atkinsoni* (*D. atkinsoni* group)	Israel	D.atkinsoni_IL-KID-3b-11	31.267059	35.233988
Daphniidae	*Daphnia*	Ctenodaphnia	*D. barbata*	Zimbabwe	D.barbata_ZW-BAR-1	-17.897591	30.791585
Daphniidae	*Daphnia*	Ctenodaphnia	*D. dolichocephala*	South Africa	D.dolichocephala_ZA-DOLI	NA	NA
Daphniidae	*Daphnia*	Ctenodaphnia	*D. hispanica*	Portugal	D.hispanica_PT-GA-1	37.05066400	-7.97930600
Daphniidae	*Daphnia*	Ctenodaphnia	*D. lumholtzi* (*D. similis* group)	India	D.lumholtzi_IN-PA-1	18.5204	73.8567
Daphniidae	*Daphnia*	Ctenodaphnia	*D. lumholtzi* (*D. similis* group)	USA	D.lumholtzi_US-AR	33.543495	-111.435493
Daphniidae	*Daphnia*	Ctenodaphnia	*D. lumholtzi* (*D. similis* group)	USA	D.lumholtzi_US-MO	37.597924	-93.711006
Daphniidae	*Daphnia*	Ctenodaphnia	*D. lumholtzi* (*D. similis* group)	Zimbabwe	D.lumholtzi_ZW-LUM	NA	NA
Daphniidae	*Daphnia*	Ctenodaphnia	*D. magna*	USA	CA-CH-1	58.770982	-93.850837
Daphniidae	*Daphnia*	Ctenodaphnia	*D. magna*	Switzerland	CH-H-1	47.557769	8.862608
Daphniidae	*Daphnia*	Ctenodaphnia	*D. magna*	Finland	FI-XINB3	59.833183	23.260387

Daphniidae	*Daphnia*	Ctenodaphnia	*D. magna*	France	FR-SA-1	43.48012	4.647302
Daphniidae	*Daphnia*	Ctenodaphnia	*D. magna*	UK	GB-EK1-32	55.702406	-2.340828
Daphniidae	*Daphnia*	Ctenodaphnia	*D. magna*	Mongolia	MN-DM1-1	45.032708	100.660481
Daphniidae	*Daphnia*	Ctenodaphnia	*D. magna*	Central Asia	RU-SAM5	52.92296	50.31727
Daphniidae	*Daphnia*	Ctenodaphnia	*D. similis* (*D. similis* group)	Israel	D.similis_IL-B-3	32.01325	34.963211
Daphniidae	*Daphnia*	Ctenodaphnia	*D. similis* (*D. similis* group)	Israel	D.similis_IL-KYN-4	32.130706	34.811168
Daphniidae	*Daphnia*	Ctenodaphnia	*D. similis* (*D. similis* group)	Israel	D.similis_IL-NS-13	31.724185	34.626269
Daphniidae	*Daphnia*	Ctenodaphnia	*D. similis* (*D. similis* group)	Israel	D.similis_IL-SIM-A20-inb3-14	32.781095	35.407369
Daphniidae	*Daphnia*	Ctenodaphnia	*D. sinensis* (*D. similis* group)	Russia	D.sinensis_RU-BU1-3	51.231333	108.3035
Daphniidae	*Daphnia*	Ctenodaphnia	*D. sinensis* (*D. similis* group)	Russia	D.sinensis_RU-NOV1-01	55.127795	77.037327
Daphniidae	*Daphnia*	Ctenodaphnia	*D. sinensis* (*D. similis* group)	Russia	D.sinensis_RU-SZB3	50.347667	114.873833
Daphniidae	*Daphnia*	Ctenodaphnia	*D. sinensis* (*D. similis* group)	Russia	D.sinensis_RU-TU2-01	55.713667	68.991
Daphniidae	*Daphnia*	Ctenodaphnia	*D. sinensis* (*D. similis* group)	Russia	D.sinensis_RU-TY6-1	50.2555	89.546833
Daphniidae	*Daphnia*	Ctenodaphnia	*D. sinensis* (*D. similis* group)	Russia	D.sinensis_RU-TY6-3	50.2555	89.546833
Daphniidae	*Daphnia*	Daphnia	*D.* cf. *obutsa* (*D. obtusa* group)	USA	D.cf.obtusa	NA	NA
Daphniidae	*Daphnia*	Daphnia	*D. pulex* (*D. pulex* group)	Switzerland	D.pulex_CH-H	47.557769	8.862608
Daphniidae	*Daphnia*	Daphnia	*D.* cf. *pulex* (*D. pulex* group)	USA	D.cf.pulex_PA42	40.2013	−87.3294
Daphniidae	*Daphnia*	Daphnia	*D.* cf. *pulex* (*D. pulex* group)	USA	D.cf.pulex_TCO	43.830013	-124.148152
Daphniidae	*Daphnia*	Daphnia	*D. pulicaria* (*D. pulex* group)	Czech Republic	D.pulicaria_CZ-RIM1-1	48.845027	14.484201
Daphniidae	*Daphnia*	Daphnia	*D. longispina* (*D. longispina* group)	Finland	D.longispina_FI-G-95-1_INB4-1	59.814996	23.248143

1212
1213

Table 2. Divergence times obtained using calibration based on fossil records and substitution rate for mtDNA (all 13 PCGs and rRNA genes) and nuclear genes (four-fold degenerate sites). The median and the 95 % CI are reported. We also reported the different estimates obtained with different priors applied to the most recent common ancestor of all Anomopoda (i.e. Late Jurassic, Early Jurassic and Permian). Four additional priors are used in the fossil calibration analysis (see main text for details).

	node age (Mya [95% HPD])					
dataset	subgenus *Daphnia-Ctenodaphnia*	*D. longispina* – *D. pulex group*	*D. magna* – *D. sinensis*	*D. magna* – *D.* cf *."similis"*	*D. similis* – *D. sinensis*	*D. magna* (Europe) – *D. magna* (Asia)
mtDNA - fossil calibration - Late Jurassic	145.2 [145.0-145.5]	104.7 [101.1-108.3]	74.1 [71.9-76.5]	66.6 [63.8-69.6]	47.8 [45.6-50.1]	6.9 [6.4-7.4]
mtDNA - fossil calibration - Early Jurassic	146.1 [145.1-147.6]	110.6 [106.9-114.2]	81.1 [78.6-83.7]	73.0 [69.9-76.2]	52.5 [49.9-55.0]	7.6 [7.0-8.2]
mtDNA - fossil calibration - Permian	175.4 [167.0-187.1]	133.9 [126.2-143.1]	99.0 [93.4-105.3]	89.4 [84.1-95.8]	64.1 [59.9-68.9]	9.3 [8.5-10.2]
mtDNA - substitution rate	26.9 [26.1-27.8]	20.3 [19.5-21.2]	14.9 [14.4-15.4]	13.5 [12.8-14.1]	9.7 [9.2-10.1]	1.4 [1.3-1.5]
nuclear genes - fossil calibration - Late Jurassic	145.0 [145.0-145.1]	31.5 [30.6-32.3]	16.3 [15.9-16.6]	10.9 [10.6-11.3]	7.6 [7.3-7.8]	1.3 [1.2-1.4]
nuclear genes - fossil calibration - Early Jurassic	145.0 [145.0-145.1]	37.0 [36.1-37.9]	19.0 [18.6-19.4]	12.8 [12.4-13.2]	8.8 [8.5-9.1]	1.5 [1.4-1.6]
nuclear genes - fossil calibration - Permian	145.1 [145.0-145.2]	47.3 [46.1-48.5]	24.2 [23.7-24.8]	16.3 [15.8-16.8]	11.3 [10.9-11.6]	1.9 [1.8-2.0]
nuclear genes - substitution rate	102.2 [100.4-104.8]	41.7 [40.7-42.8]	21.9 [21.5-22.4]	14.5 [14.1-14.9]	10.4 [10.1-10.9]	1.7 [1.6-1.8]

Supplementary materials

A Fossil-Calibrated Phylogenomic Analysis of Daphnia and the Daphniidae

Luca Cornetti[1], Peter D. Fields[1], Kay Van Damme[2,3] and Dieter Ebert[1]

[1] *University of Basel, Department of Environmental Sciences, Zoology, Vesalgasse 1, 4051 Basel, Switzerland*
[2] *Senckenberg Research Institute and Natural History Museum Frankfurt, Senckenberganlage 25, 60325 Frankfurt am Main, Germany*
[3] *Tvärminne Zoological Station, University of Helsinki, J.A. Palménin tie 260, Hanko, Finland*

Corresponding author: L. Cornetti, E-mail: luca.cornetti@unibas.ch

Figure S1. ML tree including our newly obtained sequences (highlighted in green and named according to column "Clone name" in Table 1; when only one clone per species was included, its name was shortened in order to simplify the labels) and all the taxa in Adamowicz et al., (2009) that were sequenced for 12S, 16S and COI. For the latter taxa, the same labels as in Adamowicz et al., (2009) were used. Where applicable, we provided additional labeling for species, species groups, genera and subgenera.

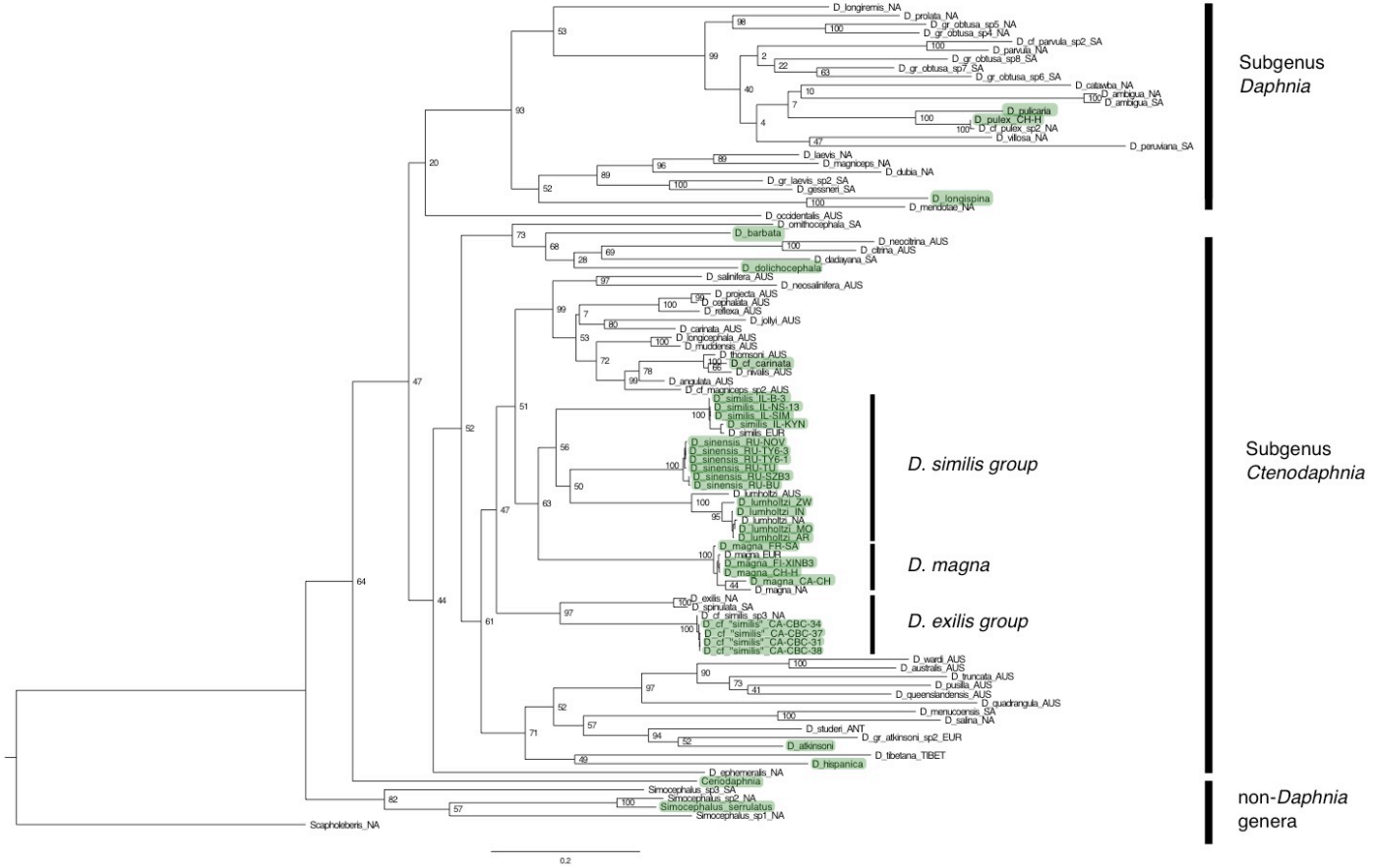


Figure S2. ML tree of the 12S gene including our newly obtained sequences (highlighted in green and named according to column "Clone name" in Table 1; in some cases the labels were shortened in order to simplify them) and the taxa in Popova et al., (2016). For the latter taxa, the same labels as in Popova et al., (2016) were used. Where applicable, we provided additional labeling for species groups.

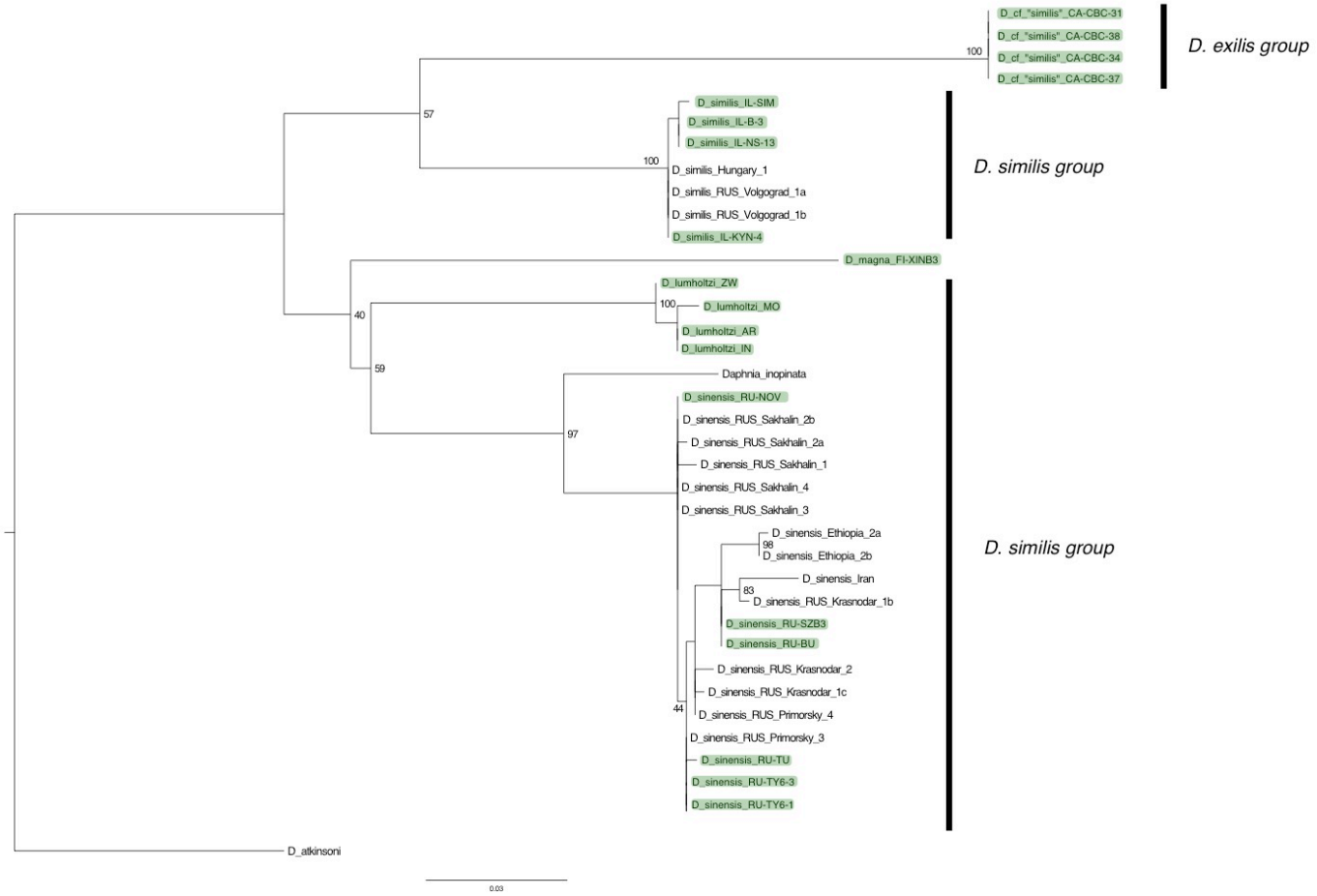


Figure S3. ML tree of the COI gene including our newly obtained sequences (highlighted in green and named according to column "Clone name" in Table 1; in some cases the labels were shortened in order to simplify them) and the taxa in Popova et al., (2016). For the latter taxa, the same labels as in Popova et al., (2016) were used. Where applicable, we provided additional labeling for species identification.

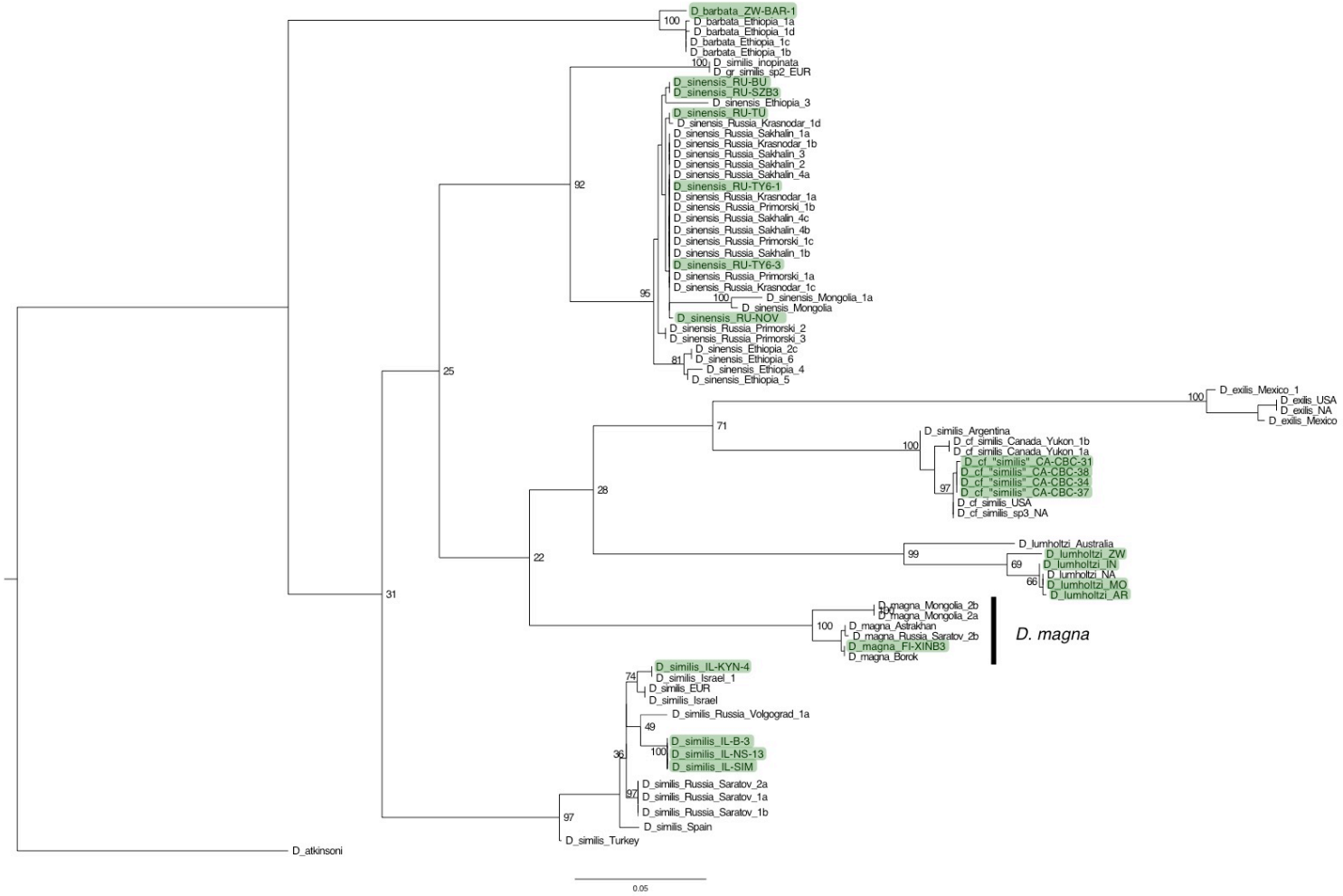


Figure S4. Resume of the BUSCO genome assessment for each of the draft nuclear genome assembly analyzed in this study

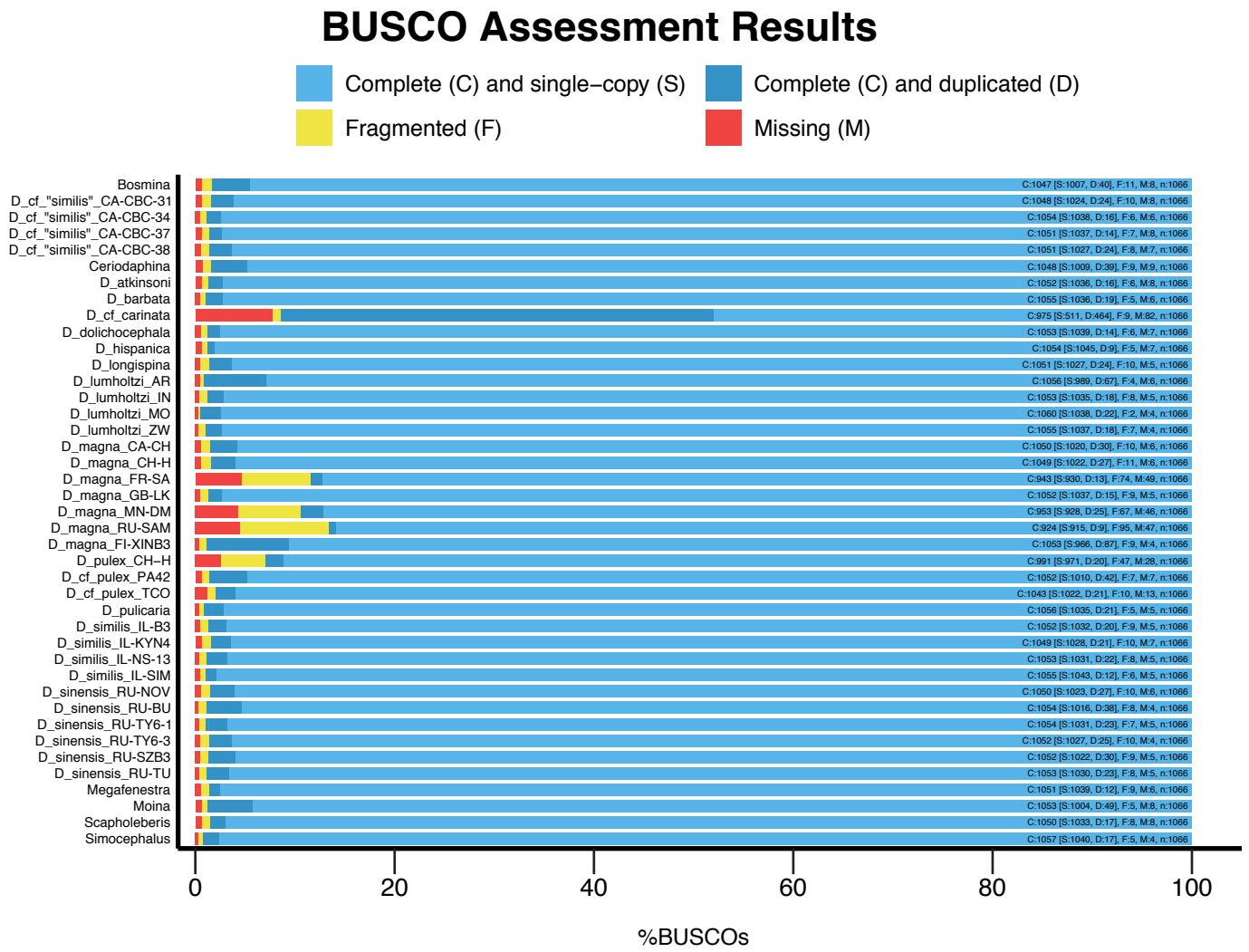


Figure S5. ML amino acid tree of the 636 nuclear genes. Bootstrap values from the RAxML analysis are reported above the nodes and local posterior probability values from the ASTRAL analyses are reported below the nodes. The ASTRAL species trees obtained with nucleotide and amino acid sequences were identical and in complete agreement with the ML topology resulting from the concatenated amino acid dataset. Local posterior probabilities obtained with nucleotide and amino acid sequences are reported on the left and on the right of "/", respectively.

1265 **Table S1.** Assembly statistics for each of the draft nuclear genome analyzed in this study.
1266

Species/Clone name	**n of scaffold**	**N50 (bp)**	**max scaffold length (bp)**	**total assembly length (Mb)**
Bosmina cf. *longispina maritima* (*B. coregoni* group)	28611	37861	699777	117
Moina brachiata (*M. brachiata* group)	14987	55219	736482	95
Megafenestra aurita	21309	104119	645684	112
Scapholeberis mucronata (*S. mucronata* group)	14674	114421	976602	102
Simocephalus cf. *serrulatus* (*S. serrulatus* group)	11681	67033	555441	116
Ceriodaphina cf. *reticulata* (*C. reticulata* group)	34395	42048	432353	158
D. cf. *carinata* (*D. carinata* group)	17888	35689	311621	145
D. cf "similis" CA-CBC-31 (*D. exilis* group)	31704	26721	286429	141
D. cf "similis" CA-CBC-34 (*D. exilis* group)	28224	27013	250664	133
D. cf "similis" CA-CBC-37 (*D. exilis* group)	27942	38342	299536	140
D. cf "similis" CA-CBC-38 (*D. exilis* group)	24715	30758	263925	136
D. atkinsoni (*D. atkinsoni* group)	15412	69719	557213	122
D. barbata	12651	131556	856311	93
D. dolichocephala	8411	133388	697702	86
D. hispanica	27804	33690	293516	119
D. lumholtzi IN-PA (*D. similis* group)	10906	95234	514921	101
D. lumholtzi US-AR (*D. similis* group)	8848	77377	504408	101
D. lumholtzi US-MO (*D. similis* group)	8933	90651	615127	102
D. lumholtzi ZW-LUM (*D. similis* group)	7787	68032	950526	104
D. magna CA-CH	17964	31682	270834	129
D. magna CH-H	24385	31607	334948	135
D. magna FI-XINB3	7664	97873	2758015	198
D. magna FR-SA	110666	1933	282211	118
D. magna GB-LK	17593	37549	285667	126
D. magna MN-DM	114151	2094	448582	123

D. magna RU-SAM	102897	1937	47969	107
D. similis IL-B3 (*D. similis* group)	17988	35365	362957	121
D. similis IL-KNY4 (*D. similis* group)	18626	35174	339006	121
D. similis IL-NIZ1-3 (*D. similis* group)	15467	43679	551099	122
D. similis IL-SIM-A20-Inb3 (*D. similis* group)	8924	54196	401850	114
D. sinensis RU-BU1-3 (*D. similis* group)	14078	58175	378024	111
D. sinensis RU-NOV1-01 (*D. similis* group)	12893	83404	564333	110
D. sinensis RU-SZB3-2 (*D. similis* group)	15095	63345	381309	112
D. sinensis RU-TU2-01 (*D. similis* group)	13596	78738	637228	110
D. sinensis RU-TY6-1 (*D. similis* group)	14908	71172	378283	112
D. sinensis RU-TY6-3 (*D. similis* group)	16245	71751	470531	110
D. pulex CH-H (*D. pulex* group)	93853	2666	77224	121
D. cf. *pulex* PA42 (*D. pulex* group)	1822	482705	1637002	143
D. cf. pulex TCO (*D. pulex* group)	5186	758069	4058679	159
D. pulicaria (*D. pulex* group)	20796	44347	390586	159
D. longispina (*D. longispina* group)	7743	73962	436802	130

Table S2. K-score comparison of single mtDNA trees with the best ML nuclear phylogeny.

mtDNA gene	K-score	Scale_factor
12S	0.40209	0.49948
16S	0.4093	0.48842
nad6	0.43029	0.16749
nad4	0.44089	0.25296
nad5	0.47433	0.25396
nad3	0.4797	0.23168
nad2	0.49619	0.19232
cox2	0.51177	0.17958
nad1	0.53062	0.26107
nad4l	0.54355	0.24242
cytb	0.56105	0.16669
atp6	0.57556	0.28293
cox1	0.60351	0.85781
cox3	0.60361	0.37814
atp8	0.66847	0.2097

Table S3. MtDNA substitution rates for each gene, when the prior on the root was placed during the Late Jurassic.

mtDNA gene	median substitution rate (per nucleotide per year)	95% HPD Interval
12S	2.36E-09	[2.1371E-09, 2.5783E-09]
16S	2.75E-09	[2.5514E-09, 2.9517E-09]
atp6	3.81E-09	[3.3984E-09, 4.1797E-09]
atp8	6.18E-09	[5.0053E-09, 7.4255E-09]
cox1	3.29E-09	[3.0757E-09, 3.5226E-09]
cox2	3.04E-09	[2.7471E-09, 3.3334E-09]
cox3	3.36E-09	[3.0847E-09, 3.6826E-09]
cytb	3.45E-09	[3.1607E-09, 3.7103E-09]
nad1	4.48E-09	[4.1026E-09, 4.8607E-09]
nad2	5.30E-09	[4.9365E-09, 5.7036E-09]
nad3	4.14E-09	[3.6142E-09, 4.687E-09]
nad4	4.97E-09	[4.6486E-09, 5.309E-09]
nad4_l	4.36E-09	[3.7827E-09, 4.9859E-09]
nad5	4.58E-09	[4.2862E-09, 4.8585E-09]
nad6	5.78E-09	[5.2166E-09, 6.3765E-09]

Table S4. MtDNA substitution rates for each gene, when the prior on the root was placed during the Early Jurassic.

mtDNA gene	median substitution rate (per nucleotide per year)	95% HPD Interval
12S	2.14E-09	[1.9449E-09, 2.3487E-09]
16S	2.49E-09	[2.3049E-09, 2.6662E-09]
atp6	3.48E-09	[3.133E-09, 3.8325E-09]
atp8	5.68E-09	[4.569E-09, 6.8037E-09]
cox1	3.00E-09	[2.7977E-09, 3.2306E-09]
cox2	2.76E-09	[2.4792E-09, 3.0282E-09]
cox3	3.05E-09	[2.7789E-09, 3.3219E-09]
cytb	3.14E-09	[2.8887E-09, 3.395E-09]
nad1	4.10E-09	[3.757E-09, 4.4552E-09]
nad2	4.85E-09	[4.4891E-09, 5.1952E-09]
nad3	3.78E-09	[3.3261E-09, 4.3197E-09]
nad4	4.53E-09	[4.2119E-09, 4.8255E-09]
nad4_l	3.98E-09	[3.4513E-09, 4.5235E-09]
nad5	4.17E-09	[3.9123E-09, 4.4284E-09]
nad6	5.28E-09	[4.7482E-09, 5.8042E-09]

Table S5. MtDNA substitution rates for each gene, when the prior on the root was placed during the Permian.

mtDNA gene	median substitution rate (per nucleotide per year)	95% HPD Interval
12S	1.74E-09	[1.5611E-09, 1.9368E-09]
16S	2.03E-09	[1.8591E-09, 2.2142E-09]
atp6	2.84E-09	[2.5299E-09, 3.1799E-09]
atp8	4.63E-09	[3.7722E-09, 5.6958E-09]
cox1	2.45E-09	[2.2373E-09, 2.6665E-09]
cox2	2.25E-09	[2.0155E-09, 2.5125E-09]
cox3	2.49E-09	[2.2410E-09, 2.7598E-09]
cytb	2.56E-09	[2.3273E-09, 2.8148E-09]
nad1	3.35E-09	[3.0131E-09, 3.6892E-09]
nad2	3.96E-09	[3.6097E-09, 4.3268E-09]
nad3	3.09E-09	[2.6615E-09, 3.5333E-09]
nad4	3.71E-09	[3.3981E-09, 4.0291E-09]
nad4_l	3.25E-09	[2.8016E-09, 3.7404E-09]
nad5	3.41E-09	[3.1382E-09, 3.6915E-09]
nad6	4.32E-09	[3.8607E-09, 4.8379E-09]