MOLECULAR PHYLOGENY OF THE FRESHWATER MUSSEL FAMILY DREISSENIDAE

A Thesis
By
SUSAN ROSE GEDA

Submitted to the Graduate School
At Appalachian State University
In partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

May 2017
Department of Biology

.

MOLECULAR PHYLOGENY OF THE FRESHWATER MUSSEL FAMILY
DREISSENIDAE

A Thesis
By
SUSAN ROSE GEDA

APPROVED BY:

_____

Michael M. Gangloff, Ph.D.
Chairperson, Thesis Committee

_____

Matt C. Estep, Ph.D.
Member, Thesis Committee

_____

Lynn Siefferman, Ph.D.
Member, Thesis Committee

_____

Zack E. Murrell, Ph.D.
Chairperson, Department of Biology

_____

Max Poole, Ph.D.
Dean, Cratis D. Williams School of Graduate Studies

Abstract

MOLECULAR PHYLOGENY OF THE FRESHWATER BIVALVE FAMILY
DREISSENIDAE

Susan R. Geda
B.S., Appalachian State University
M.S., Appalachian State University

Chairperson: Dr. Michael Gangloff

The bivalve family Dreissenidae contains some of the most economically and

ecologically important fresh and brackish-water mollusk species, including a number

of problematic invaders. There has been much uncertainty surrounding phylogenetic

resolution for members of Dreissenidae. The lineage is believed to have originated

83.6 million years ago in the Tethys and Paratethys seas. Three extant dreissenid

genera are currently recognized, *Dreissena, Mytilopsis,* and *Congeria.* However, in

2012, an un-described mussel was discovered in the Xingu River, a tributary of the

Amazon River in central Brazil. Sympatrically-occurring *Congeria* species have also

been recently described from the region. The objective of this study was to

determine the evolutionary history of the unknown South American dreissenids

(USADs) and determine their proper taxonomic placement. I examined phylogenetic

relationships among 10 described species within Dreissenidae and 6 related

outgroups using nuclear and mitochondrial genes, a molecular clock analysis and a

comparative analysis of life history characteristics to determine the evolutionary

history of these enigmatic bivalves. Recent analyses suggest that these bivalves

may comprise a distinct genus within Dreissenidae containing three species. My analyses support this hypothesis and revealed that the common ancestors of today's USADs first diverged as a distinct lineage ~27.1 MYA. Due to phylogenetic analyses, genetic distance, and life history characteristics I believe USADs are sister taxa to *Congeria* and may have dispersed to South America on large scale ocean currents during the early Miocene.

Acknowledgements

First and foremost, I would like to thank my advisor, Dr. Michael Gangloff. This project would not have been possible without your leadership and advice. You offered me so many research opportunities that created an extremely fulfilling graduate career for me. You pushed me enough so that I realized my potential and I would not be here without your support! I would also like to thank the members of my committee, Dr. Matt Estep and Dr. Sarah Marshburn, for their guidance, ideas, and edits to this thesis. This project would still be in progress without your insight.

To all of the members of the ACRL lab, thank you for keeping me sane and making me laugh when I needed it the most. Daniel Mason, thank you for being my friend in the trenches and for being a sounding board for me to bounce my, sometimes terrible, ideas off of. Your perspective helped shape what this project is today. Michael Perkins and Raquel Fagundo, thank you both for being amazing resources for me during my analyses, for always answering the phone and talking me through the next steps, and making me laugh when I was almost ready to cry.

I would also like to thank all of my collaborators. Specifically, Mark Sabaj-Perez, I would not have been able to travel to Brazil without your generous invitation. Thiago Mahlmann and Rafaela Ota at the Instituto Nacional de Pequisas da

Table of Contents

List of Tables

List of Figures

Foreword

This research will be submitted to the peer-reviewed journal, *Molecular Phylogenetics and Evolution.* It has been formatted to fit the requirements for that journal.

1. Introduction

1.1 Evolutionary History of Dreissenidae

The freshwater bivalve family Dreissenidae originated in Neogene fresh and brackish-water lake systems ~83.6 million years ago (MYA) during the Triassic period, concurrent with the time that Pangea began to separate into the continents of North and South America, Europe, and Africa (Bilandžija et al., 2013). Dreissenidae is the only extant family of the superfamily Dreissenoidea and contains three extant and two fossil genera: *Prodreissensia* and *Dreisenomya* (Nuttall 1990; Steipen et al., 2013). *Mytilopsis* (Conrad, 1857) and *Congeria* (Partsch, 1835) are sister genera and their common ancestor diverged from *Dreissena* (Van Beneden, 1835) ~37.4 MYA in the late Eocene (Bilandžija et al., 2013). The Eurasian Tectonic Plate was very close to its modern position. The Neotethys Sea was still in existence, creating shallow seas where Mediterranean, Italian, and North African coasts exist today. *Congeria* diverged from *Mytilopsis* ~22.8 MYA in the early Miocene when the Adriatic and Mediterranean seas were beginning to form and coastlines were similar to those during the present day (Bilandžija et al., 2013). *Dreissena* split from its MRCA in the late Miocene, after the draining of the Tethys Sea.

The genus *Mytilopsis* is endemic to subtropical estuarine systems in North and Central America and includes at least four extant species: *M. leucophaeata* (Conrad, 1831), *M. sallei* (Récluz, 1849), *M. trautwineana* (Tryon, 1866), and *M. lopesi* (Alverenga and Ricci, 1989). Of these, *Mytilopsis leucophaeata* and *M. sallei* are known to be highly invasive, having colonized subtropical and tropical estuarine systems worldwide (Tan and Brian, 2006) including systems as far north as the

1

Hudson River (Therriault et al., 2004). Fossil evidence suggests that *Mytilopsis* became extinct in the Old World after it dispersed across the Atlantic Ocean (Van der Velde et al., 2010). The number of *Mytilopsis* species is currently disputed. *Mytilopsis lopesi* was recently described from the Toncantins River in the Amazon drainage (Alvarenga and Ricci, 1989; Morton, 1993). As well as *M. trautwineana,* which was recently described from the Pacific Coast of Columbia and Ecuador. However, both of these descriptions were based on morphology alone.

  Dreissena is endemic to the Ponto-Caspian region of Europe, but species have been introduced to freshwater ecosystems worldwide (Therriault et al., 2004). *Dreissena* is comprised of *D. polymorpha* (Pallas, 1771) and two subgenera including *Prontodreissena,* containing *D. rostriformis* (Deshayes, 1838; quagga mussel), and *Carinodreissena,* which includes, *D. prebensis* (Kobelt, 1915)*, D. blanci* (Westerlund, 1890) and *D. stankovici* (Lvova and Starobogatov, 1982). Carinodreissenids are found primarily in the northern and central Balkan Peninsula. *D. polymorpha* and *D. rostriformis* have invaded the Laurentian great lakes and numerous riverine systems in the central and eastern United States including the Mississippi and Hudson River systems (Therriault et al., 2004).

  Until recently, *Congeria* was believed to be restricted to troglobitic ecosystems in the Dinaric Karst on the coast of the Adriatic Sea in Europe (Morton et al., 1998). However, there has been a recent species description within the Amazon drainage. *Congeria* now consists of four species. Three of which, *C. kusceri* (Bole, 1962)*, C. mulaomerovici* (Morton and Bilandzija, 2013)*, and *C. jalzici* (Morton and Bilandzija, 2013)*, are extant in the Dinaric Karst (Bilandzija et al., 2013).

Congeria hoeblichi (Schütt, 1989) was described from the Caroni River in Venezuela (Pereira et al*.,* 2014). This taxon occurs sympatrically with USADs. The Caroni (a tributary to the Orinoco River) is connected with the Negro River (an Amazon River tributary) through the Casiquiare Canal. It is unclear whether *C. hoeblichi* belongs in the genus *Congeria* because its initial description was based on morphological traits alone. It is hypothesized that all other *Congeria* became extinct in the Messinian salinity crisis ~6 MYA (Morton et al*.,* 1998). This was an event in which the Mediterranean Sea's salinity levels underwent a sharp increase after becoming isolated from the Atlantic Ocean due to a tectonic event (Duggen et al., 2003).

1.2 Dreissenid Biology

Dreissenids are colonial, epifaunal, and heteromyarian (i.e., their adductor mussels are unequally developed) bivalves. *Dreissena* and *Mytilopsis* can live three to nine years, with a large variation in age among populations (Mardsen et al*.,* 1995). Both genera are r-selected and exhibit planktotrophic larvae. Most species are short-lived with high reproductive rates that have facilitated invasion of naive ecosystems. *Congeria* is unique among dreissenids as the only k-selected genus with species that brood their larvae internally. Small, isolated *Congeria* populations occur in deep cave system habitats and individuals appear to live for 30-40 years (Puljas et al*.,* 2014). *Congeria* aggregations exhibit low recruitment and low mortality. Interestingly, *Congeria* exhibits higher levels of genetic variability than *Dreissena* and *Mytilopsis* despite its restricted range and smaller population sizes (Steipen et al*.,* 2013).

Reproduction in USADs appears more similar to *Congeria* than to *Dreissena* or *Mytilopsis* (Mansur, Personal Communication). The USADs brood their embryos in the frontal portion of the pallial cavity, whereas *Congeria* brood them near the exhalant siphon. Little is known about the life history of these novel taxa from the Amazon River basin.

1.3 Study Sites

The modern course of the Amazon River and the surrounding tributaries were formed through successive geotectonic events. Gondwana broke apart in the Mesozoic, and as a result the main drainage of the South American plate was directed to the west. Four separate river basins were formed at this time (~87-63 MYA; Pereira et al.*,* 2014). The Cenozoic era brought about a very different landscape for central Brazil with the formation of the Paranean Sea. This was formed by a 150 m deep sea stand, referred to as the Pebas Sea. Approximately 23 MYA this estuarine sea drained to the west and to the north, emptying into the Caribbean. The headwaters of this system were close to the current course of the Xingu River (Wesselingh and Salo, 2006). Before this continental sea became the Pantanal, the extensive wetland ecosystem in southern Brazil, it likely sustained large, permanent lakes between 17 and 9 MYA. This is demonstrated by the high amounts of molluscan endemism in this region (Wesselingh and Salo, 2006). In the late Miocene, the uplifting of the Guyana Shield blocked the connection of the Pebas system to the North. This uplifting (~10-9 MYA) is most likely when the system began to slowly transfer to a freshwater fluvial environment. The dynamics of the

Andes along with marine incursions and regressions are most likely responsible for forming the modern Amazon River Basin (Pereira et al*.,* 2014). The distribution of bivalves is closely related with the hydrogeological history of the continent (Pereira et al*.,* 2014). Through isolation and subsequent dispersal, speciation and adaptation allowed the bivalves of the region to diversify to the numbers known today.

There are 168 native and 5 invasive freshwater bivalve species recorded for the 52 hydrographic regions in South America, with the highest species diversity in the Brazilian Amazon (Pereira et al*.,* 2014). Brazil also has the highest species richness with 117 extant freshwater bivalve species. Mytiloida, Unionoida, and Veneroida are the three lineages present in South American freshwaters. Veneroida includes Corbiculidae, Sphaeriidae, and previously Dreissenidae (Pereira et al*.,* 2014). However, Dreissenidae was recently placed into the order Myida (Giribet and Distel, 2003). Three dreissenid taxa are known from South America; *Mytilopsis lopesi, M. trautwineana* and *Congeria hoeblichi.* The South American bivalve fauna is dominated by Unioniforme taxa (64.2% of species) but Corbiculidae (8.1%) and Dreissenidae (1.7%) are widespread. To date, phylogenetic associations, species boundaries and life history traits remain under-studied in South American drainages (Pereira et al*.,* 2014).

The purpose of this study was twofold. First I used a multi-gene dataset to elucidate the phylogenetic placement and evolutionary history of the South American dreissenid group within Dreissenidae. Second, I investigated congruence of morphological and molecular characters to determine whether there is sufficient

5

evidence to classify the USADs as a distinct genus. I hypothesize that USADs comprise a genetically distinct, monophyletic genus within Dreissenidae.

2. Methods

2.1 Tissue Collection and DNA Extraction

Tissue samples for USADs were collected between 2005 and 2016 from multiple populations within the Orinoco and Amazon basins. DNA was obtained from a single specimen among dozens collected in 2005 from the Ventuari River (Orinoco Drainage) in Venezuela. Specimens from the Ventuari River were collected, dried and transferred to 95% ethanol several years later. Xingu River specimens were collected from 5 sites in 2012 (n=26 individuals) and from 1 previously sampled site and 5 new sites in 2016 (n=28 individuals, Figure 1; Table 1). All Xingu River samples were stored in 95% ethanol at room temperature; samples collected in 2016 were stored in 70% ethanol at room temperature until samples could be transported to Appalachian State University and transferred to 95% ethanol.

DNA was extracted from adductor muscles using a MoBio Cell and Tissue DNA Extraction Kit (Carlsbad, CA) following manufacturer protocols. DNA yields were quantified using a NanoDrop 2000 UV-Vis Spectrophotometer. Xingu species 1 samples that yielded 40-160 ug/ul were used for PCR amplification. *Congeria hoeblichi* (Xingu species 2; Figure 2) samples that yielded 20-80 µg/ul were used for PCR amplification.

2.2 PCR Amplification

Regions of the mitochondrial COI subunit b and 16S mitochondrial RNA genes, as well as the nuclear 18S ribosomal RNA and 28S ribosomal RNA genes were amplified to elucidate phylogenetic relationships between these and other Dreissenidae taxa. Primers were adapted from Therriault et al.*,* 2004 and Frischer et al.*,* 2002 (Table 2). PCR amplifications for mtDNA were carried out under the following conditions: 10 µL GoTaq® Green Master Mix 2X per sample (manufacturer concentration; Promega Corporation, Madison, WI), 1 µL each primer per sample (0.5 µM final concentration), 1 µL 20-160 ng/µL DNA template per sample, and nuclease-free water to a final volume of 20 µL per sample.

Reactions were run on an Eppendorf Mastercycler Nexus thermal cycler (Hamburg, Germany) using a touchdown protocol. The initial denaturation cycle started at 94°C (5 min) followed by 24 cycles consisting of a denaturation period at 94°C (45 s), an annealing cycle that started at 68°C and decreased 1°C per cycle (2 min), and an extension cycle at 72°C (60 s). Additionally, 25 cycles were performed with a denaturation cycle at 94°C (45 s), an annealing cycle at the appropriate annealing temperature depending on the gene (COI: 50°C; 28S: 46.5°C; 16S: 48.5°C; 18S: 55°C) (60 s), and an extension cycle at 72°C (60 s). A final extension cycle at 72°C (10 min) was performed. Reactions were held at 10°C until product could be removed from the thermal cycler.

PCR products were run on 1.5% agarose gels using standard gel electrophoresis in standard 1x TBE buffer. Gels were run for 1-1.5 hours at 100 V and then visualized using UV-transillumination of ethidium bromide stained gels. For

28S and 16S there were several samples that yielded multiple products for one individual. In this case, the appropriate band was extracted from a 1.5% gel, left overnight in 1X TE buffer, and then used as DNA template in a secondary PCR reaction with the same parameters as described above. When this secondary reaction did not yield PCR product, a DNA precipitation was performed on the excised gel band and 1XTE buffer solution. 1/10th the total volume of 3M sodium acetate (pH 5.2) was added to the solution, along with 2X total volume of 100% molecular grade ethanol. The samples were then placed in a -80°C freezer for 1 hour. Samples were decanted, to leave the gel band behind, and centrifuged for 30 minutes at 14,000 RPM. The supernatant was removed, and the pellets were washed with 70% ethanol and centrifuged for an additional 5 minutes. The samples were then air dried for approximately 10-15 minutes and then rehydrated with 1X TE buffer. This solution was then used for a tertiary PCR reaction, using the parameters described above. These products were then confirmed on a gel and sent off with primary products to Retrogen, Inc. for sequencing.

2.3 Sequence Analyses

Sequences were edited, aligned, and concatenated using Geneious R7 (Biomatters Ltd., Auckland, New Zealand). All sequences were aligned using the ClustalW algorithm and manually trimmed and inspected. Sequences for the COI alignment were checked for quality by determining the presence of stop codons through translation into amino acids. Mitochondrially-derived nuclear DNA fragments (numts) and male mitotypes were identified by uncommon divergence from other sequences

8

and discarded. Sequences with HQ scores below 50% were discarded. HQ% scores are a measure of quality adapted from phred scores. The final concatenated mitochondrial DNA alignment was 486 bp of COI fragments and 402 bp of 16S fragments to comprise a total length of 888 bp. The final concatenated nuclear DNA alignment included 1,575 bp of 18S and 621 bp of 28S to comprise a total length of 2,196 bp. The concatenated alignment composed of both nuclear and mitochondrial datasets was 3,057 bp in length. There are 94 variable sites in the COI dataset, 276 variable sites in the 28S dataset, 206 variable sites in the 16S dataset, and 322 variable sites in the 18S dataset. The best-fit nucleotide substitution model was determined with jModelTest version 2 (Darriba et al., 2012). Separate iterations were run in jModelTest for each alignment.

Genetic divergence rates (uncorrected p-distance) among Dreissenidae taxa and appropriate outgroups were analyzed with MEGA version 6 using maximum composite likelihood (Tamura et al., 2013). The number of haplotypes was calculated using DNAsp v5.10.1 (Librado and Rozas, 2009). Haplotype maps were created using PopART (Clement et al., 2002). The haplotype analysis was performed with the singleton haplotypes omitted; the resulting network therefore does not include all sites or individuals sampled during this study.

2.4 Phylogenetic Analyses

Bayesian inference by Markov Chain Monte Carlo (MCMC) analyses were conducted on each single-gene dataset, the mitochondrial dataset, the nuclear dataset and a concatenated alignment consisting of all 4 genes. Each phylogeny

had 1×10^6 iterations run with 3 heated chains. Subsampling occurred every 10,000 generations. There was a burn-in length of 1×10^5 trees. These analyses were completed using the MrBayes 3.2.2 plug-in (Huelsenback and Ronquist, 2001) within Geneious. One representative per haplotype was used during analyses. Each gene's alignment was assessed in DNAsp separately to ensure that duplicate genotypes were not used in phylogenetic analyses. Maximum likelihood analyses were constructed within Geneious using the PhyML plug-in (Guindon and Gascuel, 2003). These trees were constructed with 1,000 iterations for each alignment.

I obtained outgroup sequences from Genbank for all sampled dreissenids as well as from closely-related orders within the subclass Heterodonta (Table 3). *Corbicula fluminea* was used to root all phylogenetic trees. Unfortunately, there were no sequences available for *Mytilopsis trautwineana* or *Mytilopsis lopesi*. Other than these exceptions I had data from all extant Dreissenidae.

Mapping life history traits onto a phylogenetic tree allowed visualization and interpretation of how life history traits are distributed among members of Dreissenidae. Character traits were simply mapped onto the phylogeny using Adobe Illustrator CS6 (Adobe Systems, Inc.).

2.5 Divergence Estimates

The concatenated dataset was used to estimate divergence time for the USADs utilizing BEAST version 1.8.4 (Drummond et al., 2012). An UPGMA starting tree was constructed using the GTR+I+G model and estimated base frequencies were used for this analysis. I used a constant-size coalescent model and an uncorrelated

10

relaxed clock with a lognormal distribution. The clock was calibrated from fossil evidence of the first appearance of *Dreissena* ~11.6 MYA. Nodes for *Congeria* and *Mytilopsis* were also calibrated with a lognormal prior of 5.4 and 22.8 MYA, respectively. These figures were adapted from Bizlandzija et al. (2013). The analysis was run for 1×10^7 iterations sampling every 1000 generations. Burn-in was assessed with Tracer version 1.6.

3. Results

3.1 Phylogenetic Analyses

Bayesian and maximum likelihood analyses at each locus, for the mitochondrial dataset, the nuclear dataset, and the concatenated dataset consistently revealed that USADs form a distinct monophyletic genus within Dreissenidae (Figures 3-16). Support values for a monophyletic USAD were consistently >80% in all phylogenies except the Bayesian analyses of the 28S dataset (Figure 6), the 16S dataset (Figure 8), and the maximum likelihood analysis of the 18S dataset (Figure 9). Support values for these nodes were 64, 63.8, and 74%, respectively. *Congeria hoeblichi* sequences also consistently formed monophyletic clades within the USAD clade (Figures 3-16). In most analyses, there were two USAD species, one of which corresponded to *C. hoeblichi*. The only exceptions were within the 28S analyses (Figures 5 & 6) in which USADs formed a polytomy.

Using jModelTest version 2 (Darriba et al*.,* 2012) I determined that the best model of COI, 28S, and concatenated alignment evolution fit a general time

reversible substitution model (GTR, Tavaré et al., 1997) based on the Akaike

Information Criterion (95% CI; gamma shape = 0.4530 [COI], 0.9770 [28S], 0.7190

[Concatenated]). However, jModelTest determined that the 16S and 18S datasets

best fit the Tamura Nei nucleotide substitution model (gamma shape =0.3590 [16S],

0.6530 [18S]).

Log likelihood estimates for all reconstructions performed in this study were

lowest for maximum likelihood analyses, except for the 18S dataset. Estimates were

highest for phylogenies that support *Congeria* as a sister genus to USADs (Table 4).

Of the 7 most likely phylogenies, 5 showed that this group is sister to *Congeria.* The

phylogenies with the highest log likelihood estimates are as follows: The COI

maximum likelihood analysis (Figure 3), 28S maximum likelihood analysis (Figure 5),

16S maximum likelihood gene tree (Figure 7), 18S Bayesian gene tree (Figure 10),

mitochondrial maximum likelihood analysis (Figure 11), nuclear dataset maximum

likelihood tree (Figure 13), and the 4 gene concatenated maximum likelihood tree

(Figure 15).

The concatenated phylogenies included all 4 genes (Figures 15 & 16). Within

the maximum likelihood phylogeny Dreissenidae is monophyletic with a bootstrap

value of 100%. *Dreissena* is the basal genus in this topology with a bootstrap value

of 100%. The clade linking *Mytilopsis* with the USADs + *Congeria* clade had a

bootstrap value of 93%. The node linking USADs and *Congeria* had a bootstrap

value of 71.5%. Bayesian and maximum likelihood topologies are largely congruent

except that USADs are sister to *Mytilopsis* instead of *Congeria* in likelihood

topologies*.* In the Bayesian phylogeny the node linking Mytilopsis + *Congeria* +

USADs to *Dreissena* had 100% posterior probability. *Mytilopsis* and USADs diverge

from *Congeria* with 84.2% posterior probability. The USADs and *Mytilopsis* then split

with 61.8% posterior probability.

3.2 Genetic Diversity of USADs

A total of 49 sequences were generated in this study, however only 28 were utilized

to avoid the use of duplicate haplotypes. The COI dataset generated in this study

totals 14 sequences. I found a total of 30 COI haplotypes in USAD taxa, including 16

haplotypes from a previous study (Table 1; Gangloff et al., unpublished data). There

were 20 haplotypes found within Xingu species 1 and 10 within Xingu species 2 (=

C. hoeblichi). Only cytochrome oxidase I (COI) was amplified for the Ventuari

specimen due to a lack of quality DNA, assessed with a Nanodrop 2000. DNAsp

revealed several genotypes for each of the analyzed loci. There were two 28S

genotypes for each USAD species; three 16S genotypes for species 1 (sp. 1) and

two for species 2 (sp. 2); three 18S genotypes for sp.1 and two for sp. 2. The

mitochondrial dataset had 6 haplotypes, four from sp. 1 and two from sp. 2. The

nuclear dataset had five genotypes, four from sp. 1 and one from sp 2. The

concatenated dataset had 2 USAD sequences, one from each species (Table 5).

The haplotype map (Figure 18) revealed that there are two major haplogroups

within the USAD COI dataset, representing two USAD species. Xingu sp.1 and 2 are

highly divergent. Xingu sp. 1 appears restricted to the Xingu River downstream of

the Volta Grande whereas sp. 2 is found upstream. Site 1 had the most diversity,

with 7 haplotypes. This is surprising, as Site 1 was closely located to Sites 2, 3 and

11. Sites 2, 5, 10, and 11 each had 2 haplotypes present. Sites 3 and 4 had only one haplotype present and site 1 shared 4 haplotypes with 4 other sites.

At the COI locus, inter-specific genetic distances for Dreissenidae and USAD species range from 12.8 to 19.1%. USAD taxa are 13.6 to 17.0% different from *Congeria* species, 15.7 to 19.1% different from *Dreissena* species, and 12.8 to 14.8% different from *Mytilopsis* species. Intra-specific distances between Xingu sp.1, Xingu sp. 2 and the Ventuari sp. range from 7.6 to 8.3% (Table 6). At the 28S locus, inter-specific genetic distances for Dreissenidae and USAD species range from 3.5 to 10.1%. USAD species are 3.5 to 3.9% different from *Congeria* species, 7.8 to 10.1% different from *Dreissena* species, and 5.3 to 6.5% different from *Mytilopsis* species. The intra-specific distance between Xingu sp. 1 and 2 is 0.6% (Table 7). At the 16S locus, inter-specific genetic distances for Dreissenidae and Xingu species range from 3.8 to 11.6%. USAD species are 3.8 to 5.5% different from *Congeria* species, 6.6 to 11.6% different from *Dreissena* species, and 5.6 to 8.0% different from *Mytilopsis* species. Intra-specific distance between Xingu sp. 1 and 2 is 1.7% (Table 8). At the 18S locus, inter-specific genetic distances for Dreissenidae and USADs range from 0.5 to 2.1%. The USADs are 1.0 to 1.1% different from *Congeria* species, 0.9 to 2.1% different from *Dreissena* species, and 0.5% different from *Mytilopsis* species. The intra-specific distance between Xingu sp. 1 and 2 at the 18S locus is 0.1% (Table 9).

3.3 Divergence Estimates

Molecular clock estimates revealed that the USADs split from *Congeria* in the Oligocene ~27.1 MYA (95% CI: 17.1 – 38.3 MYA). The two genera then diverged from one another ~5.1 MYA (95% CI: 1.9 – 9.4 MYA) in the late Miocene (Figure 19). Other divergence events within Dreissenidae are similar to Bizlandzija et al.'s (2013) divergence chronogram. The posterior probabilities for each node were all 100% except for the *Dreissena* + *Mytilopsis/Congeria/*USAD node (97%), the *Mytilopsis* + *Congeria/*USADs node (64%), the *Congeria* + USADs node (95%), and the *C. mulaomerovici* + *C. jalzici* node (99%).

4. Discussion

4.1 Phylogenetic Analyses

My analyses provide strong support that Dreissenidae is a monophyletic group comprised of four deeply-divergent clades and provide further support for the inclusion of a previously unknown South American bivalve lineage within this family. The multi-gene phylogenies generated by my analysis are largely congruent with those generated by prior molecular analyses that did not include data from USAD taxa (Therriault et al., 2004; Bilandzija et al., 2013). Results of all phylogenetic analyses provide compelling evidence that the taxonomy of Dreissenidae may be in need of substantial revision. At least one new genus is needed to accurately depict the unique evolutionary lineage represented by USADs within Dreissenidae.

Surprisingly, many of phylogenetic analyses suggest that *Congeria* and USADs are sister taxa. Although this seems counter-intuitive, *Congeria* (or a presumed ancestral form) was once more widespread and fossil material is known from shallow-water marine deposits in what is now Africa (Senckenberg: Collection Mollusca SMF). However, the USADs became isolated in South American interior rivers by vicariance events long before modern *Congeria* colonized the caves of the Dinaric Karst. This is also supported by the shared life history characteristics between *Congeria* and USADs (i.e., k-selection and larval brooding).

Analyses revealed some notable inconsistencies in nodal support for the USAD clade. The concatenated (4 gene) maximum likelihood and Bayesian analyses found relatively low nodal support for a monophyletic USAD (71.5% and 61.8% respectively) and this trend appears to be a characteristic of all analyses that included mitochondrial data. However, the USAD clade was well supported in nuclear datasets, except for the 28S Bayesian inference analysis (BPP = 64.5%). Other analyses of nuclear genes produced higher nodal support values (28S ML = 85.6%, 18S ML = 85.4%, 18S BPP = 74.3%) and the nuclear concatenated dataset revealed 91% and 98% BPP support for the USADs respectively.

My analyses suggest that divergence of the USADs was relatively ancient as reflected by phylogenetic reconstructions using more highly conserved nuclear loci. Explanations for low support values returned in mitochondrial phylogenies may include the phenomenon of superimposed substitutions (Springer et al., 2001). Incomplete lineage sorting and saturation are also explanations for these incongruent topologies (See Section 4.3).

4.2 Genetic Diversity of USADs

The haplotype map (Figure 18) shows the genetic structure of Xingu sp. 1 and 2.

They are highly divergent from one another, with ~300 km separating upstream

(Xingu sp. 2) and downstream (Xingu sp. 1) sites. Site 1 had the most diversity with

7 haplotypes present. Again, this is surprising due to the close location of sites 2, 3,

11. There is a high amount of gene flow occurring between haplotypes in Xingu sp.

1. However, the strange pattern of haplotypes discovered in this region may be

contributed to sample sizes from each site.

Examination of the COI and 18S loci suggest that the USADs are most

closely related to *Mytilopsis*, whereas analysis of the 28S and 16S loci suggest that

USADs are sister to *Congeria*. The closest intergeneric distances within the COI

dataset (Table 6) are as follows: the Ventuari specimen is 13.0% different than *M.*

sallei, Xingu sp. 1 is 12.8% divergent from *M. leucophaeata*, and Xingu sp. 2 is

13.9% different from *M. sallei* and *C. jalzici.* This affords doubt regarding the sister

taxa to USADs (considering this data alone) but due to the extent of these distances,

this is strong evidence that this clade is a distinct genera in the family Dreissenidae.

It is also suggested that *Congeria* is the sister genera to USADs when considering

phylogenetic analyses and life history characteristics. Interspecific distances

observed for dreissenids at the COI locus are typically ~12% (Molloy et al., 2011).

Intraspecific p-distances reported prior to this study are 0.27 – 3.9% (Wong et al.,

2011), providing evidence that there are two species within this distinct clade.

The 28S nuclear locus (Table 7) shows a different pattern. The closest

relative to Xingu sp. 1 and 2 is *Congeria kusceri* with a distance of 3.7 and 3.5%

respectively. Interspecific distances previously reported for dreissenids are 2.1 –
12.0%, while intraspecific distances are typically 0 – 2.0% (Molloy et al., 2011).
Phylogenetic reconstructions at this locus corroborate my hypothesis that the
USADs are a distinct clade that is sister to, yet divergent from, *Congeria.*

The 16S mitochondrial locus (Table 8) reports that the closest taxa to the
Xingu group are *Congeria.* Xingu sp. 1 and 2 are most closely related to *C. jalzici*,
with distances of 4.9 and 3.8% respectively. P distances for Dreissenidae at the 16S
locus are typically 1.0-8.0% between species and 0-0.9% among species (Molloy et
al., 2011).

The 18S locus is a highly conserved nuclear region, which is why the p
distances observed in Table 9 are small, even for the distantly related outgroups.
Xingu sp. 1 and 2 are most closely related to *M. leucophaeata* and *M. sallei* with
distances of 0.5%. They are 0.1% different from each other. There are no data
available comparing genetic distances among dreissenids at this locus. Based on
comparisons with more distant outgroup data as well as comparisons between
species within established genera, my data tend to support my initial hypothesis that
USADs form a distinct clade within Dreissenidae including 3 divergent species.

4.3 Incongruent Topologies

I obtained incongruent topologies among different analyses and datasets. These
results highlight the challenges of using multiple loci and analyses to recreate
evolutionary relationships. Most of my phylogenies depict the sister relationship
between *Mytilopsis* and *Congeria*, with USADs grouping with one or the other

18

depending upon the analysis performed. The USADs are sister to *Mytilopsis* in 8

(57.1%) out of the 14 phylogenies generated in this study, and sister to *Congeria* in

6 (42.9%). Mitochondrial datasets have incongruent topologies most consistently,

while nuclear datasets (with the exception of 28S) are more congruent. See Table

10 for the phylogenetic hypotheses proposed and supported within these

phylogenies.

jModelTest indicated that for 16S, 18S, and the nuclear (28S + 18S) datasets

the Tamura-Nei 93 substitution model was most appropriate. All other alignments

conformed to the General Time Reversible model. For maximum likelihood analyses,

all model parameters were applied. However, all Bayesian phylogenies were

reconstructed using the GTR model of nucleotide substitution, as the Tamura-Nei 93

substitution model could not be applied within the Geneious MrBayes 3.2.2 Plug-in.

Regardless, GTR was among the top 3 best fit models for each alignment. The GTR

substitution model allows for unequal base frequencies and different substitution

rates for all six pairs of possible substitutions, whereas the Tamura-Nei 93 model

also allows for unequal base frequencies, but equal transversion rates while allowing

transition rates to be variable (Tamura and Nei, 1993; Tavaré et al., 1997).

All of the alignments that best fit the Tamura-Nei model suggest that

Mytilopsis is sister to USADs. There is one exception, the 16S maximum likelihood

analysis suggested *Congeria* as the sister genus. For all analyses that used the

GTR model, all support *Congeria* within maximum likelihood analyses and *Mytilopsis*

within Bayesian inference, except the mitochondrial dataset which suggests

Congeria is the sister genus to USADs in both analyses.

Although maximum likelihood and Bayesian inference both utilize likelihood functions and an implicit model of evolution, the two methods of analysis differ in the fact that Bayesian analysis includes prior data, along with the current data, in the testing of hypotheses and process of phylogeny estimation (Archibald et al., 2003). Unfortunately, implementing realistic priors for Bayesian analysis is not yet possible. Priors such as which taxa to use, which are outgroups, or what data to use are, of course, prior knowledge, but this differs from a true prior of the probability of the distribution of trees. A simple prior, most often used, is assigning all trees equal prior probabilities. Valid priors assist in delineating the most probable, or true, phylogeny, but invalid priors could lead to inaccurate topologies (Archibald et al., 2003). Posterior probabilities describe the probability of topology considering the priors, the model selected, and the given data. All trees are then summarized in a majority rule consensus tree, which is used to determine posterior probability support. However, when considering the differences between maximum likelihood and Bayesian analysis, it is important to understand that Bayesian posterior probabilities and non-parametric maximum likelihood bootstrap values are inherently different. Posterior probabilities are often higher than maximum likelihood bootstrap values (Douady et al., 2003). Bootstrap values are more likely to fail to support a true node (type I error) and posterior probabilities are more likely to fail to reject a false node (type II error).

Even though Bayesian analyses typically have higher posterior probabilities than maximum likelihood bootstrap values, the values are still higher in the maximum likelihood analysis (even though these values are typically not comparable; Figures 15 and 16). The inability to properly implement realistic priors

into Bayesian analyses may bias topologies towards the support of an untrue topology. Maximum likelihood analyses are less likely to falsely support a false phylogenetic hypothesis.

Nuclear markers gained popularity in the community of phylogenetic reconstruction due to the heavy use of uniparentally inherited sequences (mitochondrial markers). By incorporating nuclear markers systematists hope to elucidate a clearer picture of evolutionary relationships of their target taxon. However, there are many disadvantages to using nuclear ribosomal RNA, such as multiple rDNA arrays, lack of complete concerted evolution, secondary structure and compensatory base changes, alignment accuracy, and homoplasy (Alvarez and Wendel, 2003).

The two mitochondrial genes utilized in this study (COI and 16S) do not have multiple rDNA arrays. COI is a protein-coding mitochondrial gene, and 16S is a small subunit ribosomal RNA. By inspecting genomes of mollusks closely related to Dreissenidae, it was determined that 16S has only one copy in the genome. However, 28S and 18S, the nuclear derived ribosomal RNA genes used in this study, may have multiple rDNA arrays. Multiple rDNA arrays can be created through polyploidization events and any event that leads to gene duplication (Mishima et al., 2002). When paralogous genes are used in phylogenetic studies, there are atypical long branch lengths in maximum likelihood analyses (Mayol and Rossello, 2001), I did not experience any of these phenomena during analyses. Dreissenidae are diploid organisms, perhaps lending evidence to the orthology of the genes

investigated within this study. However, if there were paralogous genes amplified, it would explain the incongruence among topologies.

Concerted evolution is a phenomenon where multiple copies of a gene appear to evolve in unison. This is due to unequal crossing over and high frequency gene conversion (Alvarez and Wendel, 2003). In theory, this would eliminate the problems that paralogous sequences pose. Unfortunately, when concerted evolution acts on divergent gene repeats, sometimes one sequence type is eliminated, changing the direction of homogenization, resulting in a mixture of paralogs and orthologs and tracing the evolutionary history of gene divergence events rather than organismal divergence events (Hillis et al., 1991). In the course of this study, it was not determined if the gene amplified are orthologs or paralogs. This problem may have contributed to the incongruent topologies observed in my phylogenies. However, seeing as how there is fairly high resolution, I believe that these discrepancies may be more attributable to the method of analyses (i.e., Bayesian vs. maximum likelihood).

The secondary structure of ribosomal RNAs consists of stem-loop structures which most likely experience compensatory base changes to maintain base pairing (Alvarez and Wendel, 2003). There are differences in the amount and placement of these stem-loop structures depending on the organism (Baldwin, 1992). These compensatory changes may lead to homoplasy, obscuring elucidation of accurate evolutionary history. This non-independence of nucleotide positions violates assumptions and may have consequences when reconstructing phylogenies (Alvarez and Wendel, 2003).

There were relatively few problems with sequence alignment and phylogenetic reconstruction, meaning it is unlikely that many of the problems associated with ribosomal loci manifested during my analyses. I did not observe any atypically long branch lengths, and sequences were not uncommonly divergent from other sequences within the taxonomic group (See Tables 5-8). Some of the disadvantages of these loci would explain the incongruent topologies observed in this analysis. Advantages to using nuclear ribosomal RNA sequences include biparental inheritance, universality, simplicity, intra-genomic uniformity (the advantage of concerted evolution), inter-genomic variability (phylogenetically informative), and low functional constraint (Alvarez & Wendel, 2003).

4.4 Divergence Estimates

The divergence time of the family Dreissenidae was placed at ~35.6 MYA. *Dreissena* began to speciate from a common ancestor ~21.7 MYA. A common ancestor of *Mytilopsis* and a MRCA of the *Congeria* + USADs diverged ~32.7 MYA. *Mytilopsis* then radiated ~22.4 MYA. The USADs then split from *Congeria* ancestors ~27.1 MYA, with a subsequent speciation ~5.1 MYA.

The methods used in this study to estimate divergence times account for rate variation and assume uncorrelated rates of evolution. Previous literature assessing divergence timing within Dreissenidae differ in their use of a strict clock (Stepien et al., 2001, 2003). However, methods applied here were also used in Bilandzija et al. (2013). All of these estimates, with the exception of the USAD divergence estimate, are very similar to the results found in other estimates for Dreissenidae.

The origin of the family (~35.6 MYA) corresponds to the earliest time

dreissenids were found in the fossil record. Similarly, divergence estimates for

Dreissena spp. are comparable to those obtained by other studies (e.g., the split

between *D. rostriformis* and *D. polymorpha* is estimated to be ~10-15 MYA, Stepien

et al., 2001, 2003; Bilandzija et al., 2013). *Mytilopsis* likely colonized the neotropics

in the late Oligocene, just before the isolation of the Paratethys Sea.

This Oligocene colonization event agrees with my estimate of *Mytilopsis* diverging

from a MRCA ~32.7 MYA just before the Oligocene began. My divergence estimate

that the USADs diverged from a *Congeria* MRCA ~27.1 MYA corresponds with the

timing of my hypothetical dispersal events (see below).

4.5 Phylogeography and Dispersal

The fact that a group of freshwater mussels restricted to cave systems in the Dinaric

Karst share a common ancestor with taxa found in interior South American

drainages seems unlikely. I estimate that the split between *Mytilopsis* and *Congeria*

occurred ~22.4 MYA (95% CI: 20.3–24.7) and the USADs then split from a MRCA of

Congeria ~27.1 MYA (95% CI: 17.1–38.3). During this time, the ancestors of modern

dreissenids were restricted to the Tethys and Paratethys seas (i.e., the site of the

present-day Mediterranean Sea). The Tethys and Paratethys seas were connected

with the Atlantic Ocean during the time period when I estimate that the USADs split

from the MRCA. The Atlantic Ocean was connected to what would become the

modern day Indian Ocean from ~37.8-20.44 MYA via the Tethys/Paratethys seaway

(Rögl, 1999). Prevailing ocean currents originating in the Indian Ocean may have

facilitated dispersal through the Tethys Seaway towards the Gulf Coast of South America in the Oligocene and early Miocene. Global temperatures were 6°C above the average from 1960-1990 (Miller et al., 2006). These temperatures were not as high as in the early Eocene, but the Arctic was still experiencing melted glaciers, leading to a higher amount of freshwater in the world's oceans, reducing overall salinity. Krapp 2012 modeled the climate evolution and large-scale ocean circulation of the middle Miocene. These models revealed a pathway for USADs to disperse from the Tethys Sea, across the Atlantic Ocean, and into the proto-Amazon. Wind patterns (Krapp, 2012; pg. 22), horizontal ocean circulation patterns (Krapp, 2012; pg. 27), and surface salinity (Krapp, 2012; pg. 30) all corroborate this hypothesis.

At this time, most of the ancestral members of Dreissenidae were extremely tolerant of brackish and marine like environments. The Tethys Sea and the resulting Lake Pannon were hypersaline (Harzhauser and Mandic, 2010). According to Krapp, 2012 (pg. 28) the Atlantic Ocean and the proto-Mediterranean had comparable temperature and salinity. When examining these figures it is plausible that any individuals dispersing from the Tethys Sea would have ended up in the north-western region of South America. During the Miocene the Amazon River flowed from the east to west, emptying into the southern Caribbean where the Panamanian isthmus is today. Eventually this system created the Pebas Sea (Wesselingh and Salo, 2006). This sea was a system of marginal marine environments, lakes, and swamps. There was also connection with northern Venezuela. This system has been touted as a vector for marine organisms to adapt to freshwater ecosystems. Many of the freshwater aquatic groups in the modern Amazon have obvious marine ancestry.

Some of these groups include freshwater stingrays, drums, anchovies, needlefish, dolphins, and manatees (Wesselingh and Salo, 2006). Most of these organisms are said to have made the transition to Amazonian freshwater environments ~23-15 MYA through the corridor of the Pebas Sea (Wesselingh and Salo, 2006). After the USADs diverged from *Congeria* and dispersed to South America, the Tethys Sea underwent a molluscan speciation event. There were as many as 30 species of *Congeria*. However, ~6 MYA the Tethys Sea was cut off from the Atlantic Ocean by an uplifting event. Evaporation caused what is known today as the Messinian Salinity Extinction. Most of the *Congeria* species were extirpated in the extreme conditions, except for *C. kusceri*. This relict mussel was able to find refuge in the freshwater cave systems of the Dinaric Karst. Here, speciation events gave rise to *C. jalzici* and *C. mulaomerovici* ~3 MYA (Bilandžija et al*.,* 2013). While the rest of the *Congeria* genus members were experiencing high levels of extinction, the USADs were speciating in the freshwater rivers of the modern Amazon. There are likely many more species of USADs to be discovered here.

If sister taxa to *Congeria* can disperse to the inland tributaries of the Amazon, are there *Mytilopsis* species in the inner Brazilian and Guianan Shield regions? Being that *Mytilopsis* has a widespread distribution and is native to the Gulf of Mexico, it is logical to assume *Mytilopsis* has dispersed to South America. *Mytilopsis lopesi* (Alvarenga and Ricci, 1989) was recently described from the Toncantins River, an adjacent sub basin to the Xingu. However, this species was described based upon morphological data alone. *Mytilopsis trautwineana* was described from the Pacific coast of South America (Graf, 2013). Again, no genetic studies have

examined data from this species. The ubiquitous invaders, *M. sallei* and *M. leucophaeata* are widespread and are likely introduced in Central and South America estuaries (Graf, 2013; Rizzo et al., 2014). Two scenarios are plausible. First, the MRCA of *Mytilopsis*, *Congeria* and the USADs dispersed into the Pebas Sea and evolved tolerance of freshwater sympatrically or second, the MRCA of *Mytilopsis* did not make it into the Pebas Sea and the species that have been recently described in South America are members of the USADs. *Mytilopsis* ancestors extant in the Gulf of Mexico may have been hindered from dispersal into the proto-Amazon by several factors. *Mytilus edulis* and *Mytilus trossulus* are marine bivalves that have shown interspecific differences in how they handle sheltered habitats versus wave-exposed coastal environments (Riginos and Cunningham, 2004). If such differences can be seen between two species, there may be intergeneric differences between *Mytilopsis* species and the USADs that kept these taxa separate during the evolution of the proto-Amazon.

Phylogenetic analyses need to be performed to refute or support the recent species descriptions of *Mytilopsis* species. Identifying the biodiversity and distribution of members of Dreissenidae will aid in the reconstruction of their evolutionary histories, ultimately informing management and conservation decisions. This would allow agencies to either keep these prolific invaders from spreading to new ecosystems or keep them from being extirpated in the search for hydroelectric conquest in the Amazon basin.

4.6 Comparative Analysis and Invasive Potential

Without taking life history characteristics and ecological perspectives into account, it would be difficult to determine which genus within Dreissenidae is sister to USADs due to incongruent phylogenies. However, when considering the reproductive strategy and morphological characteristics of this novel group, it is most likely closely related to the stygobitic genus *Congeria*. Mapping life history characteristics onto a phylogeny (Figure 17) informed an interpretation of how traits associated with invasive potential are distributed among the members of Dreissenidae. The maximum likelihood analysis of the concatenated dataset was the most parsimonious phylogeny for this task. Genera in the family Dreissenidae are known for their ability to invade new environments with great success. Although USADs are most closely related to *Congeria*, the only genera within the family that is not invasive, it seems likely that USADs have a high likelihood of being introduced to naïve freshwater ecosystems worldwide. USADs brood their larvae much like *Congeria* (Figure 17). There are several examples of invasive, larval-brooding freshwater mussels that are ubiquitous throughout North America. One is the prolific *Corbicula fluminea*, a distant relative to the Dreissenidae family. Much like *Congeria, C. fluminea* incubates fertilized eggs within its inner demibranches which act as modified marsupia, this zygote then develops through the trochophore stage to the veliger stage and is released as a pediveliger through the exhalant siphon and is approximately 200 um in shell length (Rajagopal et al., 2000). The life history characteristics of the USADs are still largely unknown. Along with larval brooding, they may have similar traits that would aid in settling into naïve ecosystems such as

high growth and fecundity (McMahon, 2011). Also human activities aid k-selected

freshwater organisms in invading naïve ecosystems. One of three scenarios is likely

in the Amazonian Basin, there may be freshwater fauna extinction, extirpating these

bivalves before they are fully known to science, they may be transported to new

environments through increased shipping traffic, or both scenarios may happen,

where these mollusks are extirpated locally and distributed globally. Taking the

appropriate management steps is vital to preventing any of these scenarios.

References

Alvarenga, L.C.F., Ricci, C.N., 1989. A new species of Mytilopsis, Conrad, 1857, from Toncantins River, Tucurui, Para, Brazil (Mollusca, Bivalvia, Dreissenidae). Mem. Inst. Oswaldo Cruz. 84, 27-33.

Alvarez, I., Wendel, J.F., 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Pylogenet. Evol. 29, 417-434.

Archibald, J.K., Mort, M.E., Crawford, D.J., 2003. Bayesian inference of phylogeny: a non-technical primer. Txaon, 52, 187-191.

Baldwin, B.G., 1992. Phylogenetic utility of the interal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Mol. Pylogenet. Evol. 1, 3-16.

Bilandžija, H.B., Morton, M. Podnar, H. Ćetković., 2013. Evolutionary history of relict Congeria (Bivalvia: Dreissenidae): unearthing the subterranean biodiversity of the Dinaric Karst. Front. Zool. 10, 5.

Chopra, S., 2011. Xingu River. Encyclopedia Britannica, School and Library.

Clement, M., Snell, Q., Walker, P., 2002. TCS: estimating gene genealogies. Proceedings 16th International Parallel and Distributed Processing Symposium.1530-2075/02.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods. 9, 772.

Douady, C.J., Delsuc, F., Boucher, Y., Doolittle, W.F., Douzery, E.J.P., 2003. Comparison of Bayesian and maximum likelihood boostrap measures of phylogenetic reliability. Mol. Biol. Evol. 20, 248-254.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut. A., 2012. Bayesian

    phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969-

    1973.

Duggen, S., Hoernle, K., van den Bogaard, P., Rupke, L., Morgan, J.P., 2003. Deep

    roots of the Messinian salinity crisis. Nature 422, 602-606.

Frischer, M.E., Handsen, A.S., Wyllie, J.A., Wimbush, J., Murray, J., Nierzwicki-

    Bauer, S.A., 2002. Specific amplification of the 18S rRNA gene as a method

    to detect zebra mussel (Dreissena polymorpha) larvae in plankton samples.

    Hydrobiologia 487, 33-44.

Gangloff, M.M., Lujan, N.K., Geda, S.R., Perkins, M., Abernathy, E., Siefferman, L.

    Biodiversity inventories and DNA barcoding reveal a cryptic new Neotropical

    zebra mussel genus (Mollusca: Bivalvia: Dreissenidae) from Clearwater rivers

    of the Brazilian and Guianan shields. Unpublished Data.

Giribet, G., Distel, D.L., 2003. Bivalve phylogeny and molecular data, in: Lydeard,

    C., Lindberg, D.R., Eds, Molecular Systematics and Phylogeography of

    Mollusks. Smithsonian Books, Washington, D.C. pp. 45-90.

Goulding, M., Barthem, R. Ferreira, E., 2003. The Smithsonian Atlas of the Amazon.

    Smithsonian Books, Washington, D.C.

Graf, D.L., 2013. Patterns of freshwater bivalves global diversity and the stat of

    phylogenetic studies on the Unionidae, Sphaeriidae, and Cyrenidae. Am.

    Malacol. Bull. 31, 135-153.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O.,

    2003. New algorithms and methods to estimate maximum-likelihood

phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307-321.

Harzhauser, M., Manic, O., 2010. Neogene dreissenids in Central Europe: European shifts and diversity changes, in: Van der Velde, G., Rajagopal, S., bij de Vaate, A., 2010. The zebra mussel in Europe. Margraf Publishers, Leiden. Pp. 11-28.

Hillis, D.M., Moritz, C., Porter, C.A., Baker, R.J., 1991. Evidence of biased gene conversion in converted evolution of ribosomal DNA. Science 251, 308-310.

Huelsenback, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754-755.

Krapp, M., 2012. Middle Miocene climate evolution: the role of large-scale ocean circulation and ocean gateways. Rep. Earth Syst. Res. 111, 1-100.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451-1452.

Mardsen, J. E., Spidle, A., May, B., 1995. Genetic similarity among zebra mussel populations within North America and Europe. Can. J. Fish. Aquat. Sci. 52, 836-847.

Mayol, M., Rossello, J.A., 2001. Why nuclear ribosomal DNA spacers (ITS) tell different stories in *Quercus.* Mol. Phylogenet. Evol. 19, 167-176.

McMahon, R.F., 2011. Evolutionary and physiological adaptations of aquatic invasive animals: r selection versus resistance. Can. J. Fish. Aquat. Sci. 59, 1235 – 1244.

Miller, K.G., 2006. Eocene-Oligocene global climate and sea level changes. Geol. Soc. Am. 120, 34-53.

Mishima, M., Ohmido, N., Fukui, K., Yahara, T., 2002. Trends in site numver change of rDNA loci during polyploid evolution in *Sanguisorba* (Rosaceae). Chromosoma 110, 550-558.

Molloy, D.P., Giamberini, L., Burlakova, L.E., Karatayev, A.Y., Cryan, J.R., Trajanovski, S.L., Trajanovska, S.P., 2011. Investigation of the endosymbionts of *Dreissena stankovici* with morphological and molecular confirmation of host species, in: Van der Velde, G., Rajagopal, S., bij de Vaate, A., Eds., 2010. The zebra mussel in Europe. Margraf Publishers, Leiden. Pp. 227-238.

Morton, B.,1993. The anatomy of Dreissena polymorpha and the evolution and success of the heteromyarian form in the Dreissenoidea, in: Napela, T.F., Schloesser, D.W., Eds., 1992. Zebra mussels: Biology, impacts, and control. CRC Press, pp. 185-215.

Morton, B., Puljas, S., 2012. Life history strategy, with ctenidial and pallial larval brooding, of the troglodytic 'living fossil' *Congeria kusceri* (Bivalvia: Dreissenidae) from the subterranean Dinaric karst of Croatia. Linn. Soc. Lon. 108, 294-314.

Morton, B., Velkovrh, F., Sket, B., 1998. Biology and anatomy of the 'living fossil' *Congeria kusceri* (Bivalvia: Dreissenidae) from subterranean rivers and caves in the Dinaric Karst of the former Yugoslavia. J. of Zool. 245, 147-174.

Nuttall, C.P., 1990. Review of the Caenozoic heterodont bivalve superfamily

Dreissenacea. Palaeontology 33, 707-737.

Pereira, D., Mansur, M.C.D., Duarte, L.D.S., Schramm de Oliveira, A., Pimpão, D.M., Callil, C.T., Ituarte, C., Parada, E., Peredo, S., Darrigran, G., Scarabino, F., Clavijo, C., Lara, G., Miyahira, I.G., Rodriguez, M.T.R., Lasso, C.,2014. Bivalve distribution in hydrographic regions in South America: historical overview and conservation. Hydrobiologia 735, 15-44.

Puljas, S., Peharda, M., Morton, B., Giljanovic, N. Š., Juric, I., 2014. Growth and longevity of the living fossil *Congeria kusceri* (Bivalvia: Dreissenidae) from the subterranean Dinaric Karst of Croatia. Malacologia 57, 353-364.

Rajagopal, S., Van der Velde, G., bij de Vaate, A., 2000. Reprosuctive biology of the Asiatic clams *Corbicula fluminalis* and *Corbicula fluminea* in the river Rhine. Arch. Hydrobiolo. 149, 403-420.

Riginos, C., Cunningham, C.W., 2004. Local adaptations and species segregation in two mussel (*Mytilus edulis* x *Mytilus trossulus*) hybrid zones. Mol. Ecol. 14, 381-400.

Rizzo, A.E., Miyahira, E.C., Moser, G., Santos, S.B., 2014. A new record of *Mytilopsis leucophaeata* (Bivalvia: Dreissenidae) in Rio de Janeiro (Brazil). Mar. Biod. Rec. 7, 1-6.

Rögl, F., 1999. Mediterranean and Paratethys, facts and hypotheses of an Oligocene to Miocene paleogeography (short overview). Geologica Carpathica 50, 339-349.

Sabaj-Pérez, M., 2015. Where the Xingu bends and will soon break. Amer. Sci. 395-403.

Schultz, A.R., 2014. Physiography of the river course, Amazon River, South

    America. Encyclopedia Britannica, School and Library, Chicago, IL.

Schütt, H., 1989. Eine neue, nearktische Congeria (Bivalvia: Dreissenidae). Archiv

    für Molluskenkunde der Senckenbergischen Naturforschenden Gesellschaft,

    120, 183-185.

Collection Mollusca SMF dataset. Senckenberg, 2009. Molluscs of the world.

    http://doi.org/dataset/10.15468/ohuexz.

Springer, M.S., DeBry, R.W., Douady, C., Amrine, H.M., Madsen, O., de Jong,

    W.W., Stanhope, M.J., 2001. Mitochondrial versus nuclear gene sequences in

    deep-level mammalian phylogeny reconstruction. Mol. Biol. Evol. 18, 132-

    143.

Stepien, C. A., Grigorovich, I.A., Gray, M.A., Sullivan, T.J., Yerga-Woolwine, S.,

    Kalayci, G., 2013. Evolutionary, biogeographic, and population genetic

    relationships of dreissenid mussels, with revision of component taxa. , in:

    Napela, T.F., Schloesser, D.W., Eds., 1992. Zebra mussels: Biology, impacts,

    and control. CRC Press, pp. 403-444.

Stepien, C.A., Morton, B., Dabrowska, K., Guarnera, R., Radja, R., Radja, B., 2001.

    Genetic diversity and evolutionary relationships of the troglobytic "living fossil"

    Congeria kusceri (Bivalvia: Dreissenidae). Mol. Ecol. 10, 1873 – 1879.

Stepien, C.A., Taylor, C.D., Grigorovich, I.A., Shirman, S.V., Wei, R., Korniushin,

    A.V., Dabrowska, K.A., 2003. DNA and systematic analysis of invasive and

    native dreissenid mussels: Is *Dreissena bugensis* really *D. rostriformis*?

    Aquat. Invaders. 14, 8 – 18.

Tamura, K. Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512-526.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30, 2725-2729.

Tan, K.S., Brian, M., 2006. The invasive Caribbean bivalve *Mytilopsis sallei* (Dreissenidae) introduced to Singapore and Johor Bahru, Malaysia. Raffles Bull. Zool. 54, 429-434.

Tavaré, S., Balding, D.J., Griffiths, R.C., Donnelly, P., 1997. Inferring coalescence times from DNA sequence data. Genetics 145, 505-518.

Therriault, T.W., Docker, M.F., Orlova, M.I., Heath, D.D., MacIsaac, H.J., 2004. Molecular resolution of the family Dreissenidae (Mollusca: Bivalvia) with emphasis on Ponto-Caspian species, including first report of *Mytilopsis leucophaeata* in the Black Sea basin. Mol. Gen. Evol. 30, 479-489.

Van der Velde, G., Rajagopal, S., bij de Vaate, A., 2010. The zebra mussel in Europe. Margraf Publishers, Leiden.

Wesselingh, E.P., Salo, J.A., 2006. A Miocene perspective on the evolution of Amazonian biota. Scripta Geologica 133, 439-458.

Winemiller, K.O., McIntyre, P.B., Castello, L., Fluet-Chouinard, E., Giarrizzo, T., Nam, S., Baird, I.G., Darwall, W., Lujan, N.K., Harrison, I., Stiassny, M.L.J., Silvano, R.A.M., Fitzgerald, D.B., Pelicice, F.M., Agostinho, A.A., Gomes, L.C., Albert, J.S., Baran, E., Petrere, M. Jr., Zarfl, C., Mulligan, M., Sullivan,

J.P., Arantes, C.C., Sousa, L.M., Koning, A.A., Hoeinghaus, D.J., Sabaj, M.,

Lundberg, J.G., Armbruster, J., Thieme, M.L., Petry, P., Zuanon, J., Torrente-

Vilara, G., Snoeks, J., Ou, C., Rainboth, W., Pavanelli, C.S., Akama, A., van

Soesbergen, A., Saenz, L., 2016. Balancing hydropower and biodiversity in

the Amazon, Congo, and Mekong. Science 351, 128-129.

Wong, Y.T., Meier, R., Tan, K.S., 2011. High hapolotype variability in established

Asian populations of the invasive Caribbean bivalve *Mytilopsis sallei*

(Dreissenidae). Biol. Invas. 13, 341.

Tables and Figures

Table 1. GenBank (GB) accession numbers, museum lot numbers, haplotype numbers, GPS coordinates and locality data for Xingu specimens examined in this analysis.

Taxon	Collection Locality	Lot Number (Haplotype)	COI	28S	16S	18S
Xingu sp. 1	1. Lower Volta Grande a rocky outcrop in the main channel ~500 m upstream from a campsite (-3.1292, -51.6653) Collected: 2012 and 2016	1757(6)	KU513976	--	--	--
		1758(10)	KU513977	--	--	--
		1767(9)	KU513978	--	--	--
		1769(1)	KU13972	--	--	--
		1770(4)	KU13970	--	--	--
		1754(12)	KX027437	--	--	--
		29VIII16.2.4 (27)		--	--	
		29VIII16.2.8 (30)		--		
		29VIII16.2.11 (27)			--	
		29VIII16.2.12 (28)			--	
		29VIII16.2.13 (6)			--	--
		29VIII16.2.15 (24)			--	
		29VIII16.2.49 (25)			--	
		29VIII16.2.61 (25)	--		--	--
		29VIII16.2.62 (29)				--
Xingu sp. 1	2. Lower Volta Grande downstream of Cachoeira Tamaracá, off the left bank of a large braid of river (-3.12820, -51.62143) Collected: 2012	1858(11)	KU513983	--	--	
		1859(16)	KU513984	--	--	--

Taxon	Collection Locality	Lot Number (Haplotype)	COI	28S	16S	18S
Xingu sp. 1	3. Lower Volta Grande a main straight channel running from south-southwest to north-northeast (-3.18427, -51.61735) Collected: 2012	1813(15)	KU513979	--	--	--
		1808(2)	KX027436	--	--	--
Xingu sp. 1	4. Deep channel along the right bank of the river ~38 km southeast of Vitória do Xingu (-3.09233, -51.73725) Collected: 2012	1714(3)	KU513973	--	--	--
		1715(5)	KU513974	--	--	--
		1718(13)	KU513975	--	--	--
		1720(14)	KU13971	--	--	--
Xingu sp. 2	5. Xingu River, along right bank ~ 14 km upstream of Rio Iriri confluence (-3.9392, -52.5790) Collected: 2012	1841(7)	KU513980	--	--	--
		1842(8)	KU513981	--	--	--
Ventuari sp.	6. Ventuari River, Island in deltaic confluence with the Orinoco River, 71.5 km east of San Fernando do Atabapo (-3.97841, -67.06047) Collected: 2016	Ven1	KU13969	--	--	--
Congeria hoeblichi	7. Iriri River, Cahoeira Grande, ~12 km upstream confluence with Xingu River, above and below rapids (-3.84196, 52.73487) Collected: 2016	26VIII16.1.12 (28)				
		26VIII16.1.17 (19)		--	--	
		26VIII16.1.21 (21)			--	
Congeria hoeblichi	8. Xingu River, Rebojo do Avelino, ~10 km downstream confluence with Iriri River, right	27VIII16.1.4 (22)			--	

Taxon	Collection Locality	Lot Number (Haplotype)	COI	28S	16S	18S
	descending bank (-3.75201, -52.51154) Collected: 2016	27VIII16.1.5 (18)			--	
		27VIII16.1.6 (23)			--	--
Xingu sp. 1 or 2 *Congeria hoeblichi*	9. Xingu River, Cachoeira do Espelho, ~26 km downstream confluence with Iriri River, and ~45 km upstream of Altamira. Right descending bank, below rapids (-3.64477, -52.37960) Collected: 2016	27VIII16.2.17 (19)	--	--		--
Congeria hoeblichi	10. Xingu River, Ja Bota, ~30 km downstream confluence with Iriri River, and ~40 km upstream of Altamira.Right descending bank (-3.62195, -52.36140) Collected: 2016	28VIII16.1.1 (20)		--		--
		28VIII16.1.2 (20)		--		--
		28VIII16.1.3 (20)		--		--
Xingu sp. 1 or 2 *Congeria hoeblichi*	11. Xingu River, Cachoeira Itamaraca, ~4 km upstream Belo Monte, below rapids (-3.14695, -51.65779) Collected: 2016	29VIII16.1.3 (26)				--
		29VIII16.1.4 (26)			--	

Table 2. Primers used for phylogenetic analyses. Mitochondrial and nuclear genes allowed us to evaluate deep and shallow evolutionary relationships.

Locus	Origin	Forward	Reverse	Expected Product Size
COI	Mitochondrial	GTTCCACAAATCAT AAGGATATTGG	TACACCTCAGGGTG ACCAAAAAACCA	700 bp
16S	Mitochondrial	CCGTTCTGAACTCA GCTCATGT	CGACTGTTTAACAAA AACAT	460 bp
28S	Nuclear	TCCGATAGCGCACA AGTAC	TTGCACGTCAGAATC GCTA	600 bp
18S	Nuclear	CTGCCAGTAGTCAT ATGC	ACCTTGTTACGACTT TAC	1800 bp

Table 3. Taxa, accession numbers, and localities for sequences downloaded from GenBank that were utilized for outgroups in this analysis. Dashes indicate that there was no representative for that species at that locus.

Taxon	COI	28S	16S	18S	Locality
Dreissena presbensis	EF414478.1[1]	EF414469.1[1]	EF414449.1[1]	-	[1]Greece: Lake Dojran
	EF414479.1[2]	EF414470.1[2]	EF414450.1[2]	-	[2]Greece: Lake Vegoritis
	-	EF414474.1[3]	EF414455.1[3]	-	[3]Macedonia: Lake Ohrid
	-	EF414475.1[4]	EF414460.1[4]	-	[4]Montenegro: Lake Skutari
	-	EF414476.1[4]	EF414461.1[4]	-	
Dreissena stankovici	DQ840108.1	DQ333768.1	DQ333703.1	-	Macedonia: Lake Ohrid
Dreissena blanci	EF414481.1	EF414471.1	EF414452.1	-	Greece: Lake Trichonis
Dreissena bugensis	-	-	JQ348913.1[5]	-	[5]USA: Lake Erie, OH
	JX099436.1[6]	-	JX099457.1[6]	JX099479.1[6]	[6]Netherlands: Ijsselmeer, Lelystadt
	DQ840132.1[7]	FJ455425.1[8]	AF038996.1[9]	-	[7]Mediterannean, Black sea
					[8]USA: Lake Mead
					[9]Croatia: Jama u Predolcu, Metkovic
Dreissena polymorpha	JX099437.1[10]	JX099499.1[10]	JX099458.1[10]	JX099478.1[10]	[10]Croatia: Jarun Lake, Zagreb
	EF414493.1[11]	-	-	-	[11]Turkey: Lake Buyukcekmece
	EF414494.1[12]	-	EF414465.1[12]	AM774543.1[13]	[12]Germany: Lake Tressow
	EF414495.1[14]	-	EF414466.1[14]	AF120552.1[15]	[13]Netherlands: Amsterdam
	KC429149.1[15]	-	DQ280038.1[15]	-	[14]Romania: Lake Razim
	-	-	AF507049.1[16]	-	[15]Unknown
	-	-	AF038997.1[15]	-	[16]Ukraine: Dniester Liman
Dreissena rostriformis	KP057252.1[17]	JQ700562.1[18]	AF507048.1[19]	-	[17]United Kingdom: Great Britian
	-	JQ700563.1[18]	AY302247.1[15]	-	[18]Caspian Sea: near Azerbaijan
	-	JQ700564.1[18]	-	-	[19]Ukraine: Dniester Liman, Black Sea

Taxon	COI	28S	16S	18S	Locality
Mytilopsis leucophaeata	HM100251.1[20]	EF414468.1[20]	EF414448.1[20]	KX713323.1[21]	[20]Belgium: Antwerp Harbour [21]USA: Florida Keys
Mytilopsis sallei	JX099435.1	JX099497.1	JX099455.1	JX099476.1	China: Hong Kong
Congeria kusceri	JX099430.1	-	JX099450.1	JX099471.1	Croatia: Pukotina e Tunelu Polje Jezero-Peracko Blato, Ploce, S. Dalamatia
Congeria kusceri	JX099419.1	JX099481.1	JX099439.1	JX099460.1	Bosnia and Herzegovina: Doljasnica, Popovo Polje
Congeria kusceri	JX099420.1	JX099482.1	JX099440.1	JX099461.1	Bosnia and Herzegovina: Gradnica, Neum
Congeria kusceri	JX099423.1	JX099485.1	JX099443.1	JX099464.1	Croatia: Jama u Predolcu, Metkovic, S. Dalmatia
Congeria kusceri	JX099424.1	JX099486.1	JX099444.1	JX099465.1	Croatia: Jasena Ponor, Vrgorac, S. Dalmatia
Congeria kusceri	JX099429.1	JX099491.1	JX099449.1	JX099470.1	Croatia: Pukotina e Tunelu Polje Jezero-Peracko Blato, Ploce, S. Dalamatia
Congeria jalzici	JX099421.1	JX099483.1	JX099441.1	JX099462.1	Slovenia: Izvir Jamske Skoljke, Metlika, Bela Krajina
Congeria mulaomerovici	JX099418.1	JX099480.1	JX099438.1	JX099459.1	Bosnia and Herzegovina: Dabarska Pecina, Sanski Most, N. Bosnia
Corbicula fluminea	KU905760.1[22]	AM779732.1[23]	AF152024.1[24]	AM774558.1[23]	[22]USA: MA, Choptank River [23]United Kingdom: Norfolk [24]USA: MI. Huron River, Ann Arbor
Corbula amurensis	KJ028746.1	-	-	-	China
Cyrenoida floridana	KC429123.1[15]	FM999790.1[25]	KC429280.1[15]	FM999789.1[25]	[25]USA: Big Pine Key, Blue Hole, FL
Glauconome rugosa	KC429140.1[15]	DQ184799.1[28]	KC429302.1[15]	KC429392.1[15]	[26]Vietnam: Market, Ho Chi Minh City, Ho Chi Minh Province
Moerella iridescens	JN859967.1[15]	-	AB751330.1[27]	EF613237.1[15]	[27]Japan: Yamaguchi, Estuary of Kiya River

Taxon	COI	28S	16S	18S	Locality
Mya arenaria	KX576717.1[15]	FM999792.1[28]	KT959487.1[29]	FM999791.1[28]	[28]Poland: Gydnia [29]USA: MA, Chesapeake Bay, West River
Phaxas pellucidus	-	KC429508.1	KC429309.1	KC429400.1	Unknown
Rangia cuneata	-	KC429509.1[15]	KT959495.1[30]	KC429401.1[15]	[30]USA: MA, Rhode River, Canning House Bay
Sinonovacula constricta	-	AF131005.1[15]	AB751361.1[31]	AY695800.2[15]	[31]South Korea: Suguru Ujino

Table 4. Log likelihood estimates for phylogenies created in this study.

Dataset	Maximum Likelihood	Sister Clade to SAD	Bayesian	Sister Clade
COI	-4,769.851	*Congeria*	-4,877.634	*Mytilopsis*
28S	-3,811.721	*Congeria*	-3,855.365	*Mytilopsis*
16S	-3,309.788	*Congeria*	-3,519.396	*Mytilopsis*
18S	-4,874.870	*Mytilopsis*	-4,584.214	*Mytilopsis*
Mitochondrial	-6,731.026	*Congeria*	-6,762.967	*Congeria*
Nuclear	-7,276.556	*Mytilopsis*	-7,290.938	*Mytilopsis*
Concatenated	-20,699.607	*Congeria*	-21,657.394	*Mytilopsis*

Table 5. Genotypes and representative sequences for 28S, 16S, and 18S. For collection site see Table 2.

Gene	Genotype	Sequence	Species
28S	Xingu 1	29VIII16.2.15	Xingu sp. 1
	Xingu 2	27VIII16.1.6	Xingu sp. 2
	Xingu 3	29VIII16.2.12	Xingu sp. 1
	Xingu 4	27VIII16.1.5	Xingu sp. 2
16S	Xingu 1	26VIII16.1.17	Xingu sp. 1
	Xingu 2	27VIII16.1.4	Xingu sp. 1
	Xingu 3	29VIII16.2.4	Xingu sp. 2
	Xingu 4	29VIII16.2.8	Xingu sp. 2
	Xingu 5	29VIII16.2.62	Xingu sp. 2
18S	Xingu 1	29VIII16.2.15	Xingu sp. 1
	Xingu 2	29VIII16.2.8	Xingu sp. 1
	Xingu 3	29VIII16.2.4	Xingu sp. 1
	Xingu 4	28VIII16.1.1	Xingu sp. 2
	Xingu 5	26VIII16.1.12	Xingu sp. 2
MtDNA	Xingu 1	29VIII16.2.12	Xingu sp. 1
	Xingu 2	29VIII16.2.8	Xingu sp. 1
	Xingu 3	29VIII16.2.62	Xingu sp. 1
	Xingu 4	29VIII16.2.12	Xingu sp. 1
	Xingu 5	27VIII16.1.4	Xingu sp. 2
	Xingu 6	28VIII16.1.1	Xingu sp. 2
NDNA	Xingu 1	29VIII16.2.12	Xingu sp. 1
	Xingu 2	29VIII16.2.61	Xingu sp. 1
	Xingu 3	29VIII16.1.4	Xingu sp. 1
	Xingu 4	29VIII16.2.15	Xingu sp. 1
	Xingu 5	27VIII16.1.4	Xingu sp. 2
Concatenated	Xingu 1	29VIII16.2.11	Xingu sp. 1
	Xingu 2	26VIII16.1.12	Xingu sp. 2

Table 6. Uncorrected p-distances (x100) for the COI data set used in phylogenetic analyses. Column number corresponds to taxon row numbers. Intra-generic comparisons are shaded.

Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. *M. iridescens*	---																			
2. *C. fluminea*	34.9	---																		
3. *C. floridana*	37.6	18.7	---																	
4. *G. rugosa*	37.1	18.7	15.5	---																
5. *C. amurensis*	38.1	31.0	31.0	31.0	---															
6. *M. arenaria*	41.3	36.4	35.4	34.2	27.3	---														
7. *C. kusceri*	36.2	31.9	31.9	29.6	27.6	27.4	---													
8. *C. mulaomerovici*	34.6	30.5	31.7	30.7	26.8	26.3	7.1	---												
9. *C. jalzici*	34.6	30.7	31.7	31.0	27.0	26.8	7.6	4.7	---											
10. *D. bugensis*	38.6	33.7	33.7	30.5	29.2	25.3	19.3	18.9	19.7	---										
11. *D. polymorpha*	36.3	32.8	32.9	31.9	28.9	27.8	18.3	17.2	18.4	15.7	---									
12. *D. rostriformis*	38.6	33.4	33.4	30.2	29.0	25.3	19.3	18.9	19.7	0.2	15.9	---								
13. *D. presbensis*	36.6	32.8	35.0	33.0	29.0	29.5	20.0	19.3	20.0	17.4	10.2	17.2	---							
14. *D. stankovici*	37.1	33.2	34.6	33.2	29.2	29.5	19.9	19.2	20.1	17.2	1.2	17.0	0.6	---						
15. *D. blanci*	36.9	32.4	32.9	31.9	29.2	27.8	19.5	17.4	18.9	17.2	10.6	17.0	3.6	2.9	---					
16. *M. leucophaeata*	32.4	31.7	33.9	29.5	27.8	24.6	15.8	15.5	15.2	17.2	17.4	17.2	19.2	19.7	19.2	---				
17. *M. sallei*	37.1	31.7	31.7	29.5	24.3	27.0	14.7	14.0	14.7	17.7	17.0	17.4	17.2	17.4	16.7	13.3	---			
18. Venturai n. sp.	35.1	30.2	31.0	27.5	24.6	26.3	15.6	16.0	17.0	17.0	15.7	16.7	17.8	18.2	17.0	13.8	13.0	---		
19. Xingu sp. 1	34.7	29.7	30.9	26.9	24.7	24.7	14.0	14.4	13.6	18.4	16.3	18.1	17.8	17.6	16.6	12.8	14.8	8.3	---	
20. Xingu sp. 2	34.5	30.0	30.1	27.5	24.7	26.6	14.1	14.1	13.9	19.1	17.6	18.8	18.3	18.6	17.4	14.6	13.9	8.2	7.6	---
Intra-taxon	n/a	n/a	n/a	n/a	n/a	n/a	0.9	n/a	n/a	n/a	n/a	n/a	0.2	n/a	0.3	n/a	n/a	n/a	1.2	3.1

Table 7. Uncorrected p-distances (x100) for the 28S data set used in phylogenetic analyses. Column number corresponds to taxon row numbers. Intra-generic comparisons are shaded.

Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. *R. cuneata*	---																		
2. *G. rugosa*	15.4	---																	
3. *C. floridana*	18.5	8.6	---																
4. *C. fluminea*	19.4	8.6	7.0	---															
5. *P. pellucidus*	25.3	23.9	26.2	26.2	---														
6. *M. arenaria*	18.5	18.5	20.5	19.9	23.5	---													
7. *C. kusceri*	21.5	21.3	22.4	22.9	26.2	14.5	---												
8. *C. mulaomerovici*	21.4	21.0	22.1	23.0	26.2	14.4	0.3	---											
9. *C. jalzici*	21.5	21.2	21.9	22.8	26.2	14.2	0.8	0.5	---										
10. *D. bugensis*	22.8	24.1	24.2	25.3	26.6	16.5	8.8	8.8	8.6	---									
11. *D. polymorpha*	23.2	22.4	23.0	24.1	27.5	15.3	8.0	7.9	8.3	6.8	---								
12. *D. rostriformis*	22.7	23.9	24.1	25.2	26.6	16.4	8.8	8.8	8.6	0.2	6.8	---							
13. *D. presbensis*	23.3	22.6	22.8	24.1	28.0	15.8	8.9	8.8	9.2	7.4	2.2	7.4	---						
14. *D. stankovici*	23.5	23.0	23.2	24.4	28.4	16.2	9.1	9.0	9.3	7.7	2.5	7.7	0.5	---					
15. *D. blanci*	23.7	22.8	23.3	24.4	28.2	16.3	8.9	8.8	9.2	7.9	2.3	7.9	0.7	0.9	---				
16. *M. leucophaeata*	22.4	22.4	23.3	23.7	26.0	14.7	5.1	5.0	5.2	10.6	10.4	10.6	11.3	11.5	11.3	---			
17. *M. sallei*	22.1	22.3	23.2	23.7	25.9	14.2	5.0	4.8	5.0	10.6	10.1	10.6	10.6	10.8	10.6	2.5	---		
18. Xingu sp. 1	21.0	21.5	21.7	23.2	26.4	13.1	3.7	3.8	3.9	9.9	7.9	9.9	8.4	8.6	8.8	6.5	5.6	---	
19. Xingu sp. 2	21.	21.3	21.8	22.9	26.2	13.7	3.5	3.7	3.8	10.1	7.8	10.1	8.7	8.9	8.7	6.2	5.3	0.6	---
Intra-taxon	n/a	n/a	n/a	n/a	n/a	n/a	0.1	n/a	n/a	n/a	n/a	4.0	0.0	n/a	n/a	n/a	n/a	0.0	0.5

Table 8. Uncorrected p-distances (x100) for the 16S data set used in phylogenetic analyses. Column number corresponds to taxon row numbers. Intra-generic comparisons are shaded.

Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. *S. constricta*	---																				
2. *M. iridescens*	32.2	---																			
3. *P. pellucidus*	9.1	33.9	---																		
4. *R. cuneata*	28.0	31.6	27.7	---																	
5. *C. fluminea*	28.0	33.0	25.7	25.7	---																
6. *C. floridana*	28.6	34.5	27.7	29.5	16.2	---															
7. *G. rugosa*	30.4	34.2	30.1	30.7	19.5	19.8	---														
8. *M. arenaria*	28.0	33.3	28.9	31.0	29.2	27.4	31.9	---													
9. *C. kusceri*	21.2	31.9	20.4	28.3	29.5	28.9	31.9	25.1	---												
10. *C. mulaomerovici*	22.1	32.2	21.8	29.5	30.4	30.4	32.4	24.5	1.8	---											
11. *C. jalzici*	21.5	31.9	21.2	28.9	29.8	29.8	31.9	24.5	1.2	0.6	---										
12. *D. bugensis*	25.1	32.1	23.9	29.7	28.8	30.0	32.4	25.4	8.8	10.0	9.4	---									
13. *D. polymorpha*	22.4	31.0	22.7	29.2	29.5	30.7	33.3	25.4	8.6	8.3	8.3	6.4	---								
14. *D. rostriformis*	25.1	32.2	23.9	29.8	28.9	30.1	32.4	25.4	9.1	10.3	9.7	0.3	6.5	---							
15. *D. presbensis*	21.5	30.1	21.2	28.6	29.2	29.5	32.4	25.1	6.8	7.4	7.4	5.2	2.7	5.3	---						
16. *D. stankovici*	21.5	30.1	21.2	28.6	29.2	29.5	32.4	25.1	6.8	7.4	7.4	5.2	2.7	5.3	0.0	---					
17. *D. blanci*	21.5	30.1	22.2	28.6	29.2	29.5	32.4	25.1	6.8	7.4	7.4	5.2	2.7	5.3	0.0	0.0	---				
18. *M. leucophaeata*	23.4	33.0	22.4	29.0	30.1	30.4	34.4	24.8	6.7	7.3	6.7	10.8	9.7	10.9	8.2	8.2	8.2	---			
19. *M. sallei*	22.7	31.6	22.4	28.0	31.0	29.8	31.6	25.1	7.4	8.0	7.4	11.5	11.2	11.5	8.8	8.8	8.8	7.3	---		
20. Xingu sp. 1	21.6	32.7	21.9	28.7	31.3	30.7	31.7	26.4	5.5	5.5	4.9	11.5	9.5	11.6	8.1	8.1	8.1	7.0	8.0	---	
21. Xingu sp. 2	21.6	31.3	21.9	28.6	30.4	30.5	31.9	25.4	4.5	4.4	3.8	10.0	7.9	10.1	6.6	6.6	6.6	5.6	6.6	1.7	---
Intra-taxon distance	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.0	n/a	n/a	0.6	0.0	0.0	0.0	n/a	n/a	0.4	n/a	0.4	0.1

Table 9. Uncorrected p-distances (X100) for the 18S data set used in phylogenetic analyses. Column number corresponds to taxon row numbers. Intra-generic comparisons are shaded.

Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. *M. iridescens*	---															
2. *C. fluminea*	8.8	---														
3. *P. pellucidus*	9.4	5.8	---													
4. *R. cuneate*	9.6	3.6	6.8	---												
5. *C. floridana*	8.3	0.8	5.8	3.5	---											
6. *G. rugose*	8.9	1.6	6.1	3.9	1.5	---										
7. *M. arenaria*	9.1	4.3	6.3	5.4	4.3	4.5	---									
8. *C. kusceri*	10.9	6.2	8.8	7.6	6.1	6.2	5.3	---								
9. *C. mulaomerovici*	10.8	6.3	8.8	7.6	6.2	6.2	5.3	0.1	---							
10. *C. jalzici*	10.9	6.4	8.9	7.7	6.2	6.2	5.4	0.1	0.1	---						
11. *D. bugensis*	11.2	6.6	8.2	8.2	6.6	6.8	5.4	1.5	1.5	1.5	---					
12. *D. polymorpha*	10.6	6.0	8.6	7.3	6.0	6.0	5.2	0.4	0.5	0.5	1.3	---				
13. *M. leucophaeata*	10.6	6.0	8.8	7.3	5.8	6.0	5.4	0.6	0.7	0.7	1.6	0.3	---			
14. *M. sallei*	10.6	6.0	8.8	7.3	5.8	6.0	5.4	0.6	0.7	0.7	1.6	0.3	0.0	---		
15. Xingu sp. 1	10.6	6.0	8.7	7.2	5.7	6.0	5.2	1.0	1.1	1.1	2.1	0.9	0.5	0.5	---	
16. Xingu sp. 2	10.6	6.0	8.7	7.2	5.7	6.0	5.2	1.0	1.1	1.1	2.1	0.9	0.5	0.5	0.1	---
Intra-taxon distance	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.0	n/a	n/a	n/a	0.0	n/a	n/a	0.1	0.1

Table 10. Hypotheses proposed by phylogenetic analyses. Values are nodal support (Maximum likelihood/Bayesian). X indicates that clade was not present in a given phylogeny.

Monophyletic Grouping	COI	28S	16S	18S	MtDNA	NDNA	Concatenated
Dreissenidae	88.8 / 100	100 / 100	30.7 / 96.7	99.2 / 100	97.1 / 100	100 / 100	100 / 100
USADs	87.9 / 92.1	85.6 / 64.5	94.4 / 82.9	85.4 / 74.3	94.1 / 100	91.2 / 98	100 / 100
Dreissena	92.5 / 100	100 / 100	68.5 / 100	X / X	85.6 / 100	35.1 / 76.3	99.8 / 100
Mytilopsis	23.4 / 57.9	100 / 100	X / X	81.1 / 84.2	X / X	100 / 100	100 / 100
Congeria	96.5 / 100	96.5 / 100	80 / 63.8	95.6 / 100	53.5 / 99.3	99.8 / 100	100 / 100
USADs + *Dreissena*	X / X	X / X	X / X	X / X	X / X	X / X	X / X
USADs + *Mytilopsis*	X / 77.0	X / X	48.0 / 63.8	100 / 100	X / X	100 / 100	93 / 61.8
USADs + *Congeria*	8.1 / X	41.4 / 59.9	53.7 / 50.7	X / X	69.7 / 73.7	X / X	71.5 / 84.2
Dreissena +							

Mytilopsis	65.1 / X	45.2 / X	X / 53.3	54.6 / 65.1	39.4 / 73.7	57.1 / 96.7	100 / 100
Dreissena +							
Congeria	X / 77.0	X / 100	X / X	47.5 / 66.4	X / X	35.1 / 76.3	X / X
Mytilopsis +							
Congeria	24.5 / 57.9	100 / 100	X / X	X / X	X / X	X / X	X / X
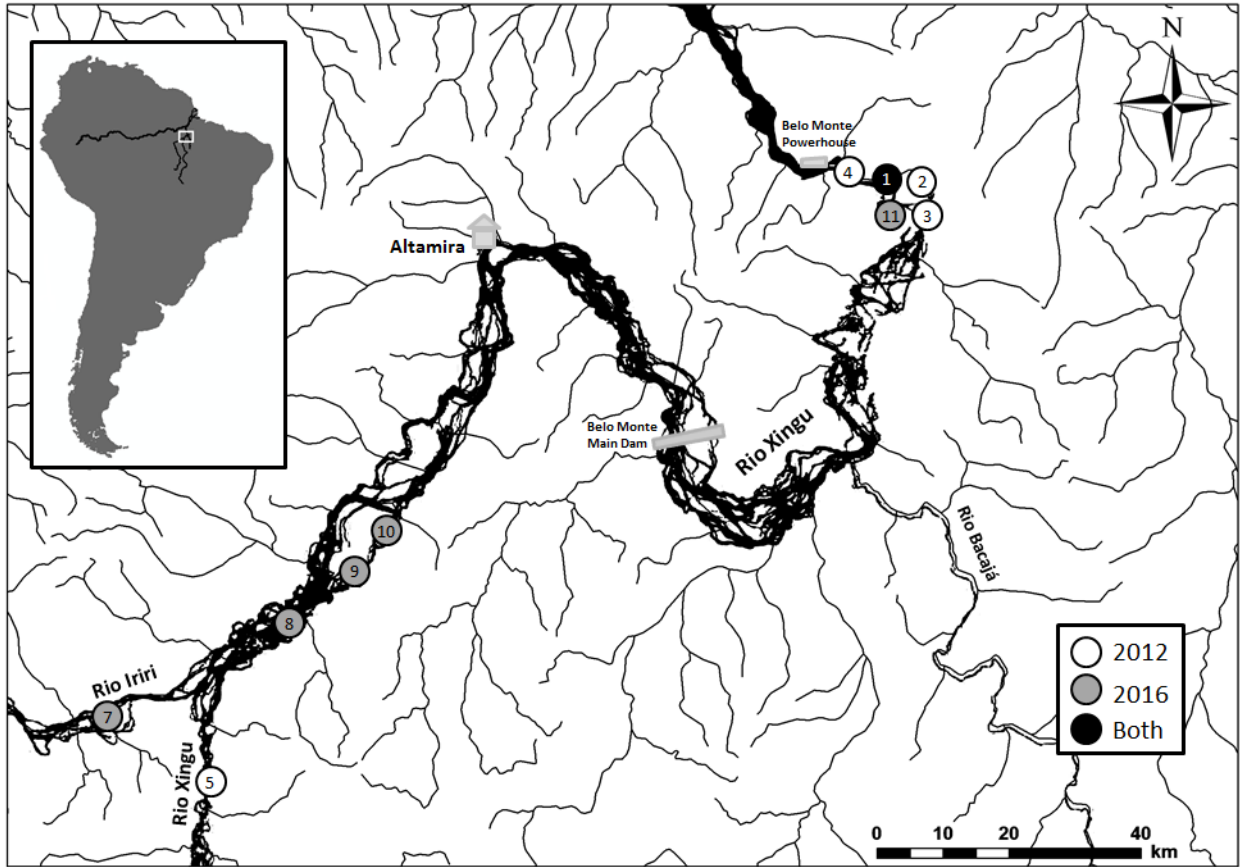
Figure 1. Map of sampling localities on the Xingu and Iriri Rivers in Para, Brazil. Sites denoted with a white dot were sampled in 2012, sites denoted with a gray dot were sampled in 2016, and sites denoted with a black dot were sampled in both years. Site number corresponds to table 2. Site 6 is not shown on this map, as it was the only site in the Ventuari River, Venezuela.

Figure 2. *Congeria hoeblichi*, Collected August 27th, 2016 from Cachoeira do Espelho, Xingu River, Para, Brazil.
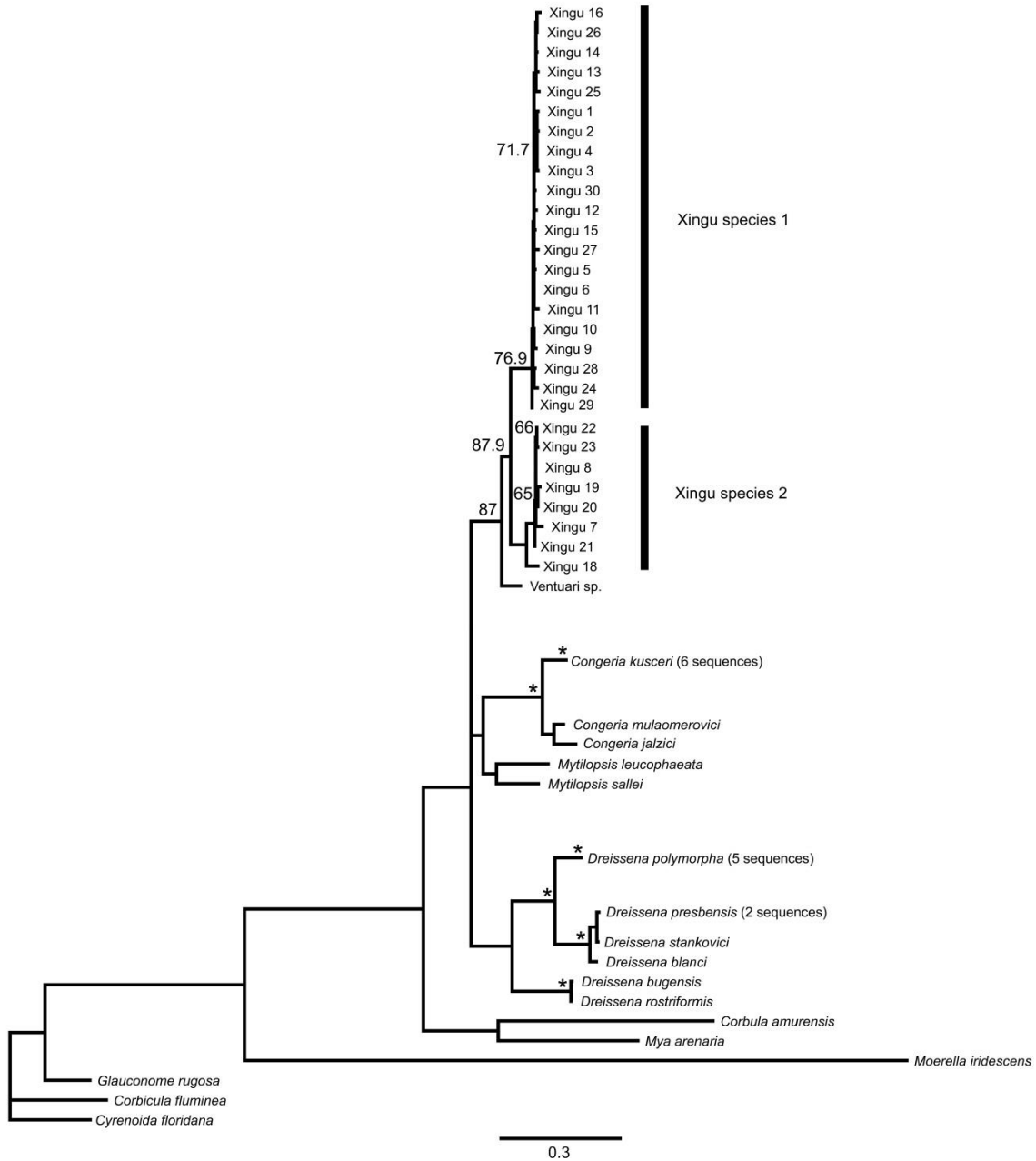
Figure 3. Maximum likelihood analysis of COI dataset. An asterisk indicates bootstrap proportions of 95% and above. Node labels (bootstrap proportions) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.
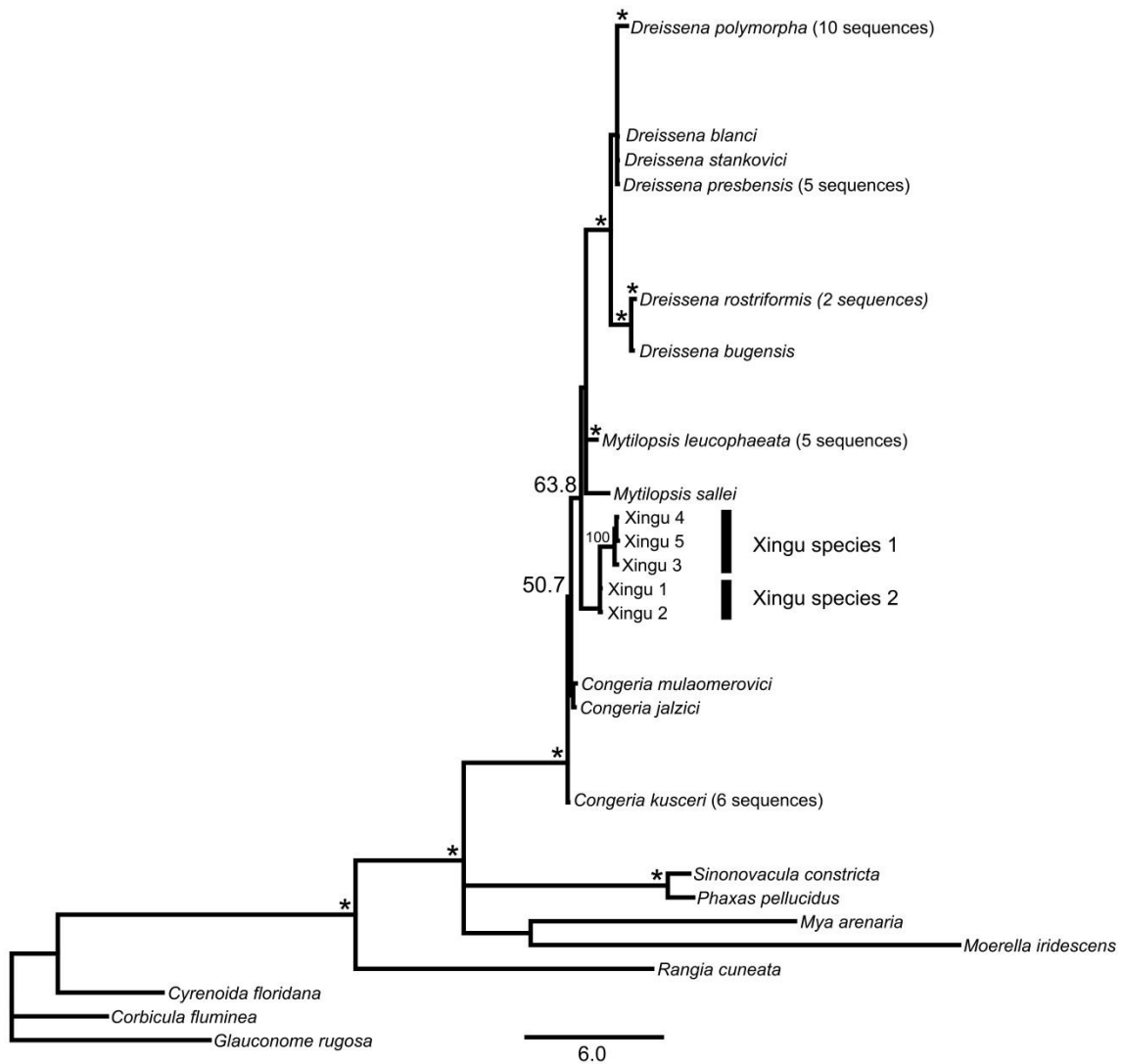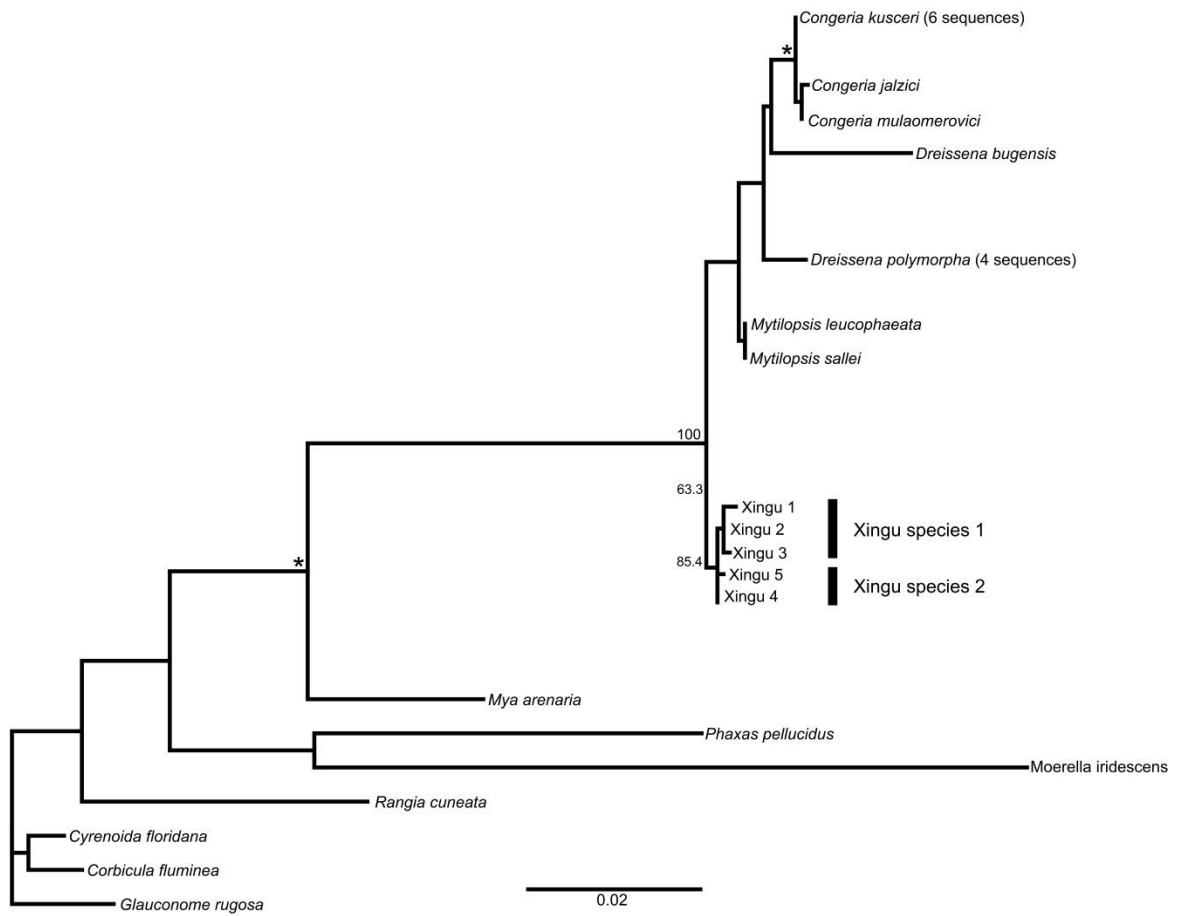
Figure 4. Bayesian analysis of COI dataset. An asterisk indicates posterior probabilities of 95% and above. Node labels (posterior probabilities) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.
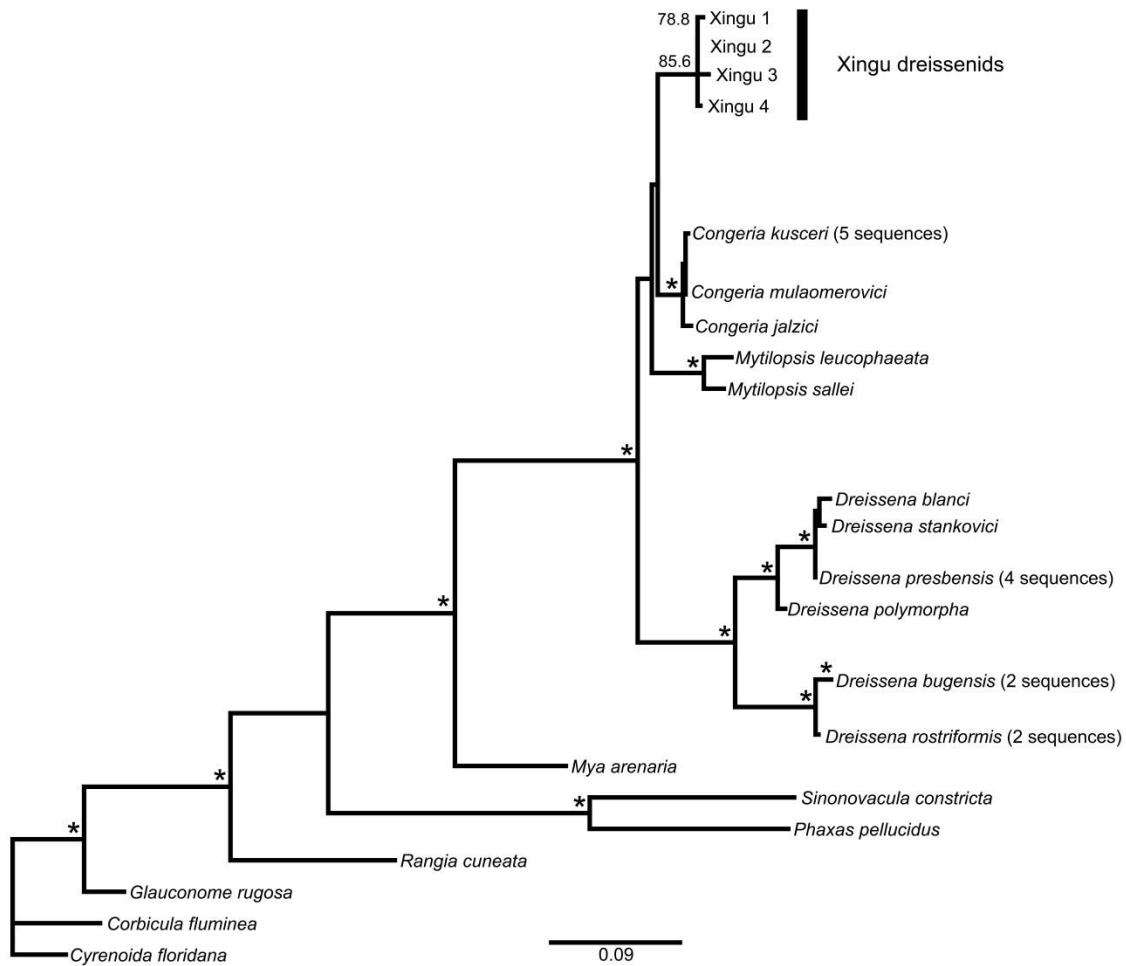
Figure 5. Maximum likelihood analysis of 28S dataset. An asterisk indicates bootstrap proportions of 95% and above. Node labels (bootstrap proportions) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.
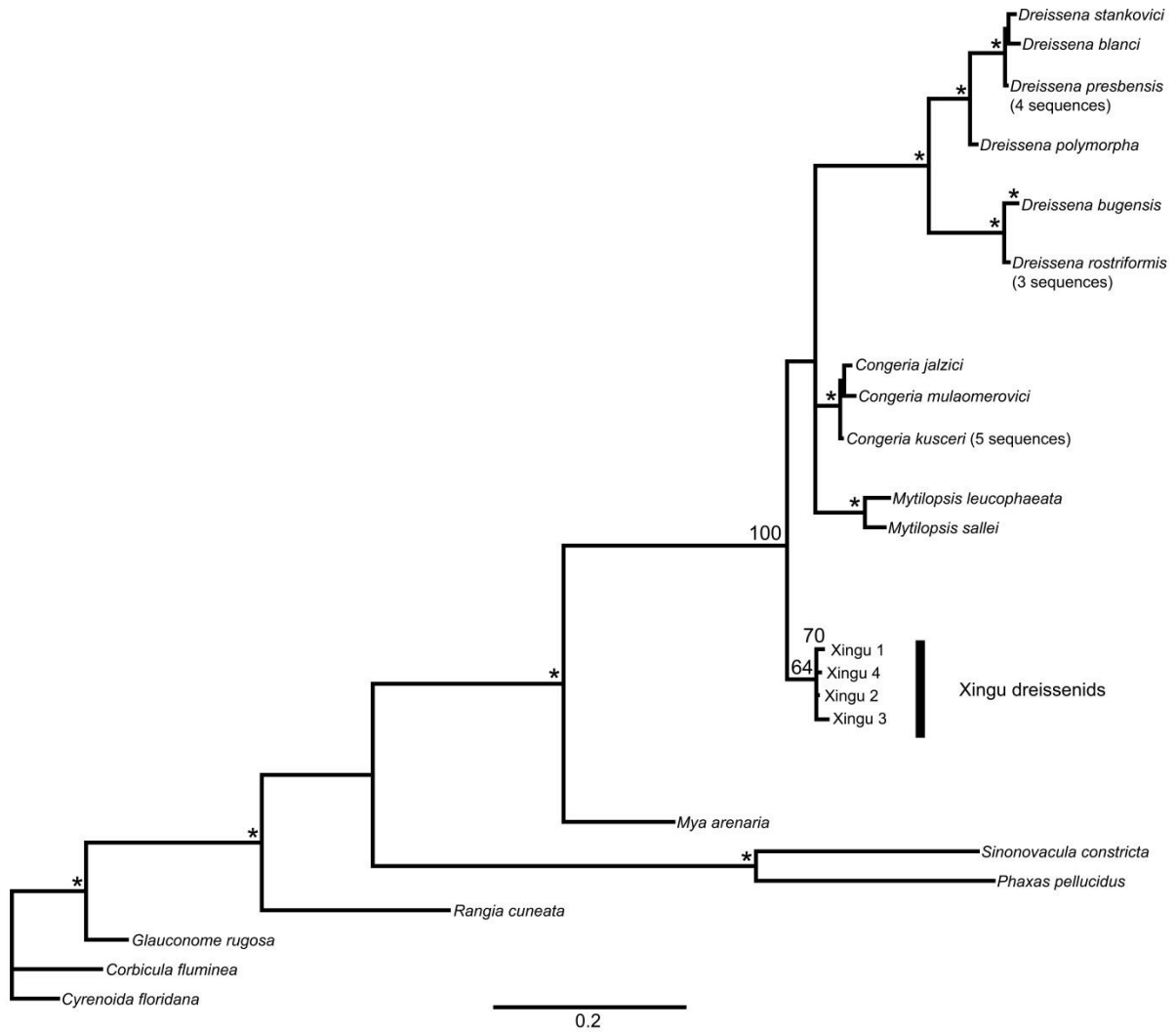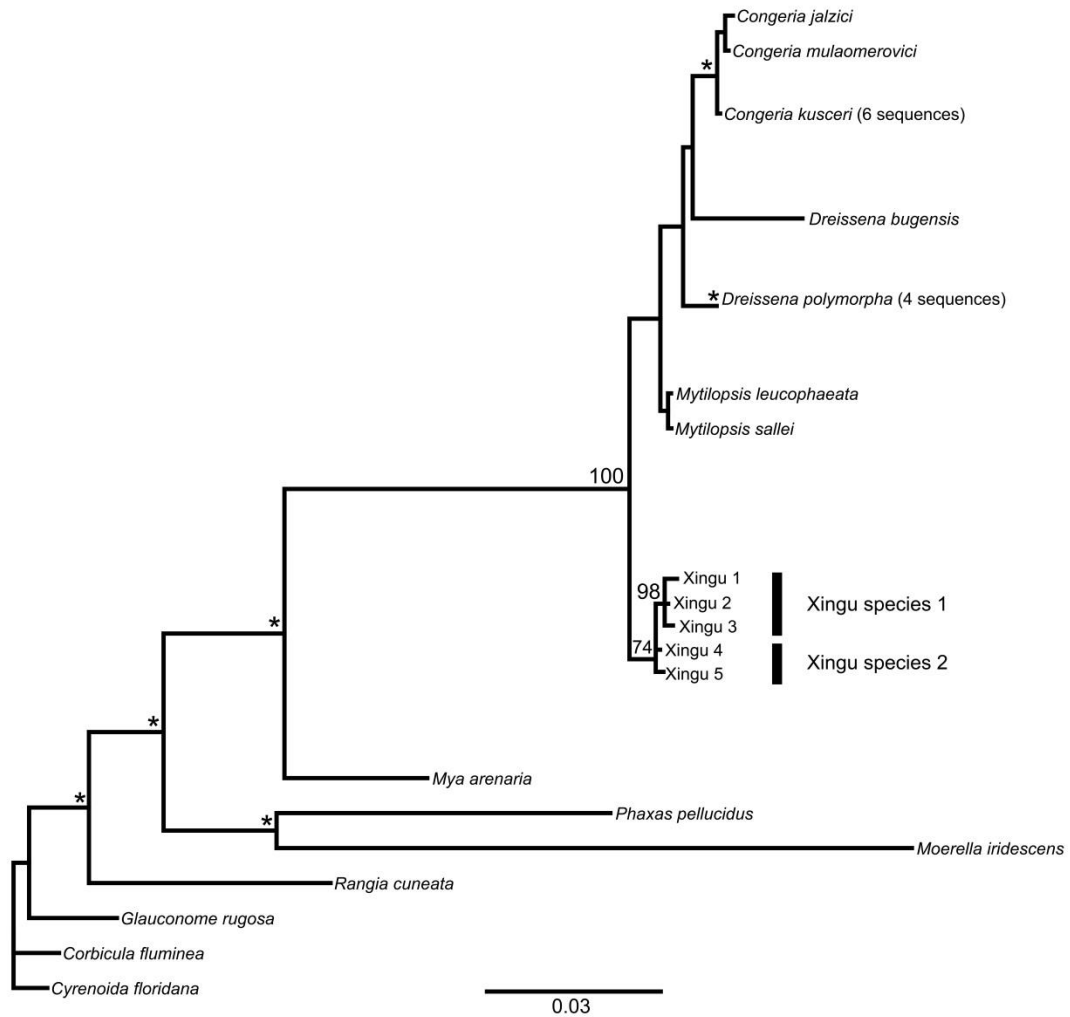
Figure 6. Bayesian analysis of 28S dataset. An asterisk indicates posterior probabilities of 95% and above. Node labels (posterior probabilities) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.
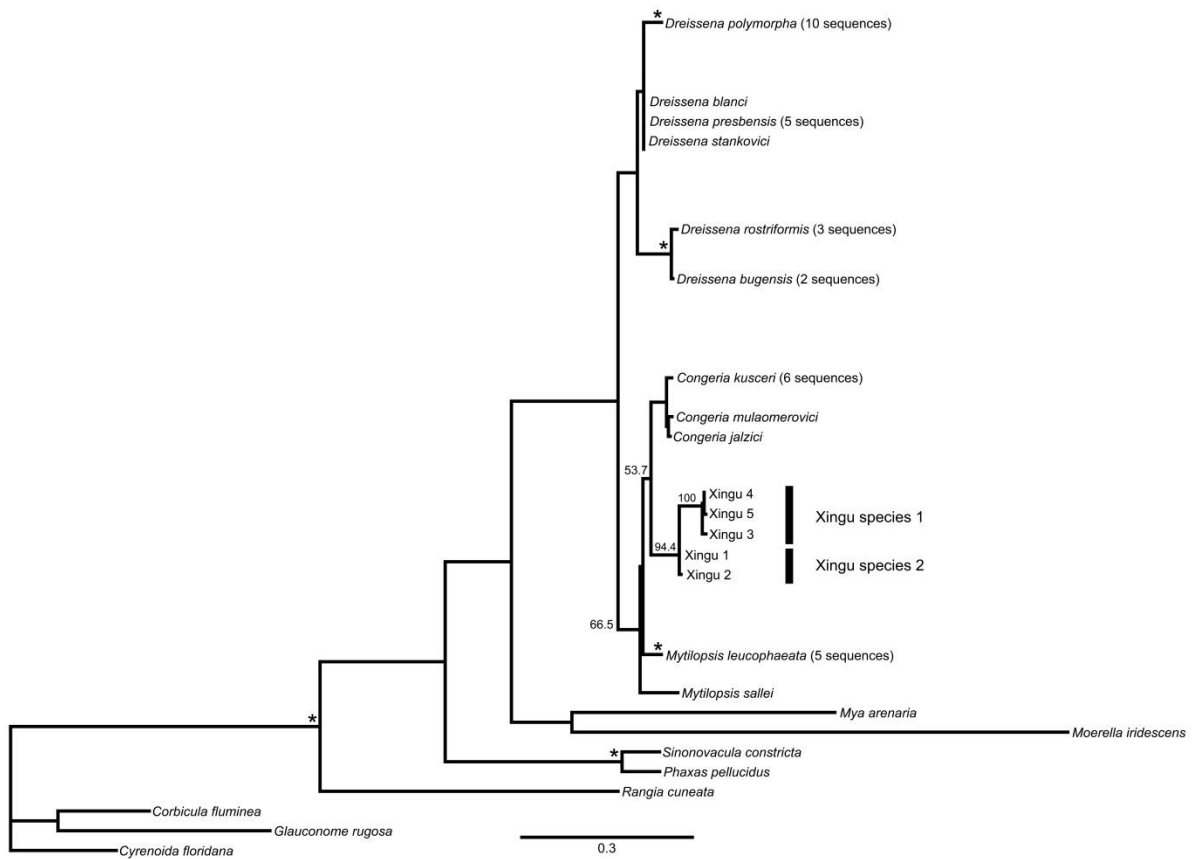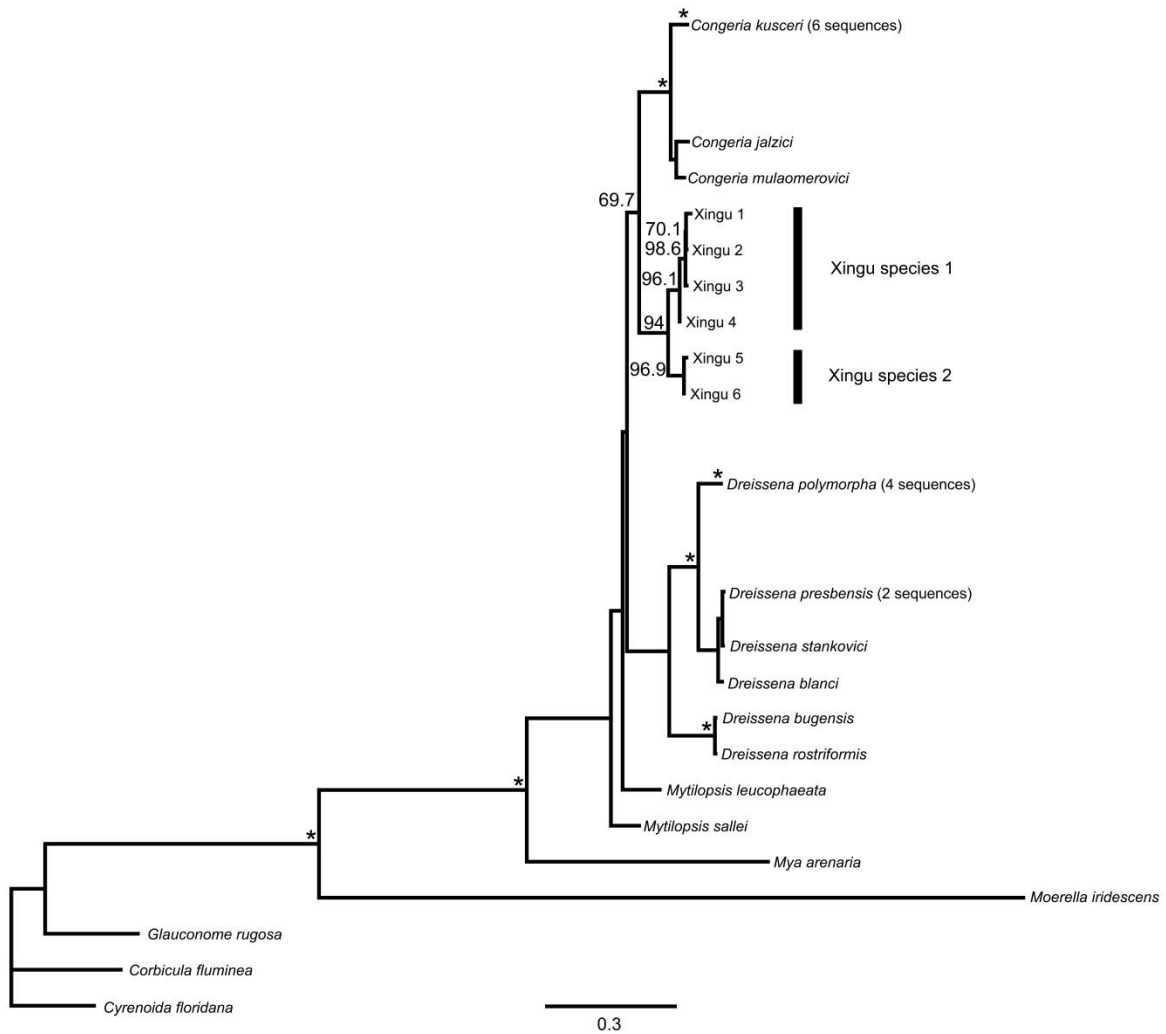
Figure 7. Maximum likelihood analysis of 16S dataset. An asterisk indicates bootstrap proportions of 95% and above. Node labels (bootstrap proportions) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.

Figure 8. Bayesian analysis of 16S dataset. An asterisk indicates posterior probabilities of 95% and above. Node labels (posterior probabilities) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.
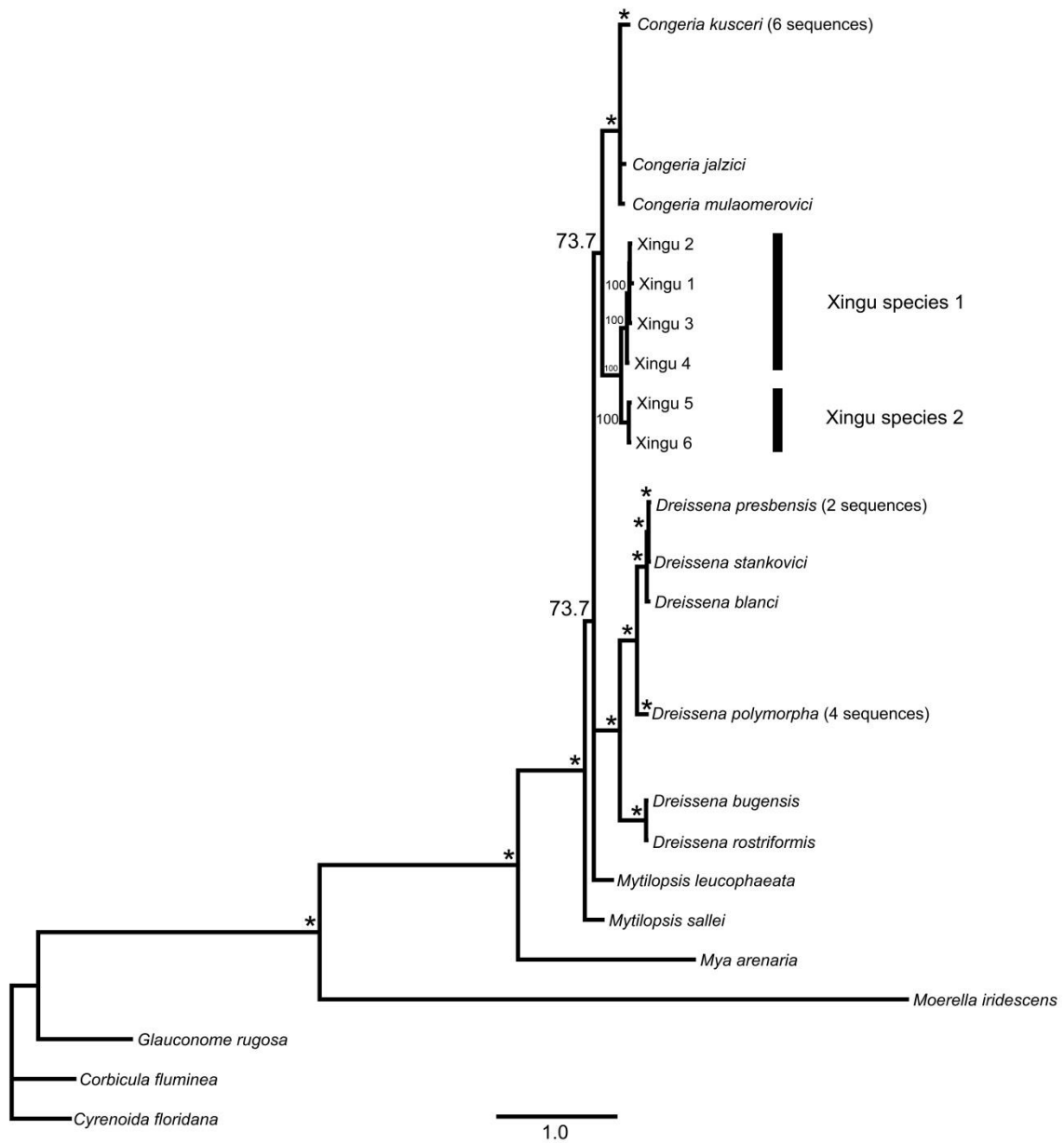
60

Figure 9. Maximum likelihood analysis of 18S dataset. An asterisk indicates bootstrap proportions of 95% and above. Node labels (bootstrap proportions) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.

Figure 10. Bayesian analysis of 18S dataset. An asterisk indicates posterior probabilities of 95% and above. Node labels (posterior probabilities) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.
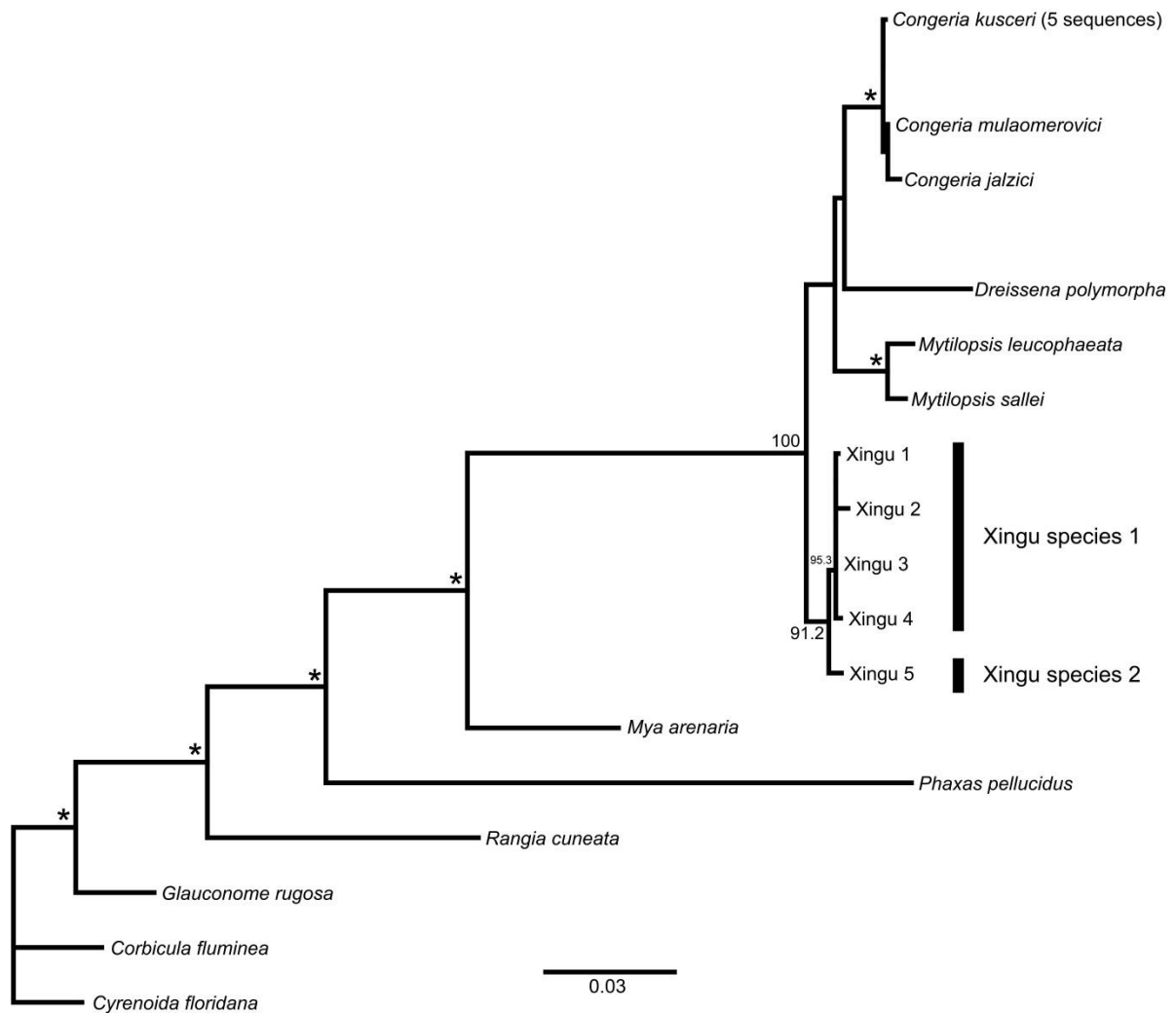
Figure 11. Maximum likelihood analysis of the mitochondrial dataset. This includes COI and 16S loci. An asterisk indicates bootstrap proportions of 95% and above. Node labels (bootstrap proportions) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.
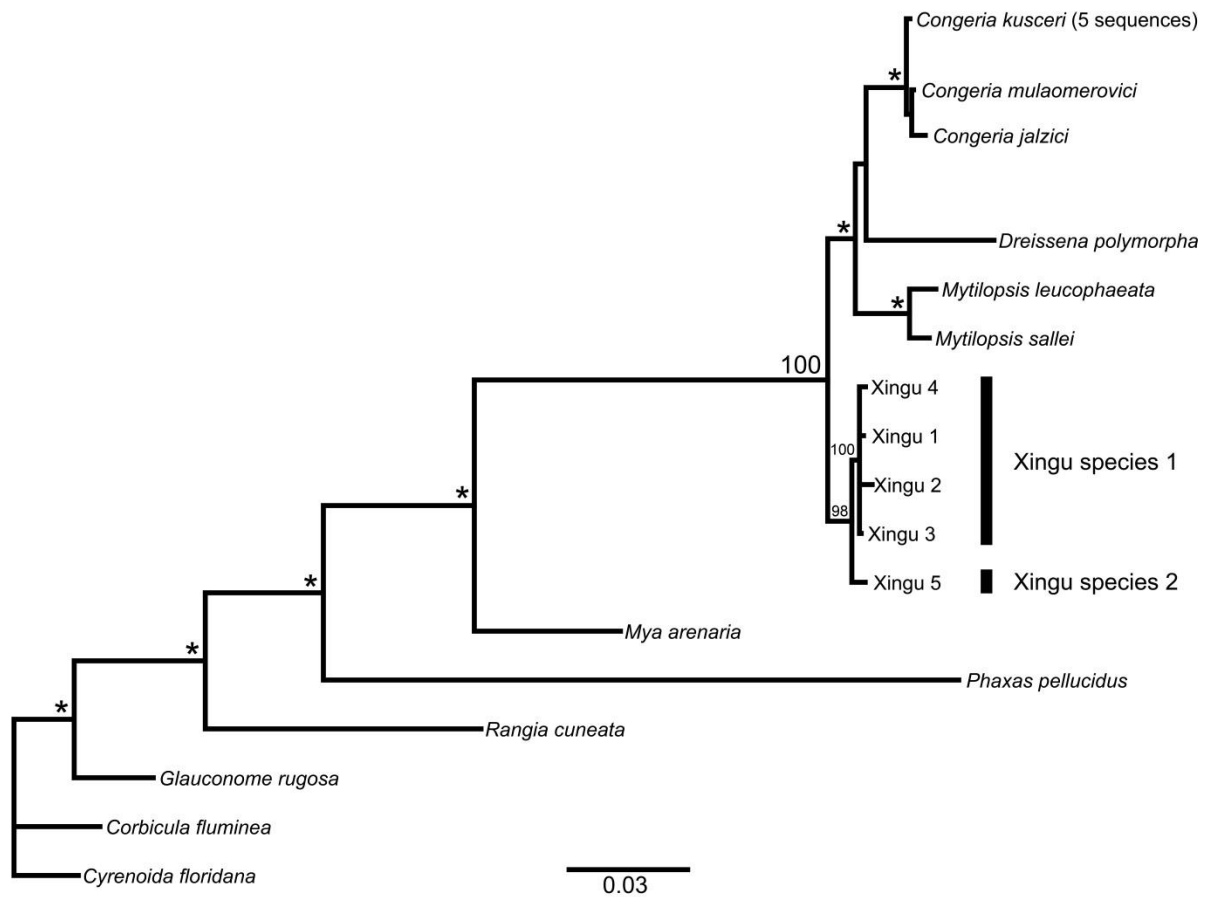
Figure 12. Bayesian analysis of the mitochondrial dataset. This includes COI and 16S loci. An asterisk indicates posterior probabilities of 95% and above. Node labels (posterior probabilities) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.
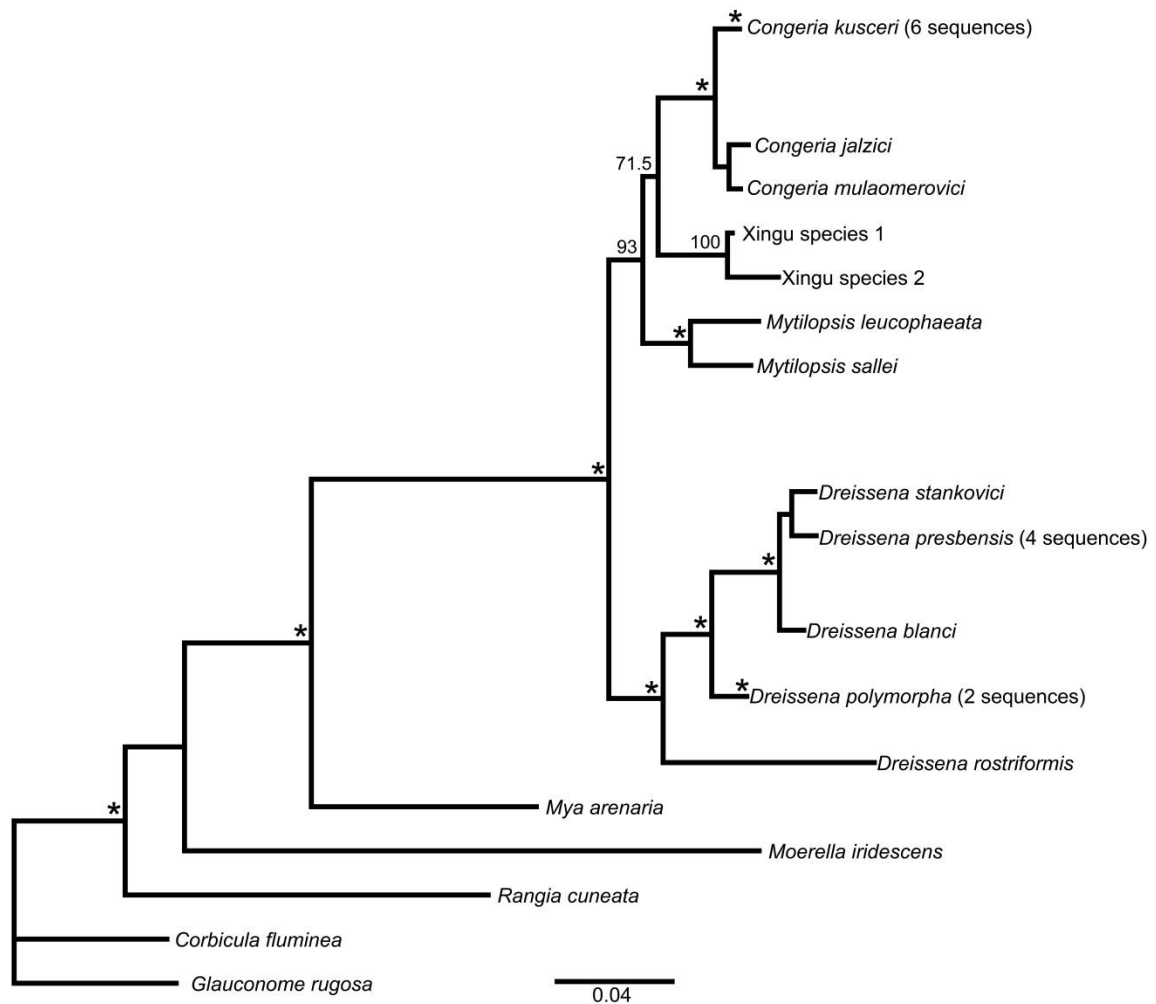
Figure 13. Maximum likelihood analysis of the nuclear dataset. This includes 28S and 18S loci. An asterisk indicates bootstrap proportions of 95% and above. Node labels (bootstrap proportions) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.

Figure 14. Bayesian analysis of the nuclear dataset. This includes 28S and 18S loci. An asterisk indicates posterior probabilities of 95% and above. Node labels (posterior probabilities) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.
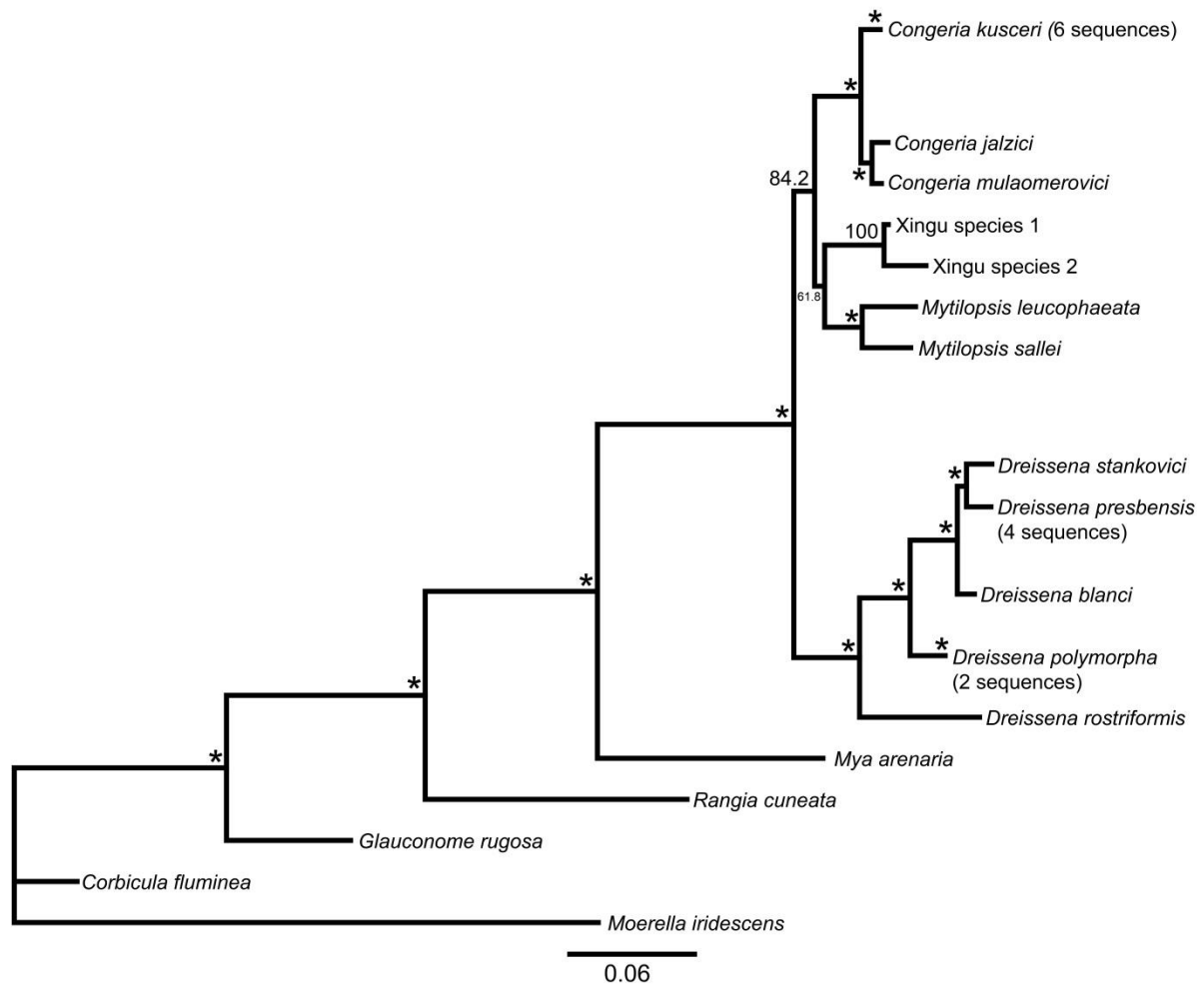
Figure 15. Maximum likelihood analysis of the concatenated dataset of all four genes. This includes all loci examined in this study. An asterisk indicates bootstrap proportions of 95% and above. Node labels (bootstrap proportions) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.

Figure 16. Bayesian analysis of the concatenated dataset of all four genes. This includes all loci examined in this study. An asterisk indicates posterior probabilities of 95% and above. Node labels (posterior probabilities) above 50% in the focal taxa were preserved for clarity. Scale bar represents nucleotide substitutions per site. Xingu haplotypes are in no particular order and are not geographically unique.
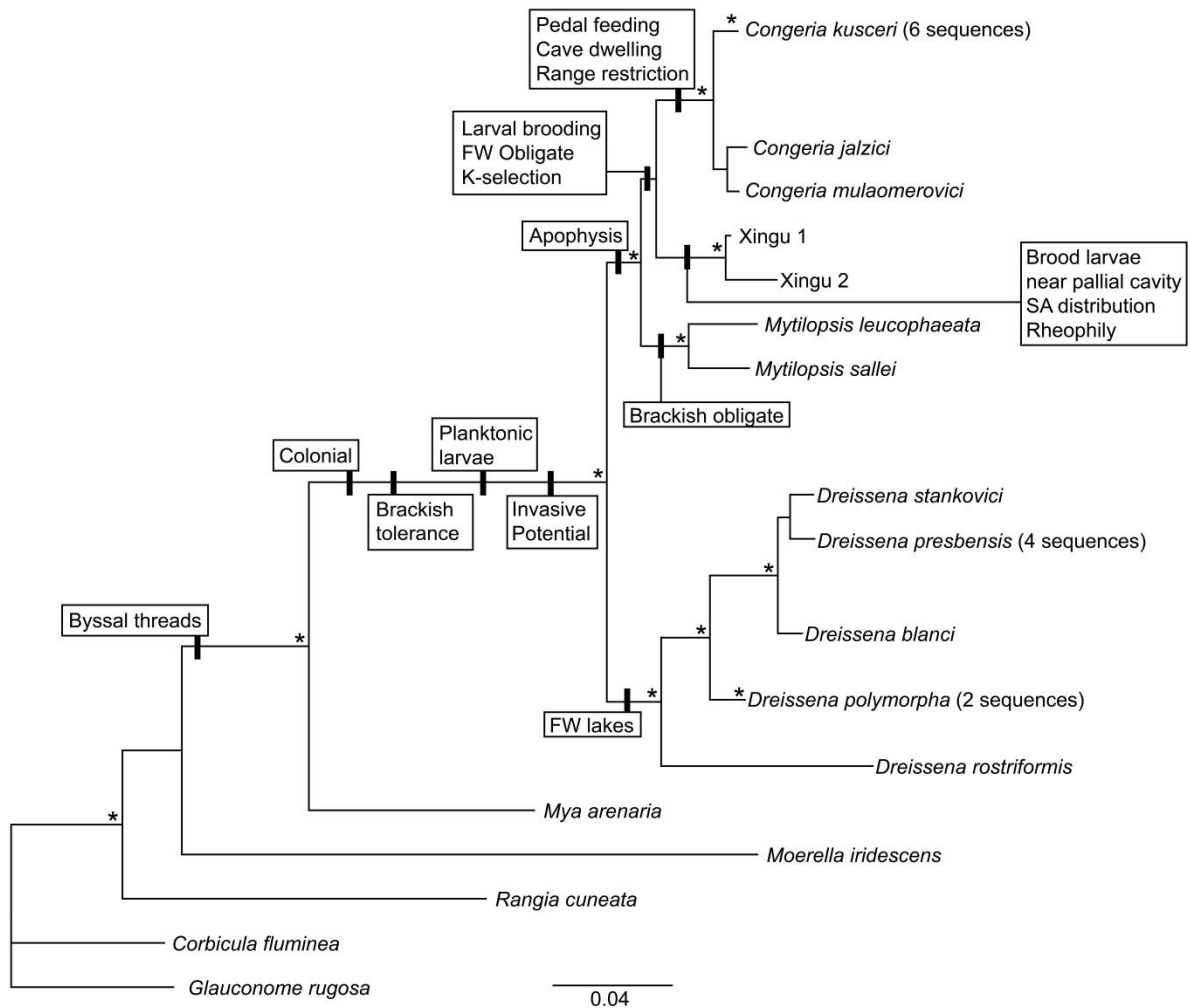
Figure 17. Comparative analysis with the concatenated (four gene) maximum likelihood tree. Character traits were mapped onto the phylogeny. Asterisks indicate bootstrap values above 95%. Traits are indicated by boxes. (Abbreviations: FW: freshwater; EU: European; SA: South American; WW: worldwide.)
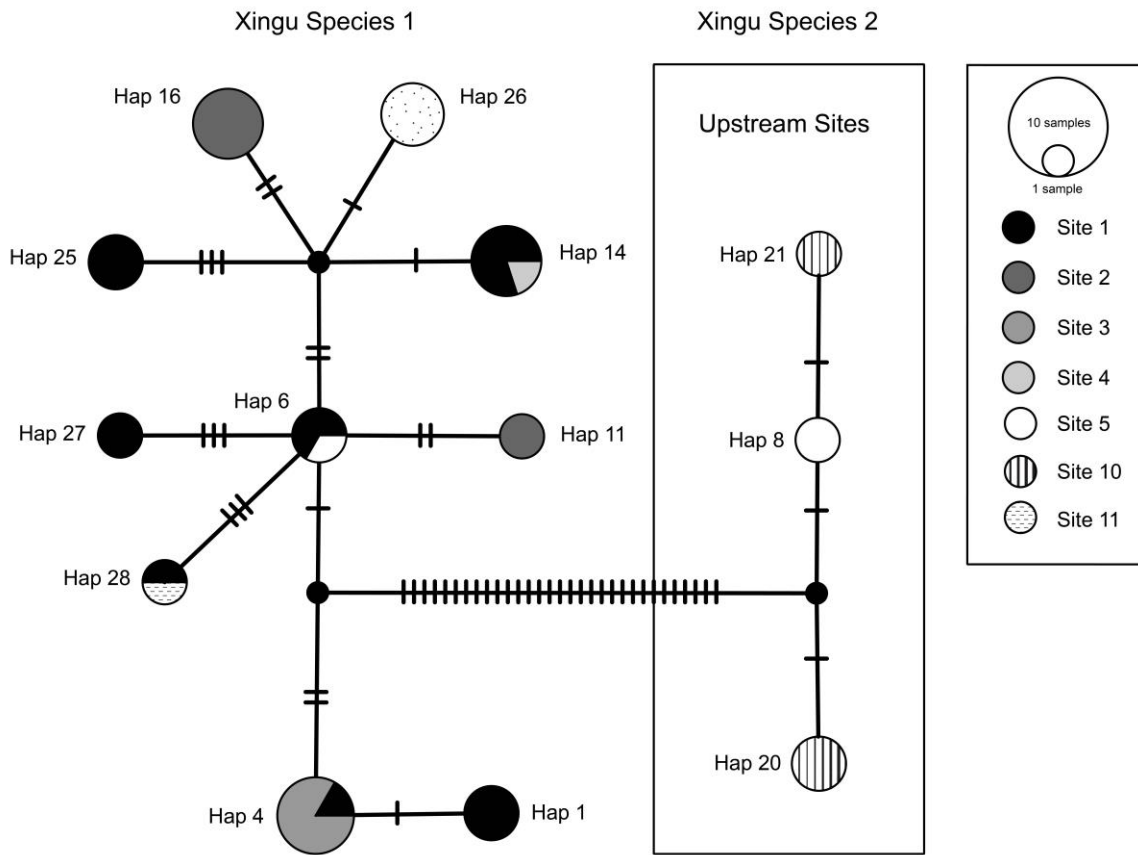
Figure 18. TCS site haplotype network of COI dataset. Circles represent haplotypes. Different shades and patterns represent collection site. Size of circles indicates how many individuals are represented in that haplotype. One dash indicates a one nucleotide difference. Small black circles represent an implied or unsampled haplotypes. Haplotypes with only one sequence were omitted from this analysis. Site numbers correspond to the sites in Table 2.
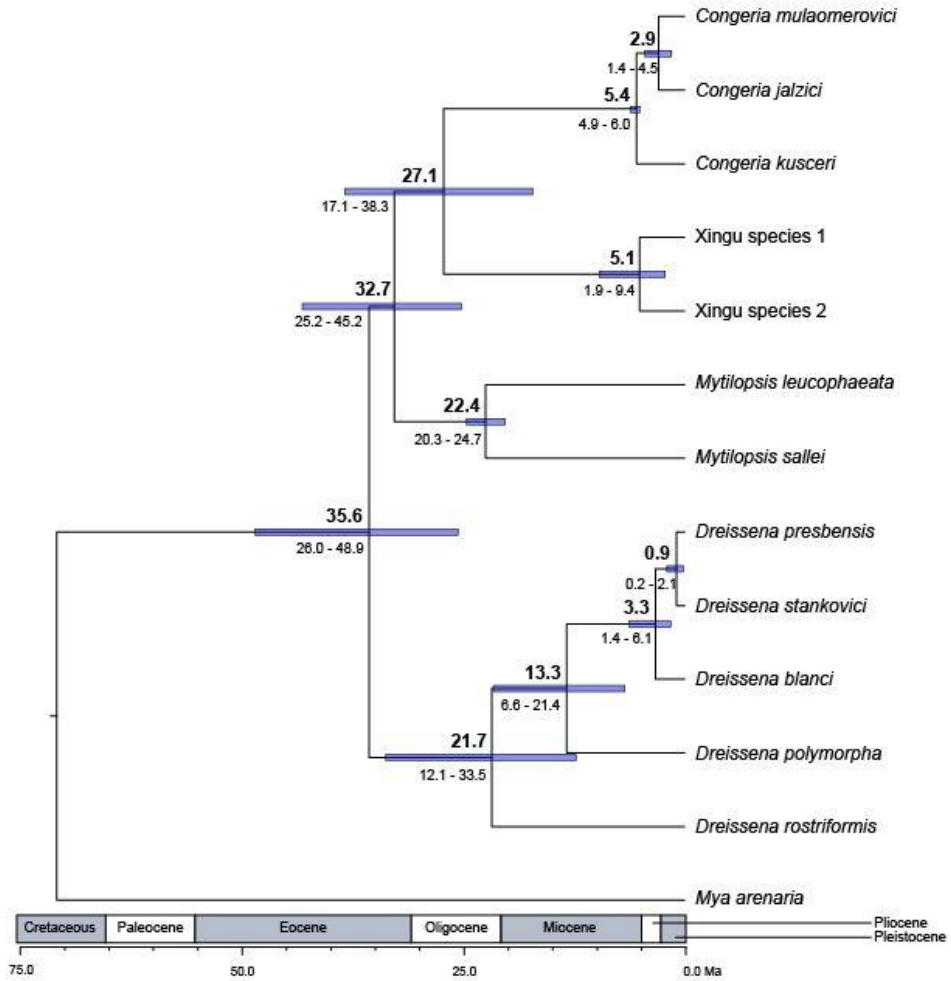
Figure 19. Maximum clade credibility chronogram created within BEAST v 1.8.4 (lognormal clock model) utilizing a concatenated alignment of all 4 genes amplified within this study (COI, 28S, 16S, and 18S). Mean divergence ages are shown above the nodes and the 95% highest posterior density intervals are denoted within parentheses below the nodes and shown by the blue horizontal bars. Geological time periods are shown on the scale bar in millions of years.

Appendix 1

Description of *Rheodreissena*, a new genus of South American freshwater bivalves (Dreissenidae).

Rheodreissena (Latin, f. "riverine dreissenid"), a new genus of freshwater mussel (family Dreissenidae). *Rheodreissena* includes at least three species, *Rheodreissena sp. novum* Ventuari, *Rheodreissena sp. novum* 1, Xingu*, Rheodreissena sp. novum* 2, Xingu

Rheodreissena, gen. nov. Gangloff et al. 2016

Type species: *Congeria hoeblichi* Schütt, 1991

Etymology

Rheodreissena is a portmanteau of the Greek word *rheos*, meaning river or stream, and the family name Dreissenidae, and is reference to the distinctively riverine, and often fast-flowing, habitats of this genus.

Diagnosis

Shell thin and very small, typically <15 mm in length. Shell subovate in outline with sharply angled and flattened ventral surface; beaks/umbos slightly inflated and strongly anterior, more acute than the other members of Dreissenidae; byssal threads originating at anterior-ventral surface. *Rheodreissena* is typically more

dorsally inflated and ventrally flattened than *Mytilopsis* and *Dreissena*. Compared to *Congeria*, *Rheodreissena* is anteriorly more rounded, and the greatest width of the shell (anterior view) is entirely ventral. *Congeria* has a more acute dorsolateral border than does *Rheodreissena*. Larger specimens of this genus have corrugations parallel to the posterior ridge, with one obvious corrugation halfway to two thirds down the shell. Smaller individuals lack this feature.

Rheodreissena is typically taller and shorter than *Mytilopsis, Dreissena,* and *Congeria.* Like members of *Dreissena*, individuals possess a small apophysis, although it is only observable through a macroscope. Other than this feature there is little structuring within the shell. Inner shells lack adductor scars and teeth. Periostracum color highly variable, from black to white, some specimens have dark stripes or spots. Individuals from the Xingu Drainage were largely grey, red, tan or black whereas individuals from the Ventuari exhibit more of a mottled brown and black color scheme. Nacre in all individuals is white to blueish-white. Preserved tissues white to brown in color.

Animals occur in small to moderately-sized aggregations, often on the undersides of interstitial cobbles (0.3 – 0.6 m) in laminar flow above or below large high-gradient shoals and rapids. Individuals occur on these rocks in a linear fashion where the rock and surrounding substrate meet. Some specimens are located in dimples in the rocks, providing crypsis. Aggregations are typically found 1-8 meters in depth, but have been observed at depths of 23 m. Larvae are brooded using a strategy that appears similar to other *Congeria* species. *Congeria* males release sperm into the water that is inhaled by females, resulting in internal fertilization.

Fertilized ova are held in the upper demibranches of females until they develop into embryos. Embryos are held in the marsupium until they are released as crawl away juveniles (Morton and Puljas, 2012). However, little is known concerning other life history characteristics.

Included species

Rheodreissena currently contains only one described species: *Rheodreissena hoeblichi* (Schütt, 1991), new combination. However, molecular evidence suggests the existence of at least two additional undescribed and reciprocally monophyletic species currently being referred to as: *Rheodreissena sp. novum* Ventuari, *Rheodreissena sp. novum* 1, Xingu*, Rheodreissena sp. novum* 2, Xingu.

Distribution

Clearwater rivers draining granitic basement rocks of the Brazilian and Guiana shields in northern South America. These rivers comprise several tributaries of the lower Amazon and upper Orinoco drainage basins, including the Iriri, Tapajos, Tocantins, and Xingu rivers in the Amazon, and the Ventuari River in the Orinoco.

Vita

Susan Rose Geda was born in Tampa, Florida to Mark and Linda Geda in 1990. She grew up in Winston-Salem, NC and her passion for aquatic ecology was rooted in the many trips her and her parents would take to numerous river systems across the Carolinas as a child. She received her Bachelor of Science degree in Ecology and Evolution from Appalachian State University in 2014 and started pursuing her Master of Science degree in August of 2015. The Master of Science degree was awarded in May 2017.