Title: Mitogenomic phylogeny and fossil-calibrated mutation rates for all F- and M-type mtDNA genes of the largest freshwater mussel family (Bivalvia: Unionidae)

Short running title: Unionidae Mitogenomic phylogeny and mutation rates

Authors names and affiliations:

Alexandra ZIERITZ[1], Elsa FROUFE[2], Ivan BOLOTOV[3,4,5], Duarte V. GONÇALVES[2], David C. ALDRIDGE[6], Arthur E. BOGAN[7], Han Ming GAN[8], André GOMES-DOS-SANTOS[2,9], Ronaldo SOUSA[10], Amilcar TEIXEIRA[11], Simone VARANDAS[12], David ZANATTA[13], and Manuel LOPES-LIMA[2,14]

[1] School of Geography, Sir Clive Granger Building, University of Nottingham, University Park, NG7 2RD Nottingham, UK

[2] CIIMAR/CIMAR - Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Terminal de Cruzeiros do Porto de Leixões, Av. General Norton de Matos s/n, 4450-208 Matosinhos, Portugal.

[3] Federal Center for Integrated Arctic Research, Russian Academy of Sciences, Severnaya Dvina Emb. 23, 163000Arkhangelsk, Russian Federation

[4] Northern Arctic Federal University, Northern Dvina Emb. 17, 163002 Arkhangelsk, Russian Federation

[5] Saint-Petersburg State University, Universitetskaya Emb. 7/9, 199034 Saint Petersburg, Russian Federation

[6] Aquatic Ecology Group, The David Attenborough Building, Department of Zoology, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, United Kingdom

[7] Research Laboratory, North Carolina State Museum of Natural Sciences, 11 West Jones St., Raleigh, NC 27601, USA.

[8] Centre for Integrative Ecology, School of Life and Environmental Sciences, Deakin University, Geelong, 3220, VIC, Australia

[9] Department of Biology, Faculty of Sciences, University of Porto, Rua do Campo Alegre 1021/1055, Porto, Portugal

[10] CBMA - Centre of Molecular and Environmental Biology, Department of Biology, University of Minho, Campos de Gualtar, 4710-057 Braga, Portugal

[11] Centro de Investigação de Montanha (CIMO), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal

[12] CITAB-UTAD – Centre for Research and Technology of Agro-Environment and Biological Sciences, University of Trás-os-Montes and Alto Douro, Apartado 1013, 5001-811 Vila Real, Portugal

[13] Biology Department, Institute for Great Lakes Research, Central Michigan University, Biosciences 2408, Mount Pleasant, MI 48859, USA

[14] CIBIO/InBIO - Research Center in Biodiversity and Genetic Resources, Universidade do Porto, Campus Agrário de Vairão, Rua Padre Armando Quintas, 4485-661 Vairão, Portugal.

Corresponding author:

Dr. Alexandra Zieritz

School of Geography, University of Nottingham, Nottingham, UK

Phone: 0115 84 68137
Email: Alexandra.zieritz@nottingham.ac.uk

Acknowledgements

This research was developed under ConBiomics: the missing approach for the Conservation of freshwater Bivalves Project Nº NORTE-01-0145-FEDER-030286, co-financed by COMPETE 2020, Portugal 2020 and the European Union through the ERDF, and by the Portuguese Foundation for Science and Technology (FCT) through national funds. FCT also supported AGS (SFRH/BD/137935/2018) and SV (UID/AGR/04033/2019). AZ was supported by an Anne McLaren Research Fellowship from the University of Nottingham. INB was supported by the Ministry of Science and Higher Education of Russia (Project No. AAAA-A18-118012390161-9, and Project No. FSRU-2020-0007), and the Russian Scientific Fund (Project No. 19-04-00066).

Abstract

Unionidae represent an excellent model taxon for unravelling the drivers of freshwater diversity, but phylogeographic studies on Southeast Asian taxa are hampered by a lack of a comprehensive phylogeny and mutation rates for this fauna. We present complete female- (F) and male-type (M) mitogenomes of four genera of the Southeast Asian clade Contradentini + Rectidentini. We calculated substitution rates for the mitogenome, the 13 protein-coding genes, the two ribosomal units and three commonly used fragments (*co1*, *nd1* and *16S*) of both F- and M-mtDNA, based on a fossil-calibrated, mitogenomic phylogeny of the Unionidae. Phylogenetic analyses, including an M+F concatenated dataset, consistently recovered a monophyletic Gonideinae. Subfamily-level topology was congruent with that of a previous nuclear genomic study and with patterns in mitochondrial gene order, suggesting Unionidae F-type 2 as a synapomorphy of the Gonideinae. Our phylogeny indicates that the clades Contradentini + Rectidentini and Lamprotulini + Pseudodontini + Gonideini split in the early Cretaceous (~125 Mya), and that the crown group of Contradentini + Rectidentini originated in the late Cretaceous (~79 Mya). Most gonideine tribes originated during the early Palaeogene. Substitution rates were comparable to those previously published for F-type *co1* and *16S* for certain Unionidae and Margaritiferidae species (pairs).

Keywords: evolutionary biogeography; freshwater mussels; palaeogeography; mitochondrial DNA; substitution rate; selection; tropical biodiversity

Introduction

Southeast Asia (SE-Asia) is a region of exceptional biodiversity, being comprised by four of the world's 35 biodiversity hotspots (Mittermeier et al., 2011). Understanding patterns and drivers of this diversity is crucial for prioritising conservation efforts, particularly in freshwater habitats, where species are being lost faster than in any other realm (Dudgeon et al., 2006; Reid et al., 2019). Freshwater mussels of the order Unionida represent an excellent model taxon for understanding causes of and threats to tropical freshwater diversity, owing to their restriction to freshwater habitats throughout their lifecycle, exceptional diversity in the region, and high sensitivity to habitat degradation and other anthropogenic factors (Wächtler et al., 2001; Gallardo et al., 2018; Lopes-Lima et al., 2018; Zieritz et al., 2018a; Zieritz et al., 2018b). Unfortunately, biogeographical studies on SE-Asian Unionida are currently hampered by a lack of a comprehensive phylogeny of this fauna. SE-Asian Unionida have received little scientific attention until recent intensification of sampling and sequencing efforts (Pfeiffer & Graf, 2015; Zieritz et al., 2016; Bolotov et al., 2017a; Bolotov et al., 2017b; Zieritz et al., 2018a; Konopleva et al., 2019; Zieritz et al., 2020).

The vast majority (99%) of the approximately 135 Unionida species from SE-Asia fall into the family Unionidae, comprising 714 currently recognised species globally (Zieritz et al., 2018a; Graf & Cummings, 2019). The evolutionary relationships among suprageneric clades of the Unionidae has been the subject of a number of recent studies, but the number of these clades and relationships between them remain contentious. Based on a two-locus dataset, Lopes-Lima et al. (2017b; Fig. 1) divided the Unionidae into seven subfamilies, two of which exhibit a SE-Asian centre of diversity: (1) the Rectidentinae (=Contradentini + Rectidentini; 36 species, 7 genera), which are exclusively distributed in SE-Asia, and (2) the Gonideinae (=Chamberlainiini + Lamprotulini + Gonideini + Pseudodontini; 69 species, 15 genera), with a disjunct distribution in Asia, Europe, Africa and North America. The Gonideinae *sensu*

Lopes-Lima et al. (2017b) were subsequently split by Bolotov et al. (2017a) into Pseudodontinae (=Pseudodontini + Pilsbryoconchini) and Gonideinae (=Gonideini) based on a three-locus dataset, which recovered a paraphyletic Gonideinae (Fig. 1). Paraphyly of the Gonideinae *sensu* Lopes-Lima et al. (2017b) was also recovered in Huang et al.'s (2019) trees, albeit with poor nodal support (Fig. 1). Finally, utilising for the first time, a genomic dataset of 596 nuclear loci, Pfeiffer et al. (2019) reintroduced the Pseudodontinae as Pseudodontini in the Gonideinae and additionally included the Rectidentinae based on the consistent recovery and strong nodal support of this large monophyletic clade (i.e. Gonideinae *sensu* Pfeiffer et al. (2019)) (Fig. 1).

As an alternative to nuclear markers, mitochondrial DNA (mtDNA) has long been popular in population genetic and phylogenetic studies, owing to their generally high mutation rates and fast lineage sorting (Brown et al., 1979; Birky et al., 1983). The unionidan mitogenome is typical of Metazoa in that it encodes 13 proteins that belong to four enzyme complexes of the respiratory chain, the small and large RNA subunits of the mitochondrial ribosomes, and 22 tRNAs (Bernt et al., 2013a). However, the Unionidae, as well as at least two other unionidan families and a number of other bivalve groups, are very distinctive in their doubly-uniparental mode of mitochondria inheritance (Breton et al., 2009; Gusman et al., 2016). Female (F)-type mtDNA is transmitted through mothers to both male and female offspring, whilst male (M)-type mtDNA is transmitted through fathers only to male offspring and established in the germ line (Breton et al., 2007). M- and F-type mtDNA often exhibit nucleotide divergences of >20% and thus provide two independent datasets for phylogenetic reconstructions (Breton et al., 2007). In addition, mitochondrial gene order has been identified as a very valuable character in itself for supporting deep nodes (Lopes-Lima et al., 2017a). This was recently confirmed by Froufe et al. (2020), who, by combining, for the first time, M- and F-type mitogenomes, revealed that the Gonideinae *sensu* Lopes-Lima et al. (2017b) exhibits unique gene orders (i.e.

UF2 and UF3) that are not found in any taxa outside this clade. To reflect more accurately the presence of several levels of highly divergent clades within the Unionidae, Froufe et al. (2020) additionally proposed a new systematic framework with three instead of two family-group levels (i.e. subfamilies, tribes and subtribes). Here, we use two levels of these higher taxa, i.e. subfamilies and tribes (see Table 1), because subtribes were erected within a few Unionidae tribes only (Froufe et al., 2020). As the phylogeny of Froufe et al. (2020) did not include any members of the Rectidentinae *sensu* Lopes-Lima et al. (2017b), hypotheses on the origin of this important SE-Asian clade have never been tested using a mitogenome dataset.

In addition to a robust phylogenetic hypothesis, identifying the drivers of past diversification events requires knowledge of the timing (as well as place) of these evolutionary events. This information can be reconstructed through phylogenetic trees (or networks) that are time-calibrated using palaeogeographical events, the fossil record or external molecular clocks (themselves calibrated by either palaeogeographical events or fossils) (Wilke et al., 2009). Fossil-calibrated phylogenies have been constructed for the unionidan families Hyriidae (Graf et al., 2015), Unionidae (Bolotov et al., 2017a), and the Unionidae + Margaritiferidae (Froufe et al., 2020). Substitution (=mutation) rates vary considerably among genes, taxa and with time (Wilke et al., 2009). However, substitution rates have been calculated for only a small number of unionidan taxa and genes so far, all of which revealed comparatively low evolutionary rates. Froufe et al. (2016) and Araujo et al. (2016) determined almost identical mean rates for F-type co1, i.e. 2.7 and 2.5 × 10^{-9} s/s/y [substitutions/site/year] for *Unio delphinus*/*foucauldianus* (Unionidae, Unioninae) and *Potomida littoralis* (Unionidae, Gonideinae), respectively, using the Messinian Salinity Crisis (5.96 to 5.33 Mya) as their calibration point. Bolotov et al. (2016) obtained similar to slightly lower mean substitution rates for four pairs of margaritiferid sister species, with an overall mean rate of 2.16 × 10^{-9} s/s/y using fossil calibration points. In the same paper, substitution rates for the margaritiferid *16S* are also provided, averaging 1.33 ×

10^{-9} s/s/y, which exceed the rate for combined Unionidae and Margaritiferidae obtained by Lydeard et al. (1996) of around 0.50×10^{-9} s/s/y. These evolutionary rates are about 2-10 times slower than for other molluscs (Stepien et al., 1999; Marko, 2002; Wilke et al., 2009), which has been attributed to the comparatively long generation times, longevity and low metabolic rates of freshwater mussels (Araujo et al., 2016; Bolotov et al., 2016). Similarly low or even lower evolutionary rates have, however, been recorded in coelacanths, anthozoan corals, salmonids, Acipenseriformes and Testudines (see Bolotov et al. (2016) and references therein).

This study aims to (1) test the validity of the Gonideinae *sensu* Pfeiffer et al. (2019) on the basis of two independent, complete mitogenomes (M-type, F-type and concatenated), for the first time including four genera of the Rectidentini; (2) infer implications on the evolutionary biogeography of this clade, including potential causes of divergence events, using a fossil-calibrated mitogenomic approach; and (3) estimate molecular substitution rates for all coding and two rRNA (*12S* and *16S*) genes, plus the three most commonly used mitochondrial gene fragments (i.e. *co1*, *16S* and *nd1*).

Materials and methods

Sampling, DNA extraction, sequencing, assembly and annotation

One male specimen from each of two species of the Contradentini, i.e. *Lens contradens* and *Physunio superbus*, and Rectidentinae, i.e. *Rectidens sumatrensis* and *Hyriopsis bialata*, respectively, was dissected for gonadal (to recover M-type mtDNA) and mantle (to recover F-type mtDNA) tissue collections. Specimens are deposited in the Museum of Zoology, University of Malaya (Freshwater mussel collection lots #12, 20, 46 and 59). DNA extractions of both tissues for each species were obtained using a standard high-salt protocol (Sambrook et al., 1989).

The genomic DNA was processed using Nextera-based library preparation (Illumina, USA) following the manufacturer's instructions. Quantification and size estimation of the library was then completed on a Bioanalyzer 2100 High Sensitivity DNA chip (Agilent, USA). The library was then normalized to 2 nM and sequenced on a MiSeq Benchtop Sequencer (2 × 250 bp paired-end reads) (Illumina, USA).

The mitochondrial genomes were reconstructed with MITObim (Hahn et al., 2013) using previously available co1 gene sequences for each species. Annotations were performed using MITOS (Bernt et al., 2013b) adjusting the final tRNAs gene limits of tRNA genes with ARWEN (Laslett & Canback, 2008). In-house scripts (https://figshare.com/s/a756ef19cec8f65d506a) were also applied to adjust the mtDNA protein-coding limits given that MITOS often underestimates gene length.

All F- and M-type mitogenomes were visualized using GenomeVx (Conant & Wolfe, 2008) and have been deposited in GenBank under the accession numbers XX000000–XX000000 [will be provided after acceptance of the manuscript].

Mitochondrial protein-coding genes nucleotide composition analysis

Nucleotide compositional bias of the sequenced mitogenomes was summarized as GC and AT skews according to the equations: AT skew = (A – T) / (A + T), GC skew = (G – C) / (G + C) (Perna & Kocher, 1995).

Phylogenetic inference

Datasets and sequence alignments

To infer the phylogenetic relationships within Unionida, we used the eight newly sequenced haplotypes together with Unionida mitogenomes available at NCBI (Table 1). In all phylogenetic analyses, we used the DNA sequences of all mtDNA protein-coding genes (PCG), except atp8 and the gender-specific open reading frames (*M-orf, H-orf* and *F-orf*; Breton et al., 2009). The sequences of the two rRNA genes were also included in all analyses. Each gene sequence was then aligned using the stand-alone version of GUIDANCE2 (Sela et al., 2015) with the MAFFT multiple sequence alignment algorithm (version 7, Katoh & Standley, 2013). GUIDANCE builds a high-quality alignment and assigns confidence scores for each sequence, column or position in the alignment based on the guide-tree uncertainty. To build our single gene alignments we used the following GUIDANCE2 parameters: score algorithm: GUIDANCE2; bootstraps replicates: 100; Sequence cutoff score: 0.0 (no sequence removal); Column cutoff score: below 0.8; Site masking score: below 0.6 (for codon and amino acids alignments) and below 0.8 (for the rRNA alignments). The resulting single gene alignments were finally concatenated into 12,779 nucleotides (nt) for the F- type and 13,370 nt for the M-type (12 PCG plus 2 rRNA sequences). The resulting M+F concatenated alignment spanned 26,150 nt.

Substitution model selection and partitioning schemes

All concatenated dataset alignments were partitioned, and the best-fit substitution models selected with PartitionFinder2 using a greedy search approach, MrBayes model set and BIC selection criterion (Lanfear et al., 2017) (Table S1).

Phylogenetic analyses

Phylogenetic inference applying a Maximum Likelihood (ML) methodology was performed to estimate the trees for all concatenated alignments using RAxML (version 8.2.10, Stamatakis, 2014) with 100 rapid bootstrap replicates and 20 ML searches.

To infer the phylogeny with Bayesian methodology we used MrBayes v3.2.7a (Ronquist et al., 2012). The alignment was partitioned according to the best scheme suggested by PartitionFinder. We applied a separate molecular substitution model for each partition. However, the prior probabilities of topology and branch lengths were linked across the partitions. Parameters were estimated as part of the analysis with four Markov chains incrementally heated with the default heating values. Each chain started with a randomly generated tree and ran for 1×10^7 generations with a sampling frequency of 1 tree for every 1000 generations. The resultant 10,000 trees, after discarding the first 25% as burn-in, were combined in a 50% majority rule consensus tree. Two independent replicates were conducted and inspected for consistency. The log files were checked visually with Tracer v. 1.7 for an assessment of the convergence of the MCMC chains and the effective sample size of parameters (Rambaut et al., 2018). The ESS values for all parameters from both BI phylogenies were recorded as >750. Tree topology differences were checked for significance with the KH- (Kishino & Hasegawa, 1989), SH- (Shimodaira & Hasegawa, 1999), and approximately unbiased (AU)-tests (Shimodaira & Hasegawa, 1999; Shimodaira, 2002) with 10,000 bootstrap replicates all implemented in IQ-Tree 2 (Minh et al., 2020).

Calibrated Time Phylogeny

The time-calibrated mitogenomic phylogenies were reconstructed in BEAST v. 1.10.1 for F-type and M-type mtDNA separately based on two reliable fossil calibrations with exponential prior distributions (Table S2), a lognormal relaxed clock and Yule speciation process as the tree prior (Suchard et al., 2018). Calculations were performed at the San Diego Supercomputer Center (SDSC, University of California, San Diego, USA) through the CIPRES Science Gateway (Miller et al., 2010). Three Margaritiferidae mitogenomes were used as outgroups for the F-type alignment, and one for the M-type alignment. Similar settings to each gene partition as in the MrBayes analyses were specified but using a simplified substitution model (HKY; see Bolotov et al. (2017a) for details). Three replicate BEAST searches were conducted, each with 50,000,000 generations and a tree sampling every 1000th generation. The log files were checked visually with Tracer v. 1.7 for an assessment of the convergence of the MCMC chains and the effective sample size of parameters (Rambaut et al., 2018). The runs were compiled with LogCombiner v. 1.10.1 (Suchard et al., 2018) using an appropriate burn-in depending on the start of convergence of MCMC chains in each run and an additional re-sampling every 5000th generation. The ESS values for all parameters were recorded as >300. The maximum clade credibility tree was obtained from the post-burn-in trees using TreeAnnotator v. 1.10.1 (Suchard et al., 2018).

Molecular rate calculations and analyses

After obtaining the age estimates in BEAST, for each node, and each gene, average sequence divergence among each node's subgroups (always 2, as the tree was perfectly dichotomous) was calculated in MEGA-X (Kumar et al., 2018), using the Maximum Composite Likelihood substitution model, with 1000 bootstraps, Gamma distributed rates among sites (G=1), and pairwise deletion of gaps. Divergence values were then divided by the respective node's average age, and 95%CI minimum and maximum ages, in order to obtain the average and

95%CI divergence rate per year, respectively. Average substitution rates per lineage (μ) were then obtained dividing those values by 2. Rates were calculated in this way for each gene, fragment and the whole mitogenome.

To examine and quantify the effects of type (M- *vs*. F-type) and node age on divergence rates, we ran General Linear Models in R version 3.5.2 for each gene, fragment and the whole mitogenome, fitting "ln (divergence rate)" as a response variable, "type" as a factor with two levels, "ln (node age)" as a covariate, and the interaction of both terms. Non-significant terms ($P \geq 0.05$) were sequentially dropped from the model.

The obtained substitution rates account for both synonymous and non-synonymous substitutions, and are therefore affected by the molecular mutation rate and the selective process. Rate variation among genes could therefore be linked to the strength of selection they are subject to. In order to better understand such constraints, we calculated the mean ratio of the number of nonsynonymous substitutions per non-synonymous site (K_a) to the number of synonymous substitutions per synonymous site (K_s) for each PCG and the whole coding portions of the mitogenome using DnaSP (Rozas et al., 2017). K_a/K_s values >1 indicate positive (directional) selection, <1 negative (stabilizing) selection, and 1 neutrality.

Finally, we examined the relationship between the differences in F-type and M-type μ and K_a/K_s-ratios across the 13 PCGs through regression analysis.

Results

Mitogenome characteristics and gene order

All the eight sequenced haplotypes include the 13 protein-coding genes (PCGs) typically found in metazoan mitochondrial genomes, the type-specific *orf* described for all Unionida mitogenomes with the DUI system and 22 transfer RNA (tRNA) and two ribosomal RNA (rRNA) genes (Fig. 2). The length of M-type mitogenomes is larger than the F-type as usual in the Unionida, ranging from 16,761 nt in *H. bialata* to 17,282 nt in *P. superbus*, while the F-type ranged from 15,956 nt in *L. contradens* to 16,057 nt in *P. superbus* (Table 2, Fig. 2). The A+T content, and GC and AT skews are similar across all sequenced species and types, averaging 62%, 0.38 (+ strand) and -0.20 (+ strand), respectively (Table 2). Gene order of the mitogenomes for all four newly sequenced species was of Unionidae F-type 2 (UF2) for F- and Unionidae M-type 1 (UM1) for M-type mitogenomes (Lopes-Lima et al., 2017a).

Phylogenetic analyses

All phylogenies inferred in this study support the monophyly of Gonideinae *sensu* Pfeiffer et al. (2019), monophyly of the six gonideine tribes (Figs. 1, 3 and 4). The only inconsistent result is that the (Gonideini + Pseudodontini) clade is not supported in M-type mtDNA phylogeny (Figs. 1 and 3). However, topology tests of the concatenated (M + F) tree constrained with the separate M-type phylogeny was not significantly different to that of the unconstrained tree ($P>0.05$ in KH-, SH- and AU-tests). The Chamberlainiini were consistently recovered as sister to other Gonideinae tribes in all phylogenies.

Fossil-calibrated phylogeny

The fossil-calibrated phylogeny of the complete F-type mitogenome dataset placed the split between Ambleminae (North American) and (Gonideinae + Unioninae) (W-North American + European + Asian + African) in the middle Jurassic (mean age = 164 Mya, 95% HPD 155-185

307 Mya) as the most ancient divergence event within the family (Fig. 5, Fig. S1A). Gonideinae
308 (W-North American + European + Asian + African) and Unioninae (European + Asian +
309 African) split in the late Jurassic (mean age = 152 Mya, 95% HPD = 138-175 Mya), coinciding
310 with a change in gene order in the most recent common ancestor (MRCA) of the Gonideinae.
311 The Chamberlainiini (SE-Asian) split from the rest of the Gonidaeinae in the early Cretaceous
312 (mean age = 142 Mya, 95% HPD = 126-164 Mya), followed by evolution of a new gene order
313 in the MRCA of *Chamberlainia hainesiana*. The clades Contradentini + Rectidentini (SE-
314 Asian) and Lamprotulini + Pseudodontini + Gonideini (W-North America, Europe + Asia) split
315 from each other in the early Cretaceous (mean age = 125 Mya, 95% HPD = 106-147 Mya).
316 The tribes Contradentini and Rectidentini (both SE-Asia-Indo-Burma-Sundaland) split from
317 each other in the late Cretaceous (mean age = 79 Mya, 95% HPD = 58-101 Mya). The crown
318 groups of most recent tribes likely originated between 64 and 37 Mya, i.e. since the Cretaceous
319 (during the Palaeocene and Eocene). In contrast, the tribe Chamberlainiini had a more ancient
320 crown group originating near the Albian-Cenomanian boundary (mean age = 103 Mya, 95%
321 HPD = 72-132 Mya).

322 **Mutation rates and K_a/K_s**

323 Mean substitution rates (μ) per mitochondrial gene ranged from 1.14 (*rrnS*) to 6.77 (*atp8*), and
324 1.42 (*rrnS*) to 7.71 (*atp8*) s/s/y × 10^{-9} in F-type and M-type, respectively (Table 3, Fig. 6a).
325 *Atp8* thereby exhibited by far the highest μ in both F- and M- type mitogenomes, followed by
326 *nad 2-6*. The lowest μ were recorded for the two ribosomal genes and the cytochrome c oxidase
327 subunits with the exception of type M-*cox2*, which exhibits a particularly variable extension.

328 Substitution rates across the whole mitogenome were strongly and significantly affected by
329 'type' and 'node age', which explained 25 and 35% of the variation, respectively (Table 3),
330 with rates being higher in M-type mtDNA and decreasing with node age (Fig. S1B). Similarly,
331 highly significant effects of 'type' and 'node age' on μ were found for 12 of the 15 mtDNA

genes. Exceptions were *atp8* (no effect of 'type' and 'node age'), and *nad4l* and *nad6* (no effect of 'node age'). Across the whole mitogenome, M-type μ was 31% higher than F-type μ (Table 3). This difference was lowest in *atp8* (13%), *cob* (22%) and the two ribosomal genes (17 and 23%, respectively), and highest in *cox2* (61%), *nad6* (51%), *nad4l* (49%) and *nad3* (46%). Absolute differences in M-type and F-type μ ranged from 0.3 (for ribosomal genes) and 0.5 (for *cox1*) to 2.1 (for *nad6*) s/s/y $\times 10^{-9}$ (Table 3, Fig. 6b).

Mean K_a/K_s per PCG ranged from 0.12 (*cob*, *cox1*) to 0.51 (*atp8*) in M-type and 0.02 (*cox1*) to 0.39 (*atp8*) in F-type mtDNA (Table 3). K_a/K_s was significantly positively correlated with μ, explaining 82% of the variation in μ ($F_{1,24}$=110.2, P<0.0001; Fig. 6a). In addition, M-type K_a/K_s was consistently and significantly higher than F-type K_a/K_s across PCGs (ANOVAs: P<0.001). This difference between M-type and F-type K_a/K_s was particularly pronounced in *nad3*, *nad4l* and *nad2*, and particularly small in *cob* and *nad1* (Fig. 6b). In addition, there was a significant relationship between the difference in μ and difference in K_a/K_s between M-type and F-type PCGs ($F_{1,11}$=5.728, P=0.0357), with the difference in K_a/K_s explaining 34% of the variation in the difference in μ (Fig. 6b).

Discussion

Phylogeny

Our mitogenomic phylogenetic hypotheses of the Gonideinae are congruent with those recovered on the basis of a nuclear genomic dataset by Pfeiffer et al. (2019). This recovers the Ambleminae as sister to the remaining, sampled unionid subfamilies and the Gonideinae as a monophyletic group, and is in contrast to most previously published phylogenies based on 2- or 3-locus datasets, which recovered a paraphyletic Gonideinae with respect to at least the Ambleminae (Fig. 1). The consistency of results across nuclear markers and mitogenomic datasets adds confidence to this phylogenetic hypothesis of the Gonideinae. In addition, the hypothesis is congruent with patterns in F-type mtDNA gene order, with the ancestral Unionidae order UF1 being retained in the Ambleminae and (at least) the Unioninae, and evolution of UF2 in the common ancestor of the Gonideinae and subsequent evolution of the UF2 variant UF3 in (at least) *Chamberlainia hainesiana*. If, for example, Lopes-Lima et al.'s (2017b) or Huang et al.'s (2019; fig. 2) hypotheses were assumed, either UF1 or UF2 would have had evolved independently twice (e.g. UF1 in the Unioninae and Ambleminae), which would represent a less parsimonious situation.

Evolutionary biogeography

Our updated fossil-calibrated phylogeny revealed a Mesozoic origin of Unionidae, placing its crown group in the early Jurassic. These results are in full agreement with earlier studies of divergence patterns within this family based on multiple fossil calibrations (Bolotov et al., 2017a; Bolotov et al., 2017b). Furthermore, the ages inferred from our fossil-calibrated phylogeny align with those of Bolotov et al. (2020), which were reconstructed using an external mutation rate, i.e. of Froufe et al. (2016).

Here, the Contradentini + Rectidentini was recovered as an ancient clade (mean age = 79 Mya), supporting the hypothesis of a late Cretaceous origin of Southeast Asian freshwater mussel fauna (Bolotov et al., 2017a; Bolotov et al., 2017b). We found that the crown groups of most gonideine tribes likely originated during the Palaeocene and Eocene. These key divergence events in the subfamily roughly coincide with a warm and humid climatic episode in the early Palaeogene (Wing et al., 2005; Gingerich, 2006), when mid-latitude mean annual air temperatures reached 23–29°C (Naafs et al., 2018). The origin of the majority of gondeine tribes could be explained by long-distance dispersal triggered by favourable climatic and hydrological conditions in the early Palaeogene (Carmichael et al., 2017) followed by diversification processes caused by range fragmentation during subsequent cold and dry periods (Feng et al., 2013). Evidence for rapid range expansion during the Palaeocene-Eocene Thermal Maximum has been found in several other groups, including reptiles (Bourque et al., 2015), mammals (Smith et al., 2006; Burger, 2012) and plants (Wing et al., 2005; Wing & Currano, 2013). A similar but younger "evolutionary burst" roughly coinciding with the Miocene Climatic Optimum, a warm and humid period, was discovered in genus-level clades of the radicine pond snails (Lymnaeidae) (Aksenova et al., 2018). The authors suggested that subsequent aridification periods led to fragmentation of continuous ranges and thus radiation in multiple radicine lineages in suitable refugia.

Variation in mutation rates

Comparison with previously published substitution rates and potential applications

Substitution rates determined in this study are comparable to those already available for other Unionida groups (Lydeard et al., 1996; Araujo et al., 2016; Bolotov et al., 2016; Froufe et al., 2016). The mean substitution rate for the (most commonly used mtDNA fragment) F-type *co1* of 1.6×10^{-9} s/s/y reported here lies in between those reported for four margaritiferid species pairs (Bolotov et al., 2016) and two unionid species (pairs) (Araujo et al., 2016; Froufe et al.,

2016). The rate for F-type *16S* of 0.83×10^{-9} s/s/y is slightly slower than those reported on Margaritiferidae (Bolotov et al., 2016) but exceeds those inferred from combined Unionidae and Margaritiferidae (Lydeard et al., 1996). It was expected that Margaritiferidae may share slower substitution rates because of their generally longer generation times and slower metabolic rates compared to Unionidae (Bolotov et al., 2016), but this supposition was not supported by our novel results.

Molecular clocks can overcome some gaps in the fossil record by producing informative date estimates for fossil-poor evolutionary events or those for which available fossil data have limited power (e.g. due to the lack of reliable distinguishing features for morphological identification) (Donoghue & Yang, 2016). Fields of application include (1) estimating clade ages in phylogenetic and phylogeographic studies when fossil and palaeobiogeographic data are not available; (2) studying complex geological histories; and (3) validating time-estimations of geological and tectonic events, such as the separation of land masses or the formation and evolution of river basins (Pennington et al., 2004; Ketmaier et al., 2006; Roxo et al., 2014; Chen et al., 2016). However, careful selection of the substitution rate applied in a given context is crucial as effectiveness and accuracy of a given rate varies depending on the time scale of the event. For example, applying the power test by Wilke et al. (2009) and a type II error of 0.05, data variability in *co1* (~710 nt), *nd1* (~830 nt) and *16S* (~570 nt) is insufficient to reliably date phylogenetic events that are younger than 2.6, 1.1 and 6.3 Mya, respectively; at least when used independently. The matter is further complicated by time dependency and incomplete coalescence in shallower phylogenetic depths. Possible rate heterogeneity across branches should be considered (e.g. using an uncorrelated clock model in phylogenetic estimation), which is why we performed rate calculations for every node in the tree, and provide a rate confidence interval that should be used to guide future applications (Table 3, Fig. S2). Our dataset allows aging of any node within the unionid phylogeny and therefore, application at

various taxonomic levels, as substitution rates can be calculated for any given node age and gene (fragment) using the regression equations given in Fig. S2.

Variation across genes

In comparison to nuclear DNA, mutations in mtDNA are rapidly accumulated, which has been attributed to an error-prone DNA repair system, the lack of protective histone-like protein, exposure of single-strand intermediates during mitochondrial replication, and/or exposure to oxidative damage (Castellana et al., 2011; Bernt et al., 2013a; and references therein). In addition, recombination is usually absent, and the genetically effective population size of animal (F-type) mtDNA is estimated to be only a quarter that of nuclear DNA (nDNA) due to (doubly)-uniparental inheritance (DUI) and haploidy (Birky et al., 1983; Castellana et al., 2011); effective population size of M-type DNA being even smaller due to fewer mtDNA copies in the sperm compared to oocyte (Zouros, 2013). However, loss of functionality of mtDNA, including genes encoding for protein complexes that play vital roles in aerobic respiration, is prevented predominantly by strong negative selection, removing nonsynonymous deleterious mutations (Castellana et al., 2011). In addition, evidence for positive selection in mtDNA has been found for some genes and taxa (Bazin et al., 2006; Oliveira et al., 2008; Śmietanka et al., 2010).

In accordance with previous studies on other taxa, our dataset showed that substitution rates vary considerably among unionid mtDNA genes. As expected, rRNA fragments evolved much more slowly, which has been shown in various taxa before and partly attributed to high AT content (Brown, 1985; DeSalle et al., 1987). With respect to PCGs, substitution rates of *atp8* exceeded those of *cox1* by a factor of about 20 in F-mtDNA and a factor of 5 in M-mtDNA, values that are comparable to those of Śmietanka et al. (2010; fig. 5) on *Mytilus* spp. To a large degree, this variation in substitution rates among PCGs appears to be associated with a variation in negative selection pressure as indicated by the strong positive correlation between K_a/K_s and

μ of our dataset. Strongly relaxed negative selection in *atp8* has been observed in several other taxa, including whitefish, vertebrates and *Mytilus* spp. (Śmietanka et al., 2010; Castellana et al., 2011; Jacobsen et al., 2016). Sun et al. (2017) suggested that negative selection of *atp8* was already strongly relaxed in the most recent common ancestor of all bivalves and that the gene appears to have been lost in several marine bivalves. However, positive directional selection or the Compensatory-Draft Feedback (CDF) process (i.e. positive selection of compensatory mutations for mild-deleterious mutations fixed by e.g. genetic drift; Oliveira et al., 2008) may also play a role as has been shown for *atp8* and other PCGs in, for example, mammals (da Fonseca et al., 2008), parasitic wasps (Oliveira et al., 2008), billfish (Dalziel et al., 2006) and *Mytilus* spp. (Śmietanka et al., 2010).

Variation between mtDNA types

As expected and observed in a number of previous studies (e.g. Liu et al., 1996; Hoeh et al., 2002), our study confirmed that the unionid M-type mitogenome evolves faster than the F-type mitogenome (overall by 31%). This has been attributed to relaxed negative selection in M-type mtDNA and the CDF (Burzyński et al., 2017 and references therein), which is at least partly supported by a significant association between the difference in M-type and F-type substitution rates and difference in M-type and F-type K_a/K_s in our dataset. That said, again, positive selection may play a considerable role, especially in the ~550 nt extension in M-type *cox2* (Chapman et al., 2008).

Conclusions

Our mitogenomic phylogeny confirmed the validity of the Gonideinae *sensu* Pfeiffer et al. (2019), which are characterised by a synapomorphic mitogenome gene order (UF2). Fossil-calibration of this phylogeny, which was in line with one constructed using an external mutation rate (Bolotov et al., 2020), further revealed novel insights into the evolutionary

history and biogeography of the Unionidae, including evidence for rapid range expansion of several clades within the Unionidae during the Palaeocene-Eocene Thermal Maximum and origin of the SE-Asian Rectidentini in the late Cretaceous. Most importantly, the set of substitution rates of 30 mtDNA genes generated in our study will provide an important basis for future studies aiming to unravel the evolutionary history and biogeography of the Unionidae. Considering their exceptionally high levels of diversity, endemism and threat, particular focus should thereby be put on tropical freshwater mussels, including the Parreysiinae and Bornean endemic genera (Zieritz et al., 2018a; Zieritz et al., 2020).

References

Aksenova OV, Bolotov IN, Gofarov MY, Kondakov AV, Vinarski MV, Bespalaya YV, Kolosova YS, Palatov DM, Sokolova SE, Spitsyn VM. 2018. Species richness, molecular taxonomy and biogeography of the radicine pond snails (Gastropoda: Lymnaeidae) in the Old World. *Scientific Reports* **8:** 11199.

Araujo R, Buckley D, Nagel K-O, Machordom A. 2016. *Potomida littoralis* (Bivalvia, Unionidae) evolutionary history: slow evolution or recent speciation? *Zoological Journal of the Linnean Society* **179:** 277-290.

Bazin E, Glémin S, Galtier N. 2006. Population size does not influence mitochondrial genetic diversity in animals. *Science* **312:** 570-572.

Bernt M, Braband A, Schierwater B, Stadler PF. 2013a. Genetic aspects of mitochondrial genome evolution. *Molecular Phylogenetics and Evolution* **69:** 328-338.

Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, Putz J, Middendorf M, Stadler PF. 2013b. MITOS: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* **69:** 313-319.

Birky CW, Maruyama T, Fuerst P. 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics* **103:** 513-527.

Bolotov IN, Kondakov AV, Vikhrev IV, Aksenova OV, Bespalaya YV, Gofarov MY, Kolosova YS, Konopleva ES, Spitsyn VM, Tanmuangpak K, Tumpeesuwan S. 2017a. Ancient river inference explains exceptional Oriental freshwater mussel radiations. *Scientific Reports* **7:** 2135.

Bolotov IN, Konopleva ES, Vikhrev IV, Gofarov MY, Lopes-Lima M, Bogan AE, Lunn Z, Chan N, Win T, Aksenova OV, Tomilova AA, Tanmuangpak K, Tumpeesuwan S, Kondakov AV. 2020. New freshwater mussel taxa discoveries clarify biogeographic division of Southeast Asia. *Scientific Reports* **10:** 6616.

Bolotov IN, Vikhrev IV, Bespalaya YV, Gofarov MY, Kondakov AV, Konopleva ES, Bolotov NN, Lyubas AA. 2016. Multi-locus fossil-calibrated phylogeny, biogeography and a subgeneric revision of the Margaritiferidae (Mollusca: Bivalvia: Unionoida). *Molecular Phylogenetics and Evolution* **103:** 104-121.

Bolotov IN, Vikhrev IV, Kondakov AV, Konopleva ES, Gofarov MY, Aksenova OV, Tumpeesuwan S. 2017b. New taxa of freshwater mussels (Unionidae) from a species-rich but overlooked evolutionary hotspot in Southeast Asia. *Scientific Reports* **7:** 11573.

Bourque JR, Howard Hutchison J, Holroyd PA, Bloch JI. 2015. A new dermatemydid (Testudines, Kinosternoidea) from the Paleocene-Eocene Thermal Maximum, Willwood Formation, southeastern Bighorn Basin, Wyoming. *Journal of Vertebrate Paleontology* **35:** e905481.

Breton S, Beaupre HD, Stewart DT, Hoeh WR, Blier PU. 2007. The unusual system of doubly uniparental inheritance of mtDNA: isn't one enough? *Trends in genetics* **23:** 465-474.

Breton S, Beaupre HD, Stewart DT, Piontkivska H, Karmakar M, Bogan AE, Blier PU, Hoeh WR. 2009. Comparative mitochondrial genomics of freshwater mussels (Bivalvia: Unionoida) with doubly uniparental inheritance of mtDNA: gender-specific open reading frames and putative origins of replication. *Genetics* **183:** 1575-1589.

Brown W. 1985. The mitochondrial genome of animals. In: MacIntyre R, ed. *Molecular Evolutionary Genetics*. New York: Plenum. 95-130.

Brown WM, George M, Wilson AC. 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences* **76:** 1967-1971.

Burger BJ. 2012. Northward range extension of a diminutive-sized mammal (*Ectocion parvus*) and the implication of body size change during the Paleocene–Eocene Thermal Maximum. *Palaeogeography, Palaeoclimatology, Palaeoecology* **363:** 144-150.

Burzyński A, Soroka M, Mioduchowska M, Kaczmarczyk A, Sell J. 2017. The complete maternal and paternal mitochondrial genomes of *Unio crassus*: mitochondrial molecular clock and the overconfidence of molecular dating. *Molecular Phylogenetics and Evolution* **107:** 605-608.

Carmichael MJ, Inglis GN, Badger MP, Naafs BDA, Behrooz L, Remmelzwaal S, Monteiro FM, Rohrssen M, Farnsworth A, Buss HL. 2017. Hydrological and associated biogeochemical consequences of rapid global warming during the Paleocene-Eocene Thermal Maximum. *Global and Planetary Change* **157:** 114-138.

Castellana S, Vicario S, Saccone C. 2011. Evolutionary patterns of the mitochondrial genome in Metazoa: exploring the role of mutation and selection in mitochondrial protein coding genes. *Genome Biology and Evolution* **3:** 1067-1079.

Chapman EG, Piontkivska H, Walker JM, Stewart DT, Curole JP, Hoeh WR. 2008. Extreme primary and secondary protein structure variability in the chimeric male-transmitted cytochrome c oxidase subunit II protein in freshwater mussels: evidence for an elevated amino acid substitution rate in the face of domain-specific purifying selection. *BMC Evolutionary Biology* **8:** 165.

Chen W, Shen Y, Gan X, Wang X, He S. 2016. Genetic diversity and evolutionary history of the *Schizothorax* species complex in the Lancang River (upper Mekong). *Ecology and Evolution* **6:** 6023-6036.

Conant GC, Wolfe KH. 2008. GenomeVx: simple web-based creation of editable circular chromosome maps. *Bioinformatics* **24:** 861-862.

da Fonseca RR, Johnson WE, O'Brien SJ, Ramos MJ, Antunes A. 2008. The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics* **9:** 119.

Dalziel AC, Moyes CD, Fredriksson E, Lougheed SC. 2006. Molecular Evolution of Cytochrome c Oxidase in High-Performance Fish (Teleostei: Scombroidei). *Journal of Molecular Evolution* **62:** 319-331.

DeSalle R, Freedman T, Prager EM, Wilson AC. 1987. Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *Journal of Molecular Evolution* **26:** 157-164.

Donoghue PC, Yang Z. 2016. The evolution of methods for establishing evolutionary timescales. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371:** 20160020.

Dudgeon D, Arthington AH, Gessner MO, Kawabata Z, Knowler D, Lévêque C, Naiman RJ, Prieur-Richard AH, Soto D, Stiassny MLJ, Sullivan CA. 2006. Freshwater biodiversity: importance, status, and conservation challenges. *Biological Reviews* **81:** 163-182.

Feng X, Tang B, Kodrul TM, Jin J. 2013. Winged fruits and associated leaves of Shorea (Dipterocarpaceae) from the Late Eocene of South China and their phytogeographic and paleoclimatic implications. *American Journal of Botany* **100:** 574-581.

Froufe E, Bolotov I, Aldridge DC, Bogan AE, Breton S, Gan HM, Kovitvadhi U, Kovitvadhi S, Riccardi N, Secci-Petretto G, Sousa R, Teixeira A, Varandas S, Zanatta D, Zieritz A, Fonseca MM, Lopes-Lima M. 2020. Mesozoic mitogenome rearrangements and freshwater mussel (Bivalvia: Unionoidea) macroevolution. *Heredity* **124:** 182–196.

Froufe E, Gonçalves DV, Teixeira A, Sousa R, Varandas S, Ghamizi M, Zieritz A, Lopes-Lima M. 2016. Who lives where? Molecular and morphometric analyses clarify which Unio species (Unionida, Mollusca) inhabit the southwestern Palearctic region. *Organisms Diversity and Evolution* **16:** 597-611.

Gallardo B, Bogan AE, Harun S, Jainih L, Lopes-Lima M, Pizarro M, Rahim KA, Sousa R, Virdis SGP, Zieritz A. 2018. Current and future effects of global change on a hotspot's freshwater diversity. *Science of the Total Environment* **635:** 750-760.

Gingerich PD. 2006. Environment and evolution through the Paleocene–Eocene thermal maximum. *Trends in Ecology & Evolution* **21:** 246-253.

Graf DL, Cummings KS. 2019. The Freshwater Mussels (Unionoida) of the World (and other less consequential bivalves), updated 15 May 2019. MUSSEL Project Web Site.

Graf DL, Jones H, Geneva AJ, Pfeiffer III JM, Klunzinger MW. 2015. Molecular phylogenetic analysis supports a Gondwanan origin of the Hyriidae (Mollusca: Bivalvia: Unionida) and the paraphyly of Australasian taxa. *Molecular Phylogenetics and Evolution* **85:** 1-9.

Gusman A, Lecomte S, Stewart DT, Passamonti M, Breton S. 2016. Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. *PeerJ* **4:** e2760.

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.

Hoeh WR, Stewart DT, Guttman SI. 2002. High fidelity of mitochondrial genome transmission under the doubly uniparental mode of inheritance in freshwater mussels (Bivalvia: Unionoidea). *Evolution* **56:** 2252-2261.

Huang X-C, Su J-H, Ouyang J-X, Ouyang S, Zhou C-H, Wu X-P. 2019. Towards a global phylogeny of freshwater mussels (Bivalvia: Unionida): Species delimitation of Chinese taxa, mitochondrial phylogenomics, and diversification patterns. *Molecular Phylogenetics and Evolution* **130:** 45-59.

Jacobsen MW, da Fonseca RR, Bernatchez L, Hansen MM. 2016. Comparative analysis of complete mitochondrial genomes suggests that relaxed purifying selection is driving high nonsynonymous evolutionary rate of the NADH2 gene in whitefish (*Coregonus* ssp.). *Molecular Phylogenetics and Evolution* **95:** 161-170.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30:** 772-780.

Ketmaier V, Giusti F, Caccone A. 2006. Molecular phylogeny and historical biogeography of the land snail genus *Solatopupa* (Pulmonata) in the peri-Tyrrhenian area. *Molecular Phylogenetics and Evolution* **39:** 439-451.

Kishino H, Hasegawa M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *Journal of Molecular Evolution* **29:** 170-179.

Konopleva ES, Pfeiffer JM, Vikhrev IV, Kondakov AV, Gofarov MY, Aksenova OV, Lunn Z, Chan N, Bolotov IN. 2019. A new genus and two new species of freshwater mussels (Unionidae) from western Indochina. *Scientific Reports* **9:** 4106.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* **35:** 1547-1549.

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. *Molecular Biology and Evolution* **34:** 772-773.

Laslett D, Canback B. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* **24:** 172-175.

Liu HP, Mitton JB, Wu SK. 1996. Paternal mitochondrial DNA differentiation far exceeds maternal mitochondrial DNA and allozyme differentiation in the freshwater mussel, *Anodonta grandis grandis*. *Evolution* **50:** 952-957.

Lopes-Lima M, Burlakova LE, Karatayev AY, Mehler K, Seddon M, Sousa R. 2018. Conservation of freshwater bivalves at the global scale: diversity, threats and research needs. *Hydrobiologia* **810:** 1-14.

Lopes-Lima M, Fonseca MM, Aldridge DC, Bogan AE, Gan HM, Ghamizi M, Sousa R, Teixeira A, Varandas S, Zanatta D, Zieritz A, Froufe E. 2017a. The first Margaritiferidae male (M-Type) mitogenome: Mitochondrial gene order as a potential

character for determining higher-order phylogeny within Unionida (Bivalvia). *Journal of Molluscan Studies* **83:** 249–252.

Lopes-Lima M, Froufe E, Do VT, Ghamizi M, Mock KE, Kebapçı Ü, Klishko O, Kovitvadhi S, Kovitvadhi U, Paulo OS, Pfeiffer JM, Raley M, Riccardi N, Şereflişan H, Sousa R, Teixeira A, Varandas S, Wu X, Zanatta DT, Zieritz A, Bogan AE. 2017b. Phylogeny of the most species-rich freshwater bivalve family (Bivalvia: Unionida: Unionidae): Defining modern subfamilies and tribes. *Molecular Phylogenetics and Evolution* **106:** 174-191.

Lopes-Lima M, Hattori A, Kondo T, Hee Lee J, Ki Kim S, Shirai A, Hayashi H, Usui T, Sakuma K, Toriya T, Sunamura Y, Ishikawa H, Hoshino N, Kusano Y, Kumaki H, Utsugi Y, Yabe S, Yoshinari Y, Hiruma H, Tanaka A, Sao K, Ueda T, Sano I, Miyazaki J-I, Gonçalves DV, Klishko OK, Konopleva ES, Vikhrev IV, Kondakov AV, Yu. Gofarov M, Bolotov IN, Sayenko EM, Soroka M, Zieritz A, Bogan AE, Froufe E. 2020. Freshwater mussels (Bivalvia: Unionidae) from the rising sun (Far East Asia): phylogeny, systematics, and distribution. *Molecular Phylogenetics and Evolution* **146:** 106755.

Lydeard C, Mulvey M, Davis GM. 1996. Molecular systematics and evolution of reproductive traits of North American freshwater unionacean mussels (Mollusca: Bivalvia) as inferred from 16S rRNA gene sequences. *Philosophical Transactions (Series B): Biological Sciences* **351:** 1593-1603.

Marko PB. 2002. Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Molecular Biology and Evolution* **19:** 2005-2021.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE), 1-8.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37:** 1530-1534.

Mittermeier RA, Turner WR, Larsen FW, Brooks TM, Gascon C. 2011. Global biodiversity conservation: the critical role of hotspots. In: Zachos FE and Habel JC, eds. *Biodiversity Hotspots*. Berlin, Heidelberg: Springer. 3-22.

Naafs B, Rohrssen M, Inglis G, Lähteenoja O, Feakins S, Collinson M, Kennedy E, Singh P, Singh M, Lunt D. 2018. High temperatures in the terrestrial mid-latitudes during the early Palaeogene. *Nature Geoscience* **11:** 766.

Oliveira DC, Raychoudhury R, Lavrov DV, Werren JH. 2008. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). *Molecular Biology and Evolution* **25:** 2167-2180.

Pennington PT, Cronk QCB, Richardson JA, Lavin M, Schrire BP, Lewis G, Pennington RT, Delgado–Salinas A, Thulin M, Hughes CE, Matos AB, Wojciechowski MF.

675 **2004.** Metacommunity process rather than continental tectonic history better explains
676 geographically structured phylogenies in legumes. *Philosophical Transactions of the*
677 *Royal Society of London. Series B: Biological Sciences* **359:** 1509-1522.

678 **Perna NT, Kocher TD. 1995.** Patterns of nucleotide composition at fourfold degenerate sites
679 of animal mitochondrial genomes. *Journal of Molecular Evolution* **41:** 353-358.

680 **Pfeiffer JM, Breinholt JW, Page LM. 2019.** Unioverse: A phylogenomic resource for
681 reconstructing the evolution of freshwater mussels (Bivalvia, Unionoida). *Molecular*
682 *Phylogenetics and Evolution* **137:** 114-126.

683 **Pfeiffer JM, Graf DL. 2015.** Evolution of bilaterally asymmetrical larvae in freshwater
684 mussels (Bivalvia: Unionoida: Unionidae). *Zoological Journal of the Linnean Society*
685 **175:** 307-318.

686 **Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018.** Posterior summarization
687 in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67:** 901-904.

688 **Reid AJ, Carlson AK, Creed IF, Eliason EJ, Gell PA, Johnson PTJ, Kidd KA,**
689 **MacCormack TJ, Olden JD, Ormerod SJ, Smol JP, Taylor WW, Tockner K,**
690 **Vermaire JC, Dudgeon D, Cooke SJ. 2019.** Emerging threats and persistent
691 conservation challenges for freshwater biodiversity. *Biological Reviews* **94:** 849-873.

692 **Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu**
693 **L, Suchard MA, Huelsenbeck JP. 2012.** MrBayes 3.2: Efficient Bayesian
694 Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic*
695 *Biology* **61:** 539-542.

696 **Roxo FF, Albert JS, Silva GSC, Zawadzki CH, Foresti F, Oliveira C. 2014.** Molecular
697 phylogeny and biogeographic history of the armored Neotropical catfish subfamilies
698 Hypoptopomatinae, Neoplecostominae and Otothyrinae (Siluriformes: Loricariidae).
699 *PLOS ONE* **9:** e105564.

700 **Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-**
701 **Onsins SE, Sanchez-Gracia A. 2017.** DnaSP 6: DNA Sequence Polymorphism
702 Analysis of Large Data Sets. *Mol Biol Evol* **34:** 3299-3302.

703 **Sambrook J, Fritsch E, Maniatis T. 1989.** Molecular Cloning: A Laboratory Manual.

704 **Sela I, Ashkenazy H, Katoh K, Pupko T. 2015.** GUIDANCE2: accurate detection of
705 unreliable alignment regions accounting for the uncertainty of multiple parameters.
706 *Nucleic Acids Research* **43:** W7-W14.

707 **Shimodaira H. 2002.** An approximately unbiased test of phylogenetic tree selection.
708 *Systematic biology* **51:** 492-508.

709 **Shimodaira H, Hasegawa M. 1999.** Multiple comparisons of log-likelihoods with
710 applications to phylogenetic inference. *Molecular biology and evolution* **16:** 1114-
711 1114.

Śmietanka B, Burzyński A, Wenne R. 2010. Comparative genomics of marine mussels (*Mytilus* spp.) gender associated mtDNA: rapidly evolving atp8. *Journal of Molecular Evolution* **71:** 385-400.

Smith T, Rose KD, Gingerich PD. 2006. Rapid Asia–Europe–North America geographic dispersal of earliest Eocene primate *Teilhardina* during the Paleocene–Eocene thermal maximum. *Proceedings of the National Academy of Sciences* **103:** 11223-11227.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312-1313.

Stepien CA, Hubers AN, Skidmore JL. 1999. Diagnostic genetic markers and evolutionary relationships among invasive dreissenoid and corbiculoid bivalves in North America: Phylogenetic signal from mitochondrial 16S rDNA. *Molecular Phylogenetics and Evolution* **13:** 31-49.

Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* **4:** vey016.

Sun Se, Li Q, Kong L, Yu H. 2017. Limited locomotive ability relaxed selective constraints on molluscs mitochondrial genomes. *Scientific Reports* **7:** 10628-10628.

Wächtler K, Mansur MCD, Richter T. 2001. Larval types and early postlarval biology in naiads (Unionoida). In: Bauer G and Wächtler K, eds. *Ecology and Evolution of the Freshwater Mussels Unionoida*. Berlin: Springer-Verlag. 93-125.

Wilke T, Schultheiß R, Albrecht C. 2009. As Time Goes by: A Simple Fool's Guide to Molecular Clock Approaches in Invertebrates. *American Malacological Bulletin* **27:** 25-45.

Wing SL, Currano ED. 2013. Plant response to a global greenhouse event 56 million years ago. *American Journal of Botany* **100:** 1234-1254.

Wing SL, Harrington GJ, Smith FA, Bloch JI, Boyer DM, Freeman KH. 2005. Transient floral change and rapid global warming at the Paleocene-Eocene boundary. *Science* **310:** 993-996.

Zieritz A, Bogan AE, Froufe E, Klishko O, Kondo T, Kovitvadhi U, Kovitvadhi S, Lee JH, Lopes-Lima M, Pfeiffer JM, Sousa R, van Do T, Vikhrev I, Zanatta DT. 2018a. Diversity, biogeography and conservation of freshwater mussels (Bivalvia: Unionida) in East and Southeast Asia. *Hydrobiologia* **810:** 29-44.

Zieritz A, Bogan AE, Rahim KAA, Sousa R, Jainih L, Harun S, Razak NFA, Gallardo B, McGowan S, Hassan R, Lopes-Lima M. 2018b. Changes and drivers of freshwater mussel diversity and distribution in northern Borneo. *Biological Conservation* **219:** 126-137.

Zieritz A, Lopes-Lima M, Bogan AE, Sousa R, Walton S, Rahim KAA, Wilson J-J, Ng P-Y, Froufe E, McGowan S. 2016. Factors driving changes in freshwater mussel (Bivalvia, Unionida) diversity and distribution in Peninsular Malaysia. *Science of the Total Environment* **571:** 1069-1078.

Zieritz A, Taha H, Lopes-Lima M, Pfeiffer J, Wah GSK, Sulaiman Z, McGowan S, Rahim KA. 2020. Towards the conservation of Borneo's freshwater mussels: rediscovery of the endemic *Ctenodesma borneensis* and first record of the non-native *Sinanodonta lauta*. *Biodiversity and Conservation* **29:** 2235–2253.

Zouros E. 2013. Biparental inheritance through uniparental transmission: the doubly uniparental inheritance (DUI) of mitochondrial DNA. *Evolutionary Biology* **40:** 1-31.

759 **Table 1**. List of species analysed, GenBank references, and country.

TAXON	CODE	F-TYPE GenBank	M-TYPE GenBank	COUNTRY
Unionida				
Unionidae				
Ambleminae				
Lampsilis ornata	LamOrn	NC_005335	-	USA
Potamilus alatus	PotAla	KU559011	KU559010	China (Introduced)
Potamilus leptodon	PotLep	NC_028522	-	USA
*Quadrula quadrula**	QuaQua	NC_013658	FJ809751	USA
Toxolasma parvum	TaxPar	NC_015483	-	USA
*Venustaconcha ellipsiformis**	VenEll	FJ809753	NC_013659	USA
Gonideinae				
Chamberlainiini				
Chamberlainia hainesiana	ChaHai	NC_044110	MK994771	Thailand
*Sinohyriopsis cumingii**	SinCum	NC_011763	KC150028	China
Sinohyriopsis schlegelii	SinSch	NC_015110	-	China (Introduced)
Gonideini				
Microcondylaea bonellii	MicBon	NC_044111	MK994773	Italy
Ptychorhynchus pfisteri	PtyPfi	KY067440	-	China
Solenaia carinata	SolCar	NC_023250	KC848655	China
Solenaia oleivora	SolOle	NC_022701	-	China
Lamprotulini				
Lamprotula leai	LamLea	NC_023346	-	China
Lamprotula caveata[1]	LamCav	NC_030258	-	China
Potomida littoralis	PotLit	NC_030073	KT247375	Portugal
Pronodularia japanensis	ProJap	AB055625	AB055624	Japan
Pseudodontini				
*Monodontina vondembuschiana**	MonVon	NC_044112	MK994775	Malaysia
Pilsbryoconcha exilis	PilExi	NC_044124	MK994777	Malaysia
Contradentini				
Lens contradens	LenCon	MW242812	MW242813	Muar River (2.766N,

				102.396E), Malaysia
Physunio superbus	PhySup	MW242814	MW242815	Pahang River (3.486N, 103.074E), Malaysia
Rectidentini				
*Hyriopsis bialata**	HyrBia	MW242816	MW242817	Pahang River (3.486N, 103.074E), Malaysia
Rectidens sumatrensis	RecSum	MW242818	MW242819	Perak River (4.885N, 100.988E), Malaysia
Unioninae				
Unionini				
Aculamprotula coreana	AcuCor	NC_026035	-	South Korea
Aculamprotula tientsinensis	AcuTie	NC_029210	-	China
Aculamprotula tortuosa	AcuTor	NC_021404	-	China
Cuneopsis pisciculus	CunPis	NC_026306	-	China
Nodularia douglasiae	NodDou	NC_026111	-	China
Schistodesmus sp.[2]	Sch	NC_023806	-	China
*Unio crassus**	UniCra	KY290447	KY290450	Poland
Unio delphinus	UniDel	KT326917	KT326918	Portugal
Unio pictorum	UniPic	NC_015310	-	Poland
*Unio tumidus**	UniTum	KY021076	KY021073	Poland
Anodontini				
Anemina arcaeformis	AneArc	NC_026674	-	China
Anemina euscaphys	AneEus	NC_026792	-	China
*Anodonta anatina**	AnoAna	NC_022803	KF030962	Poland
Cristaria plicata	CriPli	NC_012716	-	China
Lasmigona compressa	LasCom	NC_015481	-	USA
Pyganodon grandis	PygGra	NC_013661	FJ809755	USA
Sinanodonta lucida	SinLuc	NC_026673	-	China
*Sinanodonta woodiana**	SinWoo	HQ283348	KM434235	China

Utterbackia imbecillis	UttImb	NC_015479	-	USA
*Utterbackia peninsularis**	UttPen	HM856636	NC_015477	USA
Lanceolariini				
Lanceolaria grayana	LanGra	NC_026686	-	China
Lanceolaria lanceolata	LanLan	NC_023955	-	China
Lepidodesmini				
Lepidodesma languilati	LepLan	NC_029491	-	China
Margaritiferidae				
Margaritifera dahurica	MarDah	NC_023942	-	China
Margaritifera falcata	MarFal	NC_015476	-	USA
Pseudunio marocanus	PseMrc	KY131953	KY131954	Morocco

760 * F- and M-type sequences from distinct specimens

761 [1] misidentified as *Aculamprotula scripta* in original Genbank submission (see Lopes-Lima et
762 al. (2020))

763 [2] misidentified as *Lamprotula gottschei* in original Genbank submission (see Huang et al.
764 (2019))
765

766

767 **Table 2**. Main structural features of the female- and male-transmitted mitochondrial genomes of two species of the Contradentini and Rectidentini,

768 respectively.

	♀ *F-Type*				♂ *M-Type*			
	Lens contradens	***Physunio superbus***	***Hyriopsis bialata***	***Rectidens sumatrensis***	***Lens contradens***	***Physunio superbus***	***Hyriopsis bialata***	***Rectidens sumatrensis***
Tot. size (pb)	15,956	16,057	15,919	15,983	17,057	17,282	16,761	16,934
A+T %	60.97	61.26	63.55	64.50	61.85	62.67	61.70	62.27
CG (+) skew	0.36	0.41	0.37	0.38	0.32	0.36	0.45	0.41
AT (+) skew	-0.20	-0.23	-0.20	-0.19	-0.17	-0.19	-0.23	-0.20
rrnS	837	842	842	834	887	885	861	862
rrnL	1304	1313	1323	1309	1279	1285	1302	1318
	Size (nt) (Start/Stop codons)				Size (nt) (Start/Stop codons)			
cox1	1539 (TTG/TAG)	1539 (TTG/TAA)	1539 (TTG/TAA)	1539 (TTG/TAA)	1611 (ATG/TAG)	1602 (ATG/TAG)	1596 (ATA/TAG)	1644 (ATT/TAG)
cox2	681 (ATG/TAG)	687 (ATG/TAG)	711 (ATT/TAG)	699 (GTG/TAG)	1272 (TTG/TAA)	1248 (ATG/TAA)	1329 (ATT/TAG)	1230 (ATA/TAG)
nad2	963 (ATG/TAG)	963 (ATG/TAA)	963 (ATG/TAA)	963 (ATG/TAA)	990 (ATA/TAA)	993(ATT/TAA)	984 (ATA/TAG)	984 (ATA/TAG)
nad3	357 (ATG/TAG)	393 (ATG/TAA)	357 (ATG/TAG)	357 (ATG/TAA)	378 (ATG/TAG)	360 (GTG/TAG)	369 (ATG/TAG)	438 (ATT/TAG)

cob	1152 (ATT/TAA)	1152 (ATT/TAA)	1143 (ATT/TAA)	1152 (ATC/TAA)	1162 (GTG/T**)	1140 (ATT/TAA)	1176 (ATG/TAG)	1182 (GTG/TAG)
nad5	1734 (ATG/TAA)	1734 (ATA/TAA)	1851 (ATG/TAA)	1731 (ATG/TAA)	1785 (TC/TAA)	1779 (ATG/TAG)	1794 (ATG/TAA)	1791 (ATG/TAA)
nad1	900 (ATT/TAA)	894 (ATT/TAA)	897 (ATT/TAA)	897 (ATT/TAA)	906 (ATT/TAG)	906 (ATT/TAA)	906 (ATC/TAA)	909 (ATC/TAA)
nad6	489 (ATA/TAA)	489 (ATC/TAG)	489 (ATA/TAG)	489 (ATT/TAG)	510 (ATG/TAA)	489 (TTG/TAG)	516 (ATG/TAA)	516 (ATG/TAA)
nad4	1350 (ATT/TAA)	1359 (GTG/TAG)	1350 (ATT/TAG)	1380 (GTG/TAG)	1455 (ATT/TAG)	1584 (ATT/TAA)	1419 (TTG/TAG)	1389 (ATT/TAA)
nad4L	348 (ATA/TAG)	297 (ATG/TAG)	396 (ATT/TAG)	309 (ATA/TAG)	393 (GTG/TAA)	294 (ATG/TAA)	315 (ATG/TAG)	297 (ATG/TAG)
atp8	198 (TTG/TAA)	198 (TTG/TAA)	192 (TTG/TAA)	195 (TTG/TAG)	168 (ATG/TAG)	183 (GTG/TAA)	183 (ATG/TAA)	171 (ATG/TAA)
atp6	708 (ATG/TAA)	711 (ATG/TAG)	720 (TTG/TAG)	708 (ATG/TAG)	693 (ATG/TAG)	696 (ATG/TAA)	690 (ATG/TAA)	690 (ATG/TAA)
cox3	810 (ATA/TAA)	795 (TTG/TAG)	780 (ATG/TAG)	780 (ATG/TAG)	774 (ATT/TAG)	825 (TTG/TAG)	774 (ATT/TAG)	774 (ATT/TAG)

Table 3. Substitution rates (μ) and K_a/K_s-ratios for female- and male-type mtDNA genes (gene fragments) for the Unionidae, and results of General Linear Models examining effects of type and node age on divergence rate. Abbreviations: PCG, protein-coding gene; RCC, respiratory chain complex

Type of DNA	Gene (fragment)	μ [substitutions/site/year x 10^{-9}] (mean (lower-upper 95% HPD))		K_a/K_s (mean ±SD)		General Linear Model statistics; response variable = *ln*(divergence rate)				
		Female-type	Male-type	Female-type	Male-type	P (type)	P (ln(node age))	R^2 (type)	R^2 (ln(node age))	R^2 (total model)
Whole mitogenome		2.17 (1.61-3.32)	2.96 (2.23-4.41)	0.09 ±0.02[1]	0.20 ±0.04[1]	<0.0001	<0.0001	0.25	0.35	0.78
PCG, RCC5	*atp6*	2.42 (1.81-3.68)	3.33 (2.52-4.93)	0.13 ±0.04	0.25 ±0.08	<0.0001	<0.0001	0.19	0.22	0.54
PCG, RCC5	*atp8*	6.77 (5.13-10.0)	7.71 (5.86-11.3)	0.39 ±0.19	0.51 ±0.19	n.s.	n.s.	0.04	0.01	0.07
PCG, RCC3	*cob*	2.51 (1.87-3.80)	3.14 (2.36-4.71)	0.06 ±0.02	0.12 ±0.03	<0.0001	<0.0001	0.10	0.14	0.33
PCG, RCC4	*cox1*	1.50 (1.11-2.30)	2.02 (1.52-3.03)	0.02 ±0.01	0.12 ±0.03	<0.0001	<0.0001	0.13	0.52	0.83
PCG, RCC4	*cox2*	1.90 (1.41-2.90)	3.58 (2.70-5.32)	0.09 ±0.03	0.23 ±0.07	<0.0001	<0.0001	0.44	0.14	0.73
PCG, RCC4	*cox3*	1.70 (1.26-2.60)	2.49 (1.87-3.74)	0.05 ±0.01	0.16 ±0.05	<0.0001	<0.0001	0.22	0.36	0.75
PCG, RCC1	*nad1*	2.37 (1.77-3.60)	3.27 (2.47-4.87)	0.06 ±0.02	0.14 ±0.05	<0.0001	0.0004	0.23	0.05	0.36
PCG, RCC1	*nad2*	3.09 (2.30-4.70)	4.41 (3.35-6.45)	0.13 ±0.05	0.28 ±0.07	<0.0001	<0.0001	0.34	0.08	0.54
PCG, RCC1	*nad3*	2.52 (1.86-3.90)	4.03 (3.05-5.91)	0.11 ±0.03	0.32 ±0.12	<0.0001	<0.0001	0.32	0.23	0.72
PCG, RCC1	*nad4*	2.80 (2.08-4.30)	3.86 (2.92-5.70)	0.14 ±0.04	0.25 ±0.07	<0.0001	<0.0001	0.28	0.22	0.66
PCG, RCC1	*nad4l*	2.24 (1.68-3.40)	3.69 (2.81-5.36)	0.17 ±0.10	0.34 ±0.11	<0.0001	n.s.	0.39	0	0.39
PCG, RCC1	*nad5*	2.84 (2.12-4.30)	3.95 (2.99-5.84)	0.11 ±0.03	0.25 ±0.07	<0.0001	<0.0001	0.26	0.10	0.46
PCG, RCC1	*nad6*	3.18 (2.41-4.70)	5.35 (4.07-7.85)	0.14 ±0.05	0.26 ±0.10	<0.0001	n.s.	0.42	0	0.42
Large rRNA	*rrnL*	1.49 (1.11-2.30)	1.76 (1.33-2.63)	n/a	n/a	<0.0001	<0.0001	0.08	0.23	0.41
Small rRNA	*rrnS*	1.14 (0.85-1.70)	1.42 (1.08-2.13)	n/a	n/a	<0.0001	<0.0001	0.16	0.12	0.38
PCG fragment	co1	1.60 (1.18-2.50)	2.15 (1.61-3.22)	n/a	n/a	<0.0001	<0.0001	0.12	0.52	0.79
PCG fragment	nd1	2.33 (1.74-3.50)	3.22 (2.43-4.79)	n/a	n/a	<0.0001	<0.0001	0.24	0.07	0.39
Large rRNA fragment	16S rRNA	0.83 (0.62-1.30)	1.19 (0.90-1.78)	n/a	n/a	<0.0001	0.0054	0.23	0.02	0.30

774 [1] PCGs only

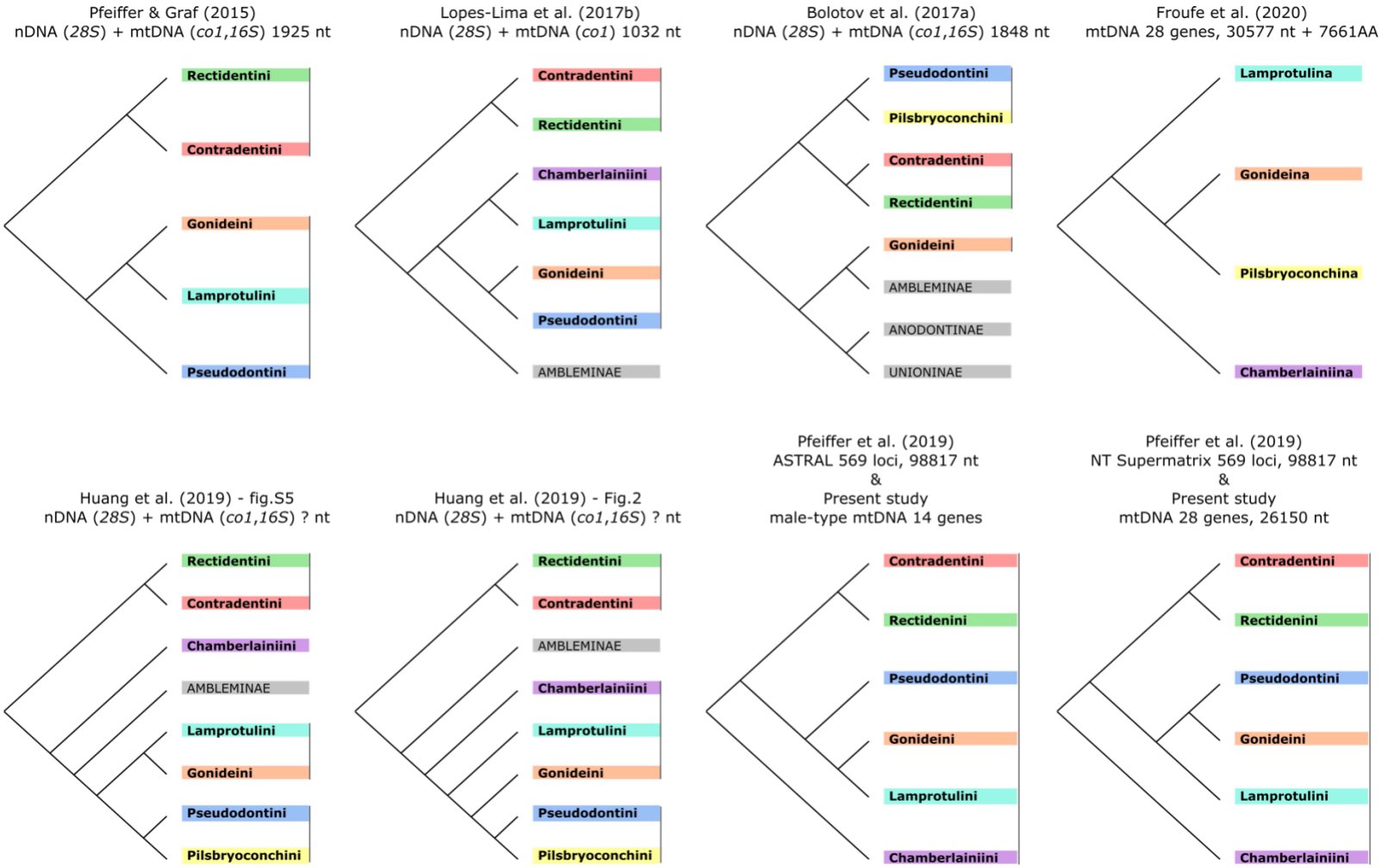


Figure 1. Recent multi-locus phylogenetic hypotheses on Gonideinae *sensu* Pfeiffer et al. (2019). Vertical bars indicate subfamilies recognised in respective publications. Note that Froufe et al. (2020) adopt a new systematic framework with three instead of two family-group levels, and thus, traditional tribes (ending -ini) are considered subtribes (ending -ina) in that study.

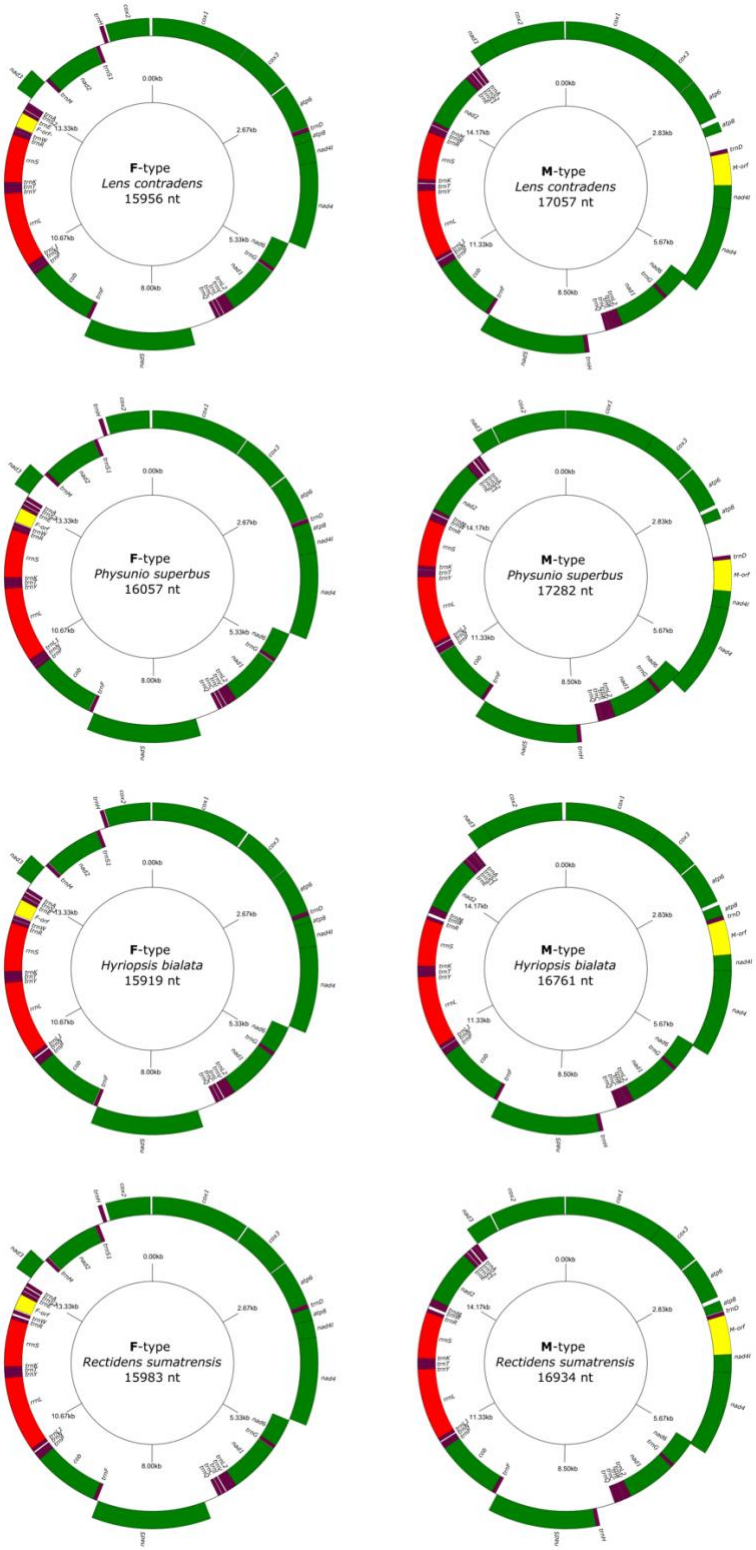
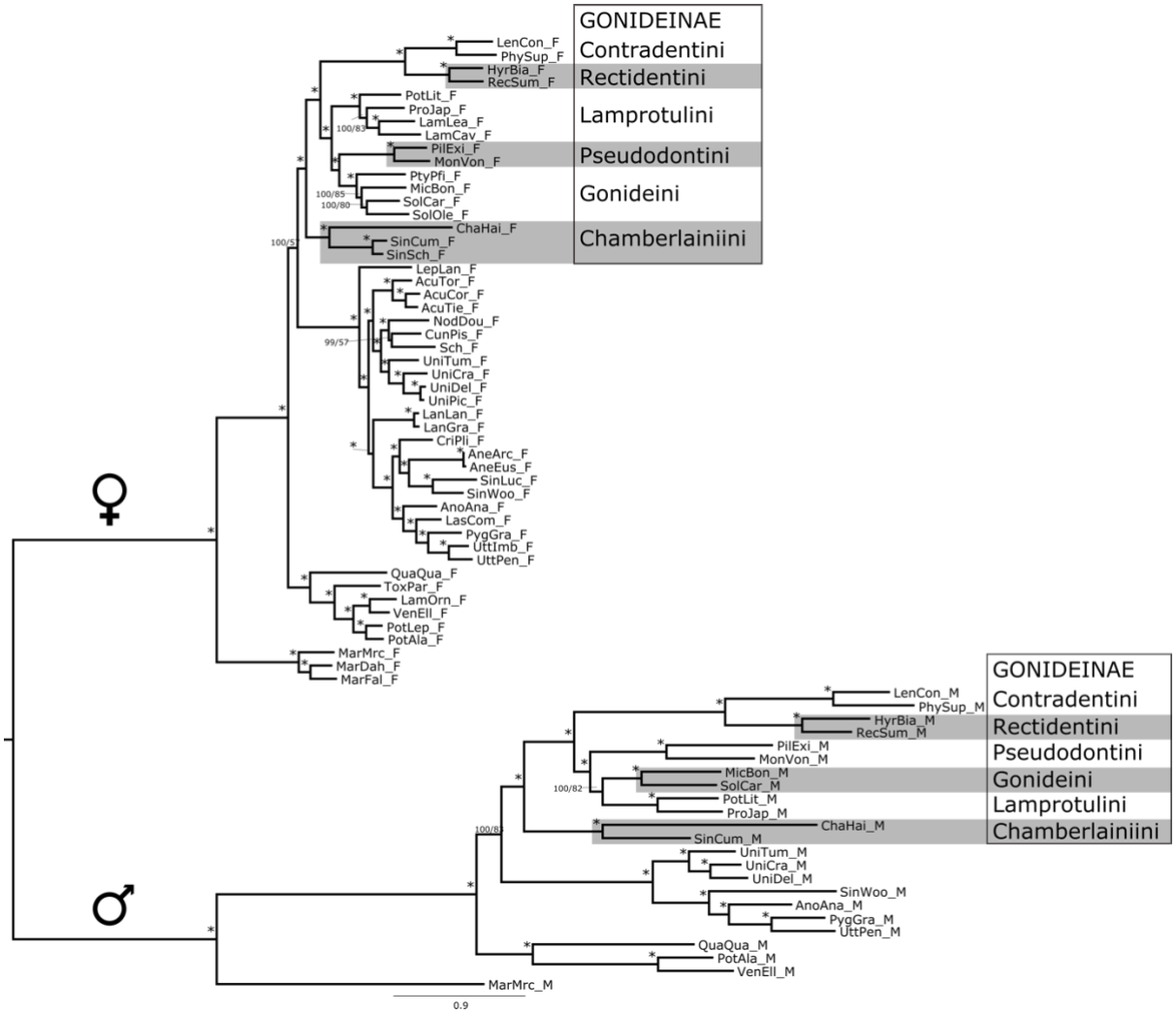


Figure 2. Gene maps of the F- and M-type mitochondrial genomes of *Lens contradens*, *Physunio superbus*, *Hyriopsis bialata* and *Rectidens sumatrensis*. Genes positioned inside the circle are encoded on the heavy strand, and genes outside the circle are encoded on the light strand. Colour codes: Small and large ribosomal RNAs (red); transfer RNAs (purple); *M-orf*, F-specific open reading frame (yellow); *M-orf*, M-specific open reading frame (yellow); protein-coding genes (green).



787

788 **Figure 3**. Phylogenetic tree of the Unionidae + Margaritiferidae estimated from 14
789 concatenated individual mtDNA gene sequences (12 protein-coding and 2 rRNA genes).
790 Values for branch support above each node represent Bayesian posterior probabilities
791 percentage/Maximum Likelihood bootstrap support

792

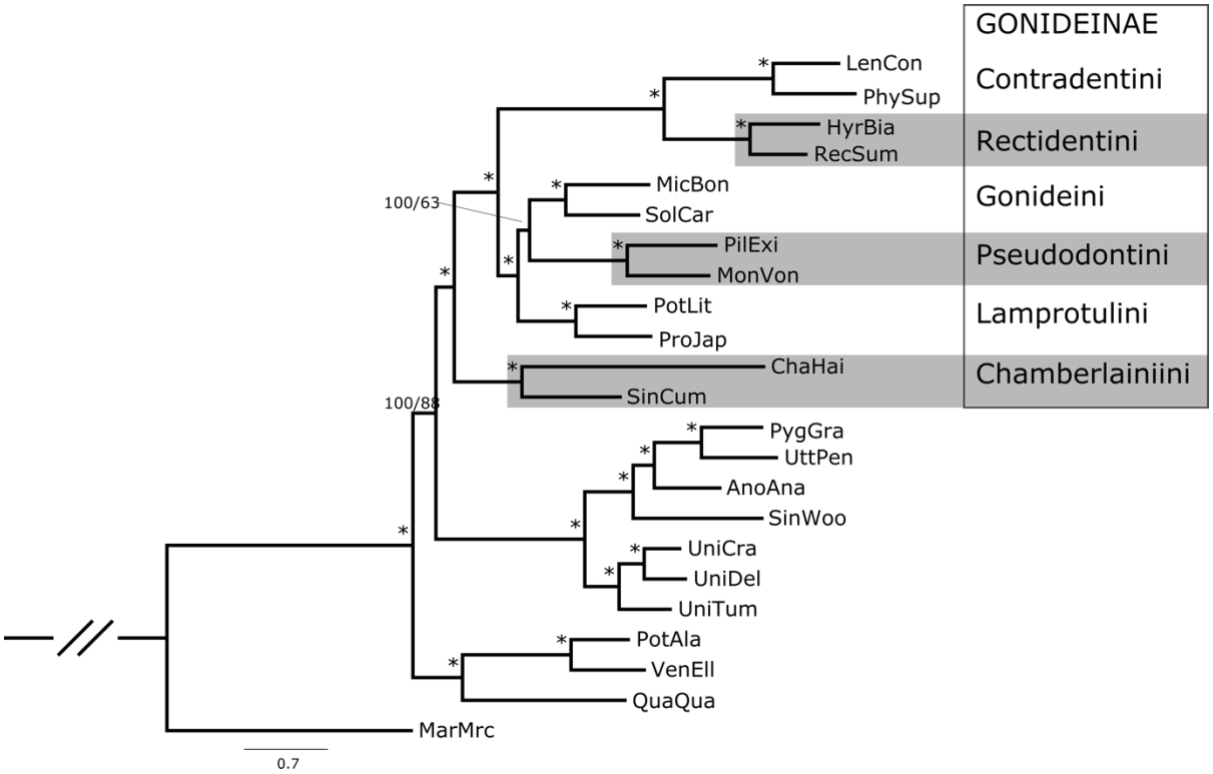


Figure 4. Phylogenetic tree of the Unionidae + Margaritiferidae estimated from 28 concatenated individual mtDNA gene sequences, i.e. 14 from female-type (12 protein-coding and 2 rRNA genes) and 14 (12 protein-coding and 2 rRNA genes) from male-type mitochondria. Values for branch support above each node represent Bayesian posterior probabilities percentage/Maximum Likelihood bootstrap

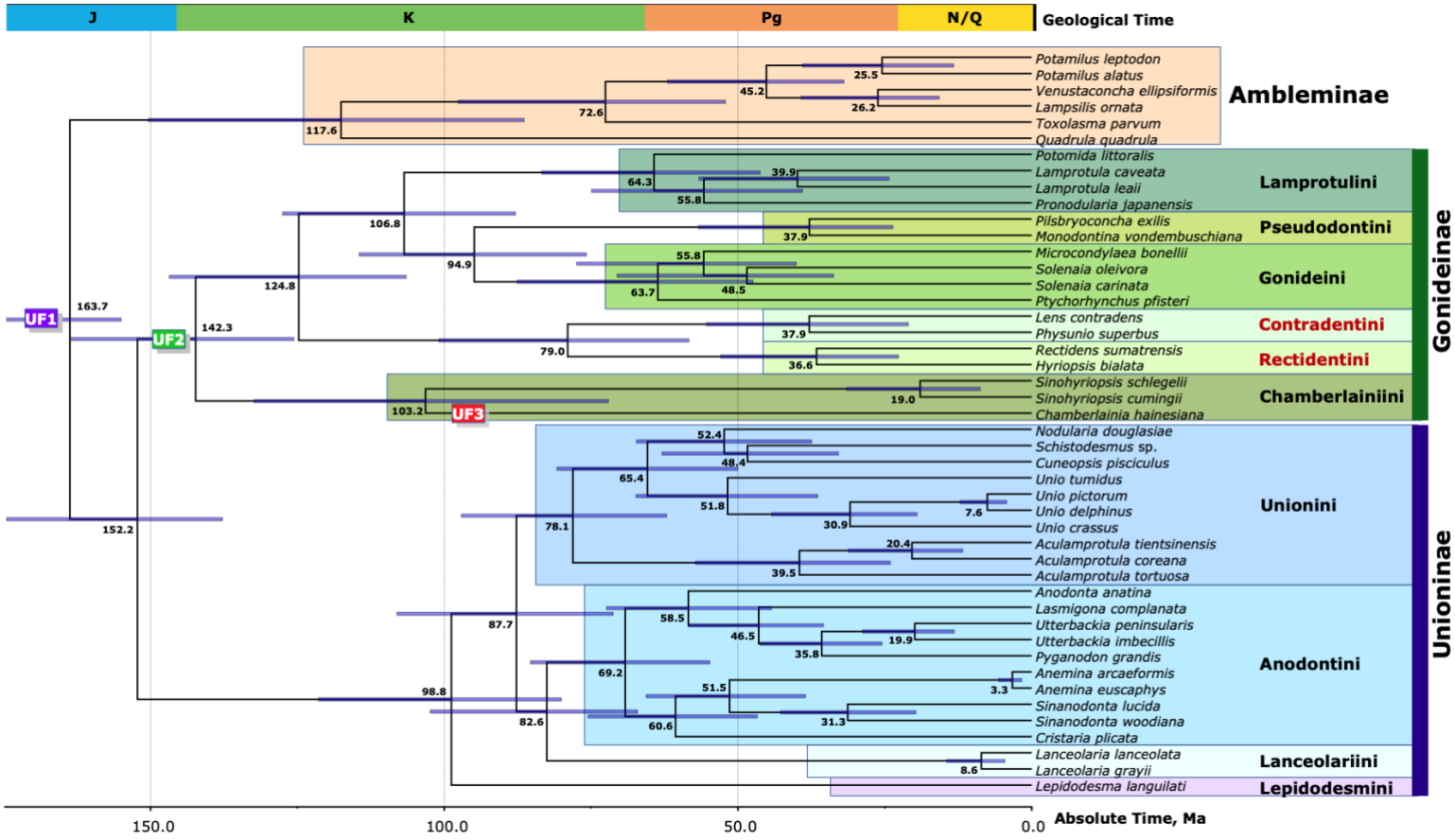


Figure 5. Fossil-calibrated ultrametric chronogram of the Unionidae calculated under a lognormal relaxed clock model and a Yule process speciation implemented in BEAST v. 1.10.1 and obtained for the complete F-type mitogenome data set. The newly sequenced tribe-level taxa are coloured red. An outgroup sample (Margaritiferidae) has been removed for better visualisation (but see original BEAST tree in Fig. S1A). Bars indicate 95% confidence intervals of the estimated divergence times between lineages (Mya). Black numbers near nodes are mean ages (Mya). Colour labels indicate the F-mtDNA gene order (UF1, UF2, and UF3). Stratigraphic chart according to the International Commission on Stratigraphy v. 2018/08 (www.stratigraphy.org).

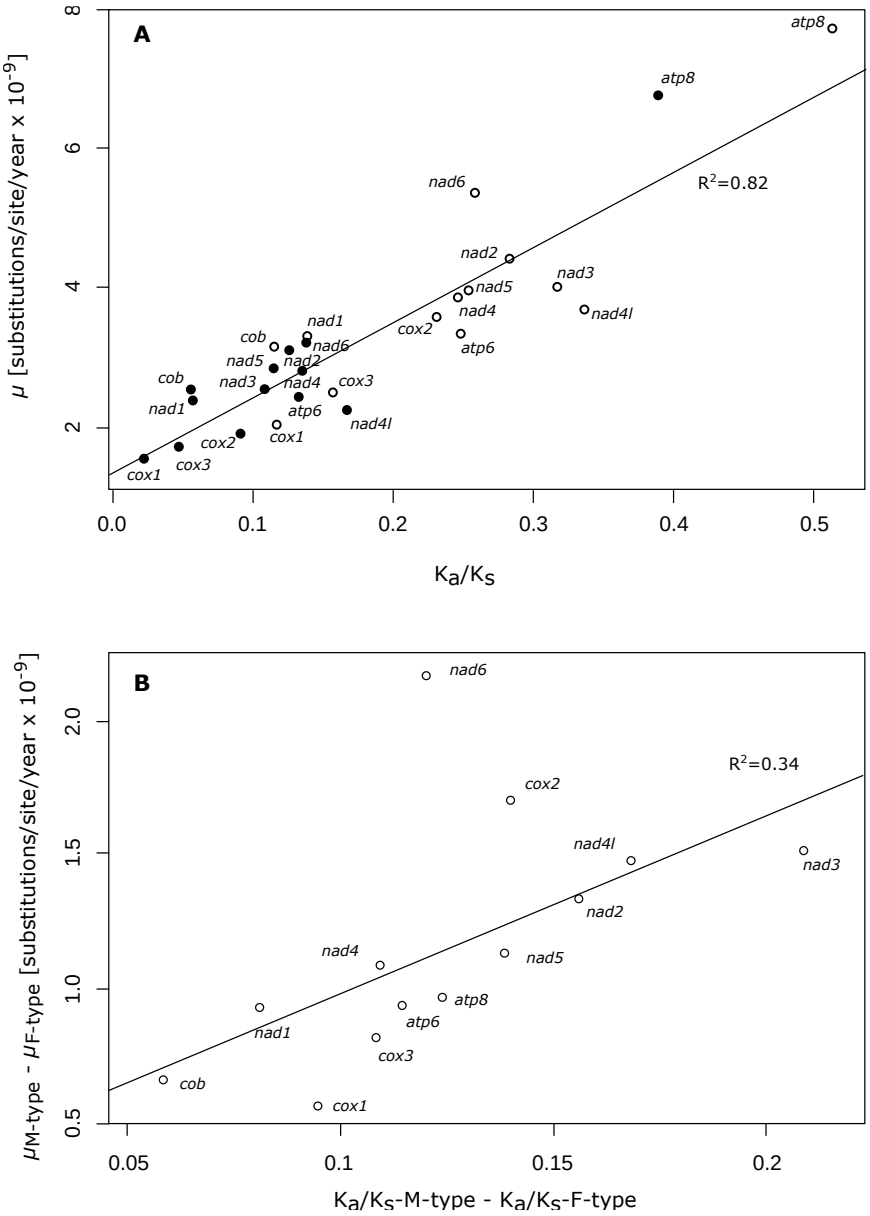


Figure 6. Relationship between (A) mean K_a/K_s and substitution rate (μ) per female-type (full circles) and male-type (empty circles) mtDNA protein-coding gene; and (B) differences between male- and female-type K_a/K_s and μ per protein-coding mtDNA gene.

Supporting information

Figure S1. Original BEAST v. 1.10.1 trees obtained for the (A) complete F-type and (B) M-type mitogenome data set, respectively.

Figure S2. Relationship of divergence rates and node age for each female- (F) and male-type (M) mtDNA gene, fragment and the whole mitogenome, calculated based on a fossil-calibrated mitogenomic phylogeny of the Unionidae. For each regression line, R^2, slope (m) and intercept (b) values are given.

Table S1. Best-fit models of nucleotide substitution for each partition subset based on Bayesian Information Criteria (BIC) using PartitionFinder2 (version 2.1.1 Lanfear et al., 2017) for the Bayesian Inference (BI) and Maximum Likelihood (ML) analyses.

Table S2. List of fossil calibrations that were used in BEAST analyses