SHELL CONVERGENCE: AN INTERSPECIFIC MOLECULAR PHYLOGENY OF *NEOHELIX* (GASTROPODA: POLYGYRIDAE)

A Thesis
by
AMANDA C. WILKINSON

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

August 2020
Department of Biology

SHELL CONVERGENCE: AN INTERSPECIFIC MOLECULAR PHYLOGENY OF *NEOHELIX* (GASTROPODA: POLYGYRIDAE)

A Thesis
by
AMANDA C. WILKINSON
August 2020

APPROVED BY:

Matt C. Estep, Ph.D.
Assistant Professor, Chairperson, Thesis Committee

R. Wayne Van Devender, Ph.D.
Professor Emeritus, Thesis Committee

Amy S. Van Devender
Graduate Advisor, Thesis Committee

Shea R. Tuberty, Ph.D.
Professor and Assistant Chair for Student Affairs, Thesis Committee

Michael McKenzie, Ph.D.
Dean, Cratis D. Williams School of Graduate Studies

Abstract

SHELL CONVERGENCE: AN INTERSPECIFIC MOLECULAR PHYLOGENY OF *NEOHELIX* (GASTROPODA: POLYGYRIDAE)

Amanda C. Wilkinson
B.S., Appalachian State University
M.S., Appalachian State University

Chairperson: Matt C. Estep

Due to convergence in *Neohelix* von Ihering, 1892 and *Mesodon* Rafinesque in Férussac, 1821, species boundaries among taxa remain ambiguous. Prior hypotheses of polygyrid ancestry using morphological and behavioral characteristics by Ken Emberton have been compared to molecular phylogeny. However, a robust interspecific phylogenetic tree for *Neohelix*, a highly convergent genus of land snails in Polygyridae, has not yet been attempted. During this study, authors sequenced 28 specimens of *Neohelix* and outgroup species, testing four mitochondrial markers and two nuclear markers for specificity. Primers for the COI locus were modified to eliminate binding site polymorphism and increase amplicon length. The COI and 16S loci produced high node support for this interspecific phylogeny, but authors recommend testing 18S to expand nuclear support. Individuals were collected from known species ranges and identified using shell morphology but were not anatomically dissected to confirm species identity. Suspected misidentification among morphologically similar species like *Neohelix major, Neohelix albolabris,* and *Mesodon normalis* was present, resulting in a paraphyletic *Neohelix* clade. These findings confirm the necessity of reproductive dissections for identification and suggests that current hypotheses for species ranges need further investigation.

Acknowledgments

The completion of this research was supported by many mentors and faculty instrumental to my personal and academic success. I would like to personally thank my advisor Matt C. Estep who habitually collects lost undergraduate students to reveal the brilliance and stochasticity behind genetics. Without his guidance, encouragement, and advice, I and many others would not be where we are today. The inspirations behind my land snail fervor are R. Wayne and Amy S. Van Devender – experts who eagerly accepted my interest and bolstered my passion. Their guidance in all aspects of identification, collection, and preservation of land snails has perpetually enhanced my understanding of understory communities. I would also like to thank my advisor Shea R. Tuberty for providing an outstanding basis for my grasp of invertebrate zoology as well as his insightful edits and comments. Learning new techniques and conducting research is challenging, but the companionship and support from my lab family has been an invaluable resource and a beacon of light at the end of many hard days. Thank you to David, Ty, Ben, Patrick, Rachel, Logan, and Marietta for being my siblings in science. Lastly, I would like to thank my partner Nickolas A. Brand for his assistance with field collections and steadfast love.

This research would not have been possible without the support of Appalachian State University, which provided me with four Office of Student Research grants and a travel grant awarded by the Graduate Student Government Association. Thank you for supporting the pursuit of knowledge.

Dedication

The opportunity to conduct research and sate your hunger for the truth is a possibility that few are afforded, and even fewer pursue. This work is dedicated to those rare opportunists and pragmatic dreamers I call my friends, my greatest inspirations, and my biggest love.

Table of Contents

List of Tables

List of Figures

Foreword

During the dash toward advancement – that which is stimulated by new technologies, the drive of academia, and the thrill of unlocking hidden secrets – pausing to breathe humid earth and listen to woodland whispers has been the crux of my sanity. This thesis is the living representation of my fascination with the frequently overlooked land snail. Gentle and chiefly blind, these hermaphroditic detritivores are concealed in the understory of woody Appalachian giants. It is my hope that this quest to delineate evolutionary relationships among morphologically similar, silent forest caretakers is useful for expanding our understanding of their niche. Furthermore, I hope to encourage an appreciation for animals who harvest decay and cultivate beauty. I could not have accomplished this feat without the guidance of my mentors who supported and guided my passion for science, truth, and conservation.

This thesis will be submitted to The American Malacological Bulletin, an international peer-reviewed journal owned and published by The American Malacological Society. This thesis has been formatted according to the style guide for the AMB.

Boone, North Carolina

SHELL CONVERGENCE: AN INTERSPECIFIC PHYLOGENY OF *NEOHELIX* (GASTROPODA: POLYGYRIDAE)

INTRODUCTION

Study justification

The North American land snail family Polygyridae has been the subject of studies regarding sympatric convergence and distribution of low vagility taxa (Pilsbry 1940; Emberton 1988; Emberton 1991a-b; 1994a; 1995a-b; 1996). As an autochthonous, charismatic family of large, globose land snails, polygyrids are useful as models for iterated evolution, especially in tribes Triodopsini and Mesodontini which contain striking examples of shell convergence. Evolutionary parallelisms in taxa of these tribes are rampant (Emberton 1988; 1991a-b; 1994a-b; 1995a-b; 1996), primarily attributed to rapid Stylommatophoran radiation (Solem and Yochelson 1978; Solem 1981) and similar selective pressures (Solem 1978; 1985). *Neohelix* (Triodopsini) and *Mesodon* (Mesodontini) are two genera that are well documented to share convergence in sympatry, especially in *Neohelix major* (A. Binney) and *Mesodon normalis* (Pilsbry) (Emberton 1988; 1994a-b; 1995a-b; 1996; Perez *et al*. 2014). Reproductive dissections have been widely supported as the only definitive method for distinguishing convergent species (Pilsbry 1940; Solem 1976; Emberton 1988; 1991a-b; 1995a-b; 1996; Perez *et al.* 2014), as shell morphology and dentition may vary among populations (Tongkerd *et al*. 2004; Perez *et al*. 2014). While species boundaries for morphologically similar *Neohelix* taxa remain ill-defined, especially for subspecies (Emberton 1988; 1995a), hypotheses for ranges based on elevation and type locality have been presented (Pilsbry 1940; Hubricht 1985; Emberton 1988; Dourson 2012). The use of DNA barcoding for delimitation of species boundaries is ineffective in low vagility taxa (Bergsten *et al*. 2012; Perez *et al*. 2014). Therefore, the use of phylogenetic markers in addition to reproductive dissection is required for further study of evolutionary relationships in land snails. Due to limitations,

this study used morphological identification based on previously described species ranges to build an interspecific phylogeny of *Neohelix*. This research was conducted primarily to enhance understandings of evolutionary relationships in *Neohelix*, a genus of convergent, charismatic land snails that span the Southern Appalachians, Piedmont, and Ozarks. This study also hopes to highlight the necessity for development of further phylogenetic markers and delineation of species boundaries for polygyrid taxa.

The Significance of *Neohelix*

Neohelix species are of a capacious, depressed-globose shell shape with an imperforate umbilicus, having either a toothless aperture or a single parietal tooth (Pilsbry 1940). Although the Family Polygyridae is a North America endemic, *Neohelix* species are distributed from Eastern N.A. through the Midwest (Hubricht 1985). Seven species comprise the genus: *Neohelix dentifera*, *N. divesta*, *N. albolabris*, *N. major*, *N. alleni*, *N. solemi*, and *N. lioderma*. Morphological similarity due to convergence in sympatry (Hubricht 1985; Dourson 2012) are especially common in species *N. albolabris*, *N. major*, and *N. solemi*. Hubricht (1985) mapped species distributions for North American land snails based on his own experience and on museum specimens. The published range of *N. albolabris* extends from Northeast N.A. through Northern Louisiana and is not distributed further West than the Mississippi River (Hubricht 1985). Subspecies *N. albolabris bogani* (Emberton 1988), however, was described as extending West of the Mississippi from Texas through Arkansas and Oklahoma. The published range of *N. major* extends as far north as Maryland and is distributed through Northeast Mississippi (Hubricht 1985). However, *N. major* was later split into two species: *N. major* and *N. solemi* by Emberton in 1988. Distribution of *N. solemi* was described as ranging from Coastal Plain Maine to South Carolina (Emberton 1988). The taxonomic split of *N. major* into species *N. major* and *N. solemi* by Emberton in 1988 confounds previous *N. major* distributions (Emberton 1988). Additionally, *N. albolabris* has been recovered further southeast than Hubricht's 1985 map previously suggested (Dourson 2013). Homoplasy and ambiguous species ranges among

Neohelix, especially *N. albolabris*, *N. major*, and *N. solemi*, make *Neohelix* an appropriate candidate for phylogenetic study.

Further motivation for studying *Neohelix* was the presence of rare species *Neohelix lioderma*. In 1940, Pilsbry described *N. lioderma* as having a somewhat translucent, glossy, imperforate shell with weaker striation than the adjacent *N. divesta* (Pilsbry 1940). This species has only been reported from Rogers and Tulsa counties in Oklahoma. In the Oklahoma Comprehensive Wildlife Conservation Strategy guide, the Oklahoma Department of Wildlife Conservation reports *N. lioderma* occurrences within the following regions: Cross Timbers and Tallgrass Prairie. Cross Timbers consists of the central one-third of Oklahoma, composed primarily of oak woodlands and prairies, while the Tallgrass Prairie region incorporates the Osage Plain and Flint Hills ecoregions (ODWC 2015). Inclusion of this species in the phylogenetic interpretation of *Neohelix* interspecific relationships was prioritized, as no previous genetic analysis has been conducted for *N. lioderma*.

DNA Markers

Recent phylogenetic work in Polygyridae (Perez *et al.* 2014) has uncovered discrepancies in the monophyly of Triodopsini and *Neohelix* as reported by Emberton 1988; 1994a; 1995a based on reproductive anatomy. Further genetic investigation of *Neohelix* and triodopsine relationships as well as geographically distinct subspecies like *N. alleni alleni* and *N. alleni fuscolabris* should be conducted to test previous relationship hypotheses. Although universal markers are useful for delineating higher taxonomic relationships, interspecific phylogenetic interpretation may require the introduction of family-specific or microsatellite markers. Before advances in marker development proceed, the efficacy of current phylogenetic markers for building a well-supported interspecific *Neohelix* phylogeny should be investigated. To verify relationships among *Neohelix* species, this study utilized common universal mitochondrial and nuclear markers as well as three related outgroup taxa to root the *Neohelix* and Triodopsini clades.

MATERIALS AND METHODS

Taxon Sampling

Tail clippings from land snail specimens were sampled from catalogued individuals at the North Carolina Museum of Natural Sciences (NCMS) and Florida Museum of Natural History (UF) or were field collected and morphologically identified by Amy S. Van Devender and R. Wayne Van Devender. Collections permits included a Wildlife Collection License through the Missouri Department of Conservation (#18168), the Arkansas Game & Fish Commission (#031820191), the North Carolina Wildlife Resources Commission (#19-SC00091), and the Blue Ridge Parkway (#BLRI-2019-SC1-0017). Individuals collected from Oklahoma were contributed by correspondent Alex C. Cooper under the Oklahoma Department of Wildlife Collection permit #7461. In total, 29 specimens were successfully sequenced (Table 1). Individuals collected from the field were preserved in 95% EtOH and donated to the NCMS collection after 10-20 mg of foot tissue was harvested. Polygyrid outgroups *Mesodon normalis* (Pilsbry 1900), *Patera perigrapta* (Pilsbry 1894), and *Xolotrema fosteri* (F.C. Baker 1932) were collected. A minimum of two individuals per species of *Neohelix* were collected, excluding *Neohelix lioderma*, from which one tissue sample was collected (Table 1). Duplicate species were sampled from geographically distinct locations to minimize genetic bias resulting from closely related populations (Figure 1). Exact coordinates of collected individuals will not be published to inhibit illegal poaching of land snails but are available from authors on request.

Table 1. Samples sequenced for phylogenetic analysis. Includes GenBank accession numbers per locus, locality (County, State), field collection numbers, and museum accession numbers. Absence of GenBank number indicates sequence failure for that locus. Dash in place of museum deposition number indicates a sample which has yet to be deposited in NCMS collections.

Species	**GenBank #'s**				**Locality**	**Field**	**Museum #**
	COI	16S	H3	28S	**(County, State)**	**Collection #**	
Neohelix albolabris	-	MT254102	MT266961	MT254113	Harlan, KY	ASV_2007-098	NCMS 64366 A
	MT252628	MT254101	MT266962	MT254114	Watauga, NC	ASV_2008-049	NCMS 64039
	-	-	MT266963	MT254115	Giles, VA	ASV_2010-068	NCMS 44717
	MT252631	MT254099	MT266966	MT254118	Alexander, NC	ASV_2015-171	NCMS 64495 B
	-	MT254098	MT266967	MT254119	Surry, NC	ASV_2017-066	NCMS 64846 B
	-	-	MT266970	MT254122	Polk, TN	ASV_2018-036	-
Neohelix dentifera	-	-	MT266958	-	Watauga, NC	ASV_2005-042	NCMS 41422 B
	-	-	MT266960	MT254112	Ashe, NC	ASV_2007-091	NCMS 64359
	MT252634	MT254095	MT266972	MT254124	Bedford, VA	ASV_2018-067	-
	MT252635	-	MT266973	MT254125	Botetourt, VA	ASV_2018-069	-
Neohelix solemi	MT252629	MT254100	MT266964	MT254116	Moore, NC	ASV_2013-040	NCMS 64109 A
	MT252630	-	MT266965	MT254117	Beaufort, NC	ASV_2014-072	NCMS 100472 B
	MT252632	MT254097	MT266968	MT254120	Moore, NC	ASV_2017-076	NCMS 65051 B
	MT252638	MT254093	MT266976	MT254128	Scotland, NC	ASV_2019-008	-
Neohelix major	-	-	MT266982	MT254134	Winston, AL	INVERT_06940	NCMS 35102
	-	-	MT266956	MT254109	Pickens, SC	ASV_2005-010	NCMS 41467
	MT252626	MT254104	MT266957	MT254110	Jackson, NC	ASV_2005-016	NCMS 41446
	MT252627	MT254103	MT266959	MT254111	Stanly, NC	ASV_2016-061	NCMS 65102
	MT252633	-	MT266969	MT254121	Walker, GA	ASV_2017-086	NCMS 64998
Neohelix alleni							-
	MT252639	MT254092	MT266977	MT254129	Madison, AL	ASV_2019-039	-
	MT252641	MT254090	MT266979	MT254131	Pushmataha, OK	ASV_2019-078	-
	-	MT254089	MT266980	MT254132	Washington, AR	ASV_2019-086	-
							-

Species	GenBank #'s				Locality (County, State)	Field Collection #	Museum #
	COI	16S	H3	28S			
Neohelix divesta	MT252640	MT254091	MT266978	MT254130	Garland, AR	ASV_2019-046	-
	-	MT254106	MT266954	MT254107	Benton, AR	ACW_2019-001	-
Neohelix lioderma	MT252642	MT254088	MT266981	MT254133	Tulsa, OK	Coles-A188	- UF 449948
Xolotrema fosteri	MT252637	-	MT266975	MT254127	Lyon, KS	ASV_2018-081	- -
Patera perigrapta	MT252625	MT254105	MT266955	MT254108	Benton, AR	ACW_2019-002	- -
Mesodon normalis	MT252636	MT254094	MT266974	MT254126	Botetourt, VA	ASV_2018-075	-
	-	MT254096	MT266971	MT254123	Transylvania, NC	ASV_2018-037	-

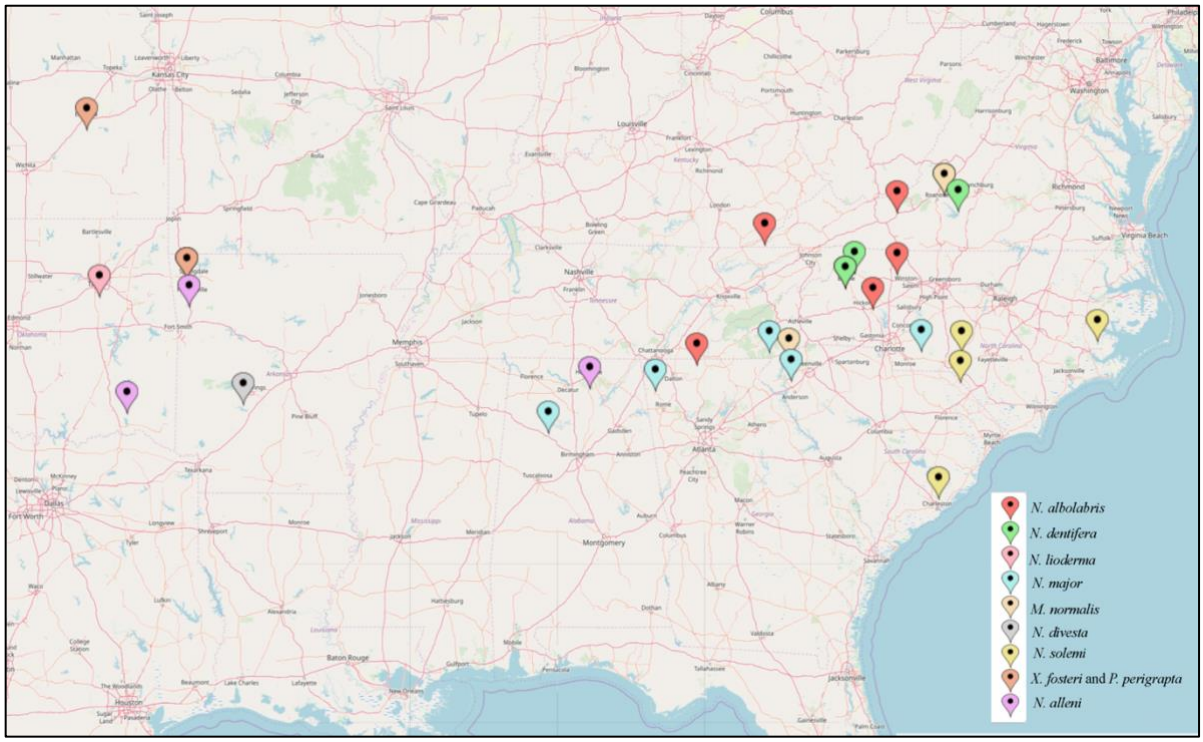

Figure 1. Geographic distribution of sampled individuals. Locality is approximate position of county marker, not the exact coordinates. Total number of sampled individuals was 28, but 26 markers are displayed due to two overlapping counties for different species. Key to specimen ID is located in bottom right corner:

Note. Red = *N. albolabris*; Green = *N. dentifera*; Pink = *N. lioderma*; Blue = *N. major*; Tan = *M. normalis*; Gray = *N. divesta*; Yellow = *N. solemi*; Orange = *X. fosteri* and *P. perigrapta* outgroups; Purple = *N. alleni*

DNA Extraction

Samples were stored in 80% EtOH at –20ºC before DNA extraction using the Invitrogen PureLink™ Genomic DNA Mini Kit following manufacturer protocol (Thermo Fisher Scientific, USA). Tissue was patted dry with sterile ChemWipes to remove excess EtOH, then weighed to determine mass in milligrams (mg). Approximately 2-10 mg of tissue was soaked in 1xTE (10 mM Tris, 0.5 mM EDTA, pH 8.0) buffer for 30 minutes to diffuse excess EtOH, then digested overnight using a Proteinase K (20 mg/mL) master mix. To improve quality and concentration of extracted DNA from museum specimens predating 2000, length of tissue diffusion in 1xTE buffer was extended from 30 minutes to 1 hour. Digestion length was increased to ≥15 hours as opposed to ≥8

hours. Once cells were lysed, DNA was extracted and purified from the surrounding cell contents. Extracted DNA was analyzed for purity and concentration using a ND-1000 Nanodrop Spectrophotometer (Thermo Fisher Scientific, USA) and 1% agarose Gel Electrophoresis. DNA samples were discarded and re-extracted if Nanodrop absorption ratios deviated significantly from 1.8 for 260/280, from 2.0 for 260/230, or if the concentration was less than 50 ng/µL. Concentration and purity was confirmed using 1% agarose Gel Electrophoresis. All DNA samples that passed quality control measures were diluted to 30 ng/µL for downstream applications.

PCR Amplification

Published primer pairs (Table 2) were used to amplify four mitochondrial markers: Cytochrome oxidase I (COI) using primers COIL1490 and COIH2198 (Folmer *et al.* 1994), cytochrome b (cytb) using primers Ucob151F and Ucob270R (Merritt *et al.* 1998), small ribosomal subunit 12S using primers 12Sai and 12Sbi (Simon *et al.* 1994), and large ribosomal subunit 16S using primers 16Sar and 16Sbr (Palumbi *et al.* 1991). Two nuclear markers were also amplified: Large ribosomal subunit 28S using primers VI and X (Hillis *et al.* 1996), and Histone 3 (H3) (Hillis *et al.* 1996) using primers H3F and H3R (Hillis *et al.* 1996).

Table 2. Nucleotide sequences of utilized phylogenetic markers and associated information.

Locus	Primer	Sequence (5' – 3')	Reference
COI	COIL1490	GGTCAACAAATCATAAAGATATTGG	Folmer *et al.* 1994
	COIH2198	TAAACTTCAGGGTGACCAAAAAATCA	
Cytb	Ucob151F	TGTGGRGCNACYGTWATYACTAA	Merritt *et al.* 1998
	Ucob270R	AANAGGAARTAYCAYTCNGGYTG	
12S	12Sai	AAACTAGGATTAGATACCCTATTAT	Kocher *et al.* 1989
	12Sbi	AAGAGCGACGGGCGATGTGT	
16S	16Sar	CGCCTGTTTAHYAAAAACAT	Palumbi *et al.* 1991
	16Sbr	CCGGTCTGAACTCAGMTCAYGT	
28S	VI	AAGGTAGCCAAATGCCTCATC	Hillis *et al.* 1996
	X	GTGAATTCTGCTTCATCAATGTAGGAAGAGCC	
H3	H3F	ATGGCTCGTACCAAGCAGACVGC	Colgan *et al.* 1998
	H3R	ATATCCTTRGGCATRATRGTGAC	

PCR reactions were prepared in 10 µL using dH_2O, 5X Colorless GoTaq Flexi Buffer, 2.5 mM $MgCl_2$, 800 µM dNTP's, 0.5 µM of right and left primers, 0.5 units of GoTaq DNA Polymerase, and ~30 ng of template DNA. Five of the loci (cytb, 12S, 16S, 28S, and H3) were amplified using thermocycler conditions from Hamstead *et al*. (2015).

COI primers COIL1490 and COIH2198 produced absent or nonspecific amplification under Folmer *et al*. (1994) and Hamstead *et al.* (2015) thermocycling conditions. Other published thermocycling conditions for COIL 1490 and COIH2198 were tested, including Thaewnon-ngiw *et al.* (2004) and Campbell *et al*. (2005) (Table 3).

Table 3. PCR Amplification Conditions for COIL1490-COIH2198.

Locus	Reaction Condition	Reference
COI	94°C for 3 min, 10 x (94°C for 1 min, 45°C for 1 min, 72°C for 1 min), 35 x (94°C for 1 min, 53°C for 1 min, and 72°C for 1 min), 72°C for 7 min. 94°C for 5 min. Hold at 10°C.	Thaewnon-ngiw *et al*. 2004
COI	92°C for 2 min, 5 x (92°C for 40 sec, 40°C for 40 sec, 72°C for 90 sec), 25 x (92°C for 40 sec, 50°C for 40 sec, 72°C for 90 sec), 72°C for 10 min. Hold at 10°C.	Campbell *et al.* 2005
Cytb, 12S, 16S, 28S, H3	13 x (94°C for 45 sec, 68°C for 2 min with -0.5°C per cycle, then 72°C for 1 min), 25 x (94°C for 45 sec, 53°C for 1 min, 72°C for 1 min), 72°C for 10 min. Hold at 10°C.	Hamstead *et al.* 2015

PCR reaction conditions were modified for the COI primers, including adjustment of $MgCl_2$, template, and Taq Polymerase concentrations following guidelines by Palumbi *et al*. 1991. No PCR modification of COI primers resulted in robust amplification across all *Neohelix* species. Alternative published COI primers for mollusks (Table 4) were evaluated for amplification improvement in *Neohelix* using standard published conditions as well as those from Table 3.

Table 4. Published Primers for COI Locus.

Primer	Sequence (5'-3')	T_m (°C)	Length	Reference
LCO1490	GGTCAACAAATCATAAAGATATTGG	56	25	Folmer *et al.* 1994
HCO2198	TAAACTTCAGGGTGACCAAAAAATCA	61	26	Folmer *et al.* 1994
COIL	GGTCAACAAATCATAAAGATATTGG	56	25	Dayrat *et al.* 2011
COIH	TAAACTTCAGGGTGACCAAARAAYCA	61 - 64	26	Dayrat *et al.* 2011
COI 14F	WYTCNACDAAYCAYAAAGAYATTGG	51 - 66	25	Dayrat *et al.* 2011
COI 698R	TADACYTCNGGRTGHCCRAARAAYCA	58 - 75	26	Dayrat *et al.* 2011
COI 839R	AAYATRTGHGCYCANACAATAAAWCC	54 - 67	26	Dayrat *et al.* 2011
LCOI	GGTCAACAAATCATAAAGATATTGG	56	25	Klussmann-Kolb *et al.* 2008
HCOI	TAAACTTCAGGGTGACCAAAAAATCA	61	26	Klussmann-Kolb *et al.* 2008
COI long f	GGTCAACAAATCATAAAGATATTGG	56	25	Klussmann-Kolb *et al.* 2008
COI long r	TAAAGAAAGAACATAATGAAAATG	50	24	Klussmann-Kolb *et al.* 2008
JB3	TTTTTTGGGCATCCTGAGGTTTAT	63	23	Bowles et al. 1992
JB4.5	TAAAGAAAGAACATAATGAAAATG	52	24	Bowles et al. 1992

Due to the failure of previous modifications to improve COI amplification, LCO1490 and HCO2198 were tested for annealing specificity to *Neohelix* template sequences. Partial *Neohelix* COI sequences were amplified using LCO1490 and HCO2198 with Thaewnon-ngiw *et al.* 2004 thermocycler conditions and Sanger sequenced at Georgia Genomics and Bioinformatics Core.

In order to develop new primers for COI, a nucleotide BLAST of non-redundant databases on NCBI was performed using limited available sequence data. Two whole mitochondrial genome references were identified for *Praticolella mexicana* (Polygyridae: Polygyrinae) (KX240084.1) and *Polygyra cereolus* (Polygyridae: Polygyrinae) (KX278421.1). Published COI primers (Table 4) were aligned to both reference sequences using Blast2Seq to test for specificity and identify primer site polymorphism (Figure 2).

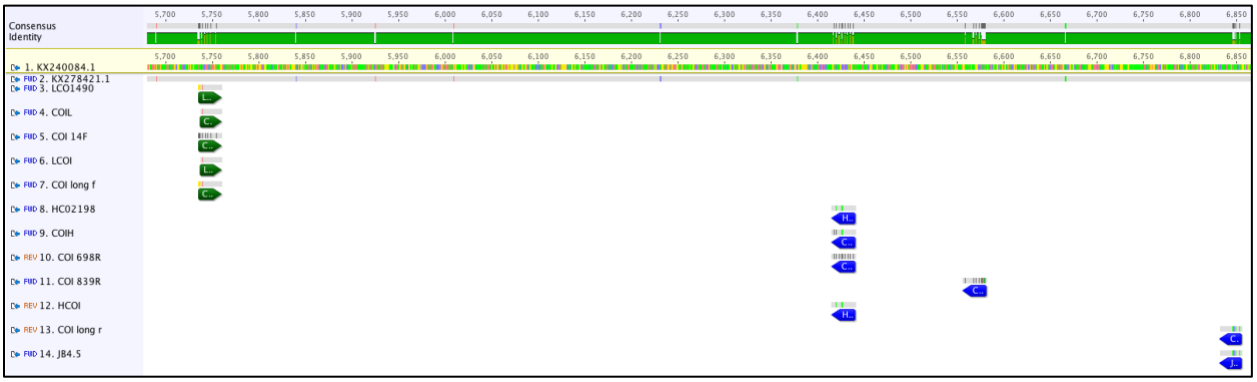

Figure 2. COI published primers aligned to KX240084.1 (*Praticolella mexicana*) and KX278421.1 (*Polygyra cereolus*) reference sequences. Dark green annotations represent published forward primers and dark blue annotations signify published reverse primers.

Published COI primers were modified to eliminate primer site polymorphism, and custom COI primers were produced using Primer3 (Rozen and Skaletsky 2000). Parameters included an optimal primer length of 20 bp, T_m of 60°C, and product range of 300-800 bp. Naming conventions for primers utilized the prefix "poly-" to denote relevance to Family Polygyridae. Modified primers were named in reference to the original published primers, including the location of the 5' end. Custom primers referenced the distance between the 5' ends of primers polyLCO-F5735 or polyHCO-R6441, the modified Folmer *et al*. 1994 COI primers. Modified and custom COI primers from Table 5 were aligned to Polygyridae reference sequences (KX240084.1 and KX278421.1) and mapped.

DNA Sequencing

Products successfully amplified from *Neohelix* and outgroup species were analyzed via 1% agarose gel and compared to a 1 kbp ladder (Thermo Fisher Scientific US). Products were manually excised from the gel and purified via DNA precipitation. Sequencing templates were diluted to 30 ng/µL and 1 µL then arrayed onto a 96-well plate. To map the sequence array, samples were assigned a number and randomly distributed within the same locus of the 96-well plate in separate forward and

reverse reads. Primer dilutions (3.3 μM) for each locus were added to the corresponding wells. Plates were Sanger sequenced at Georgia Genomics and Bioinformatics Core.

Sequence Analysis

Raw chromatograms were scored using Geneious version 9.0.5. Single read amplicons were trimmed to eliminate ambiguous ends, then contiged with their complimentary read and locally aligned using polygyrid reference sequences from GenBank. Protein coding loci H3 and COI were translated to amino acid sequences and aligned. Noncoding loci 16S and 28S were aligned according to nucleotide identity. Biologically improbable hypervariable regions in 16S were eliminated from tree-building analyses. Alignments used for this study are available from authors on request. JModelTest2 (Darriba *et al*. 2012; Guindon and Gascuel 2003) was used to generate deltaBIC scores based on each single locus alignment to determine the best nucleotide substitution model for tree building. Single gene trees were constructed under the appropriate substitution model using Maximum Parsimony (MP), Maximum likelihood (ML), and Bayesian Inference (BI). Single locus MP trees were constructed using PAUP* version 4.0a (build 167) (Swofford 2002). Maximum Likelihood single locus trees were constructed in RAxML-NG (Kozlov *et al.* 2019), while BI trees were constructed using MrBayes v3.2.6 (Huelsenbeck and Ronquist 2001). After single locus trees were topologically compared, a consensus tree was constructed from concatenated sequences and analyzed to seek a robust phylogeny for *Neohelix*. *Neohelix* phylogenetic relationships were estimated based on well supported nodes.

RESULTS

Taxon Sampling

During this study, a total of 128 specimens were collected from the field and/or museum repositories, but only 28 were used for sequence analysis (Table 1) Individuals were identified based on shell morphology, as no reproductive dissections were conducted. Specimens chosen for

phylogenetic analysis were selected for their proximity to known species ranges (Figure 1) and best approximations of species identity based on shell morphology.

DNA Extraction

Extraction of DNA using ≤5 mg of tissue yielded higher purity and greater concentration than tissue quantities ≥6-10 mg. Museum specimens collected prior to 1990 typically produced high 260/230 ratios and low DNA concentrations (10-30 ng/µL). Increased diffusion time in 1xTE for 1 hour, as opposed to 30 minutes, prior to digestion significantly improved concentration (50-60 ng/µL). The Invitrogen PureLink™ Genomic DNA Mini Kit was effective for consistent high-quality extraction of DNA.

PCR Amplification

Nuclear loci H3 and 28S produced clear amplification across *Neohelix* and outgroup samples using standard primers and thermocycler conditions. Nonspecific amplification in 16S for *Neohelix* species was rectified by applying Hamstead *et al.* (2015) thermocycler conditions, as opposed to standard 16Sar and 16Sbr conditions. Manual excision and purification of bands with ideal base size (~550 bp) for the 16S locus ensured sequence accuracy.

The amplification of COI in *Neohelix* species as well as *Mesodon normalis* and *Xolotrema fosteri* using primers LCO1490 and HCO2198 (Folmer *et al.* 1994) failed under standard conditions. Alternate published primers (Table 4) for pulmonate gastropods were tested with inconsistent results. Attempts to optimize thermocycler conditions (Table 3), $MgCl_2$ concentration, and template concentration failed to improve amplification. The most successful amplification of taxa was achieved using Thaewnon-ngiw *et al*. (2004) thermocycler conditions with LCO1490 and HCO2198 primers. Species amplified include: *Neohelix albolabris*, *Neohelix alleni*, *Neohelix divesta*, *Neohelix solemi*, and *Xolotrema fosteri*. Amplification was nonspecific for *N. dentifera* and failed for *N. major* (Figure 3). The remaining *Neohelix* species and closely related outgroup taxa produced nonspecific or absent amplification under alternate thermocycler conditions (Table 3).

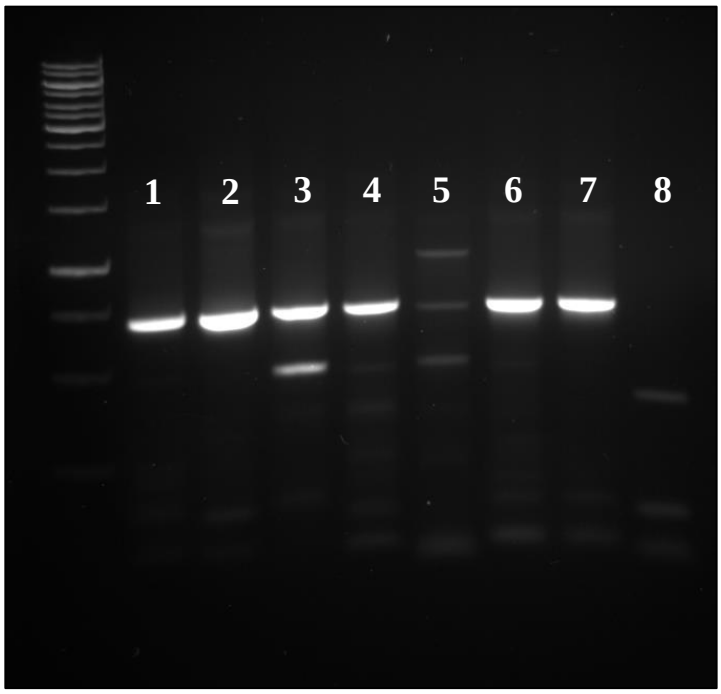

Figure 3. LCO1490 and HCO2198 (Folmer *et al*. 1994) amplification of *Neohelix* spp. and *X. fosteri* outgroup under Thaewnon-ngiw *et al*. 2004 thermal cycling conditions.

Note. Ladder is 1 kb, identity of lanes (from left to right) are (1) *Neohelix albolabris*, (2) *N. alleni*, (3) *N. dentifera*, (4) *N. divesta*, (5) *N. major*, (6) *N. solemi*, (7) *X. fosteri*, and (8) negative control.

Melting temperatures (T_m) for published COI primers (Table 4) were calculated using published concentrations, but many fell outside the standard 55-60$_o$C annealing range (Palumbi *et al.* 1991). Twelve primers aligned to Polygyridae reference sequences and were mapped in Figure 2. Forward primer JB3 was not specific to either template. Three primers: LCO1490, COI 14F, and COI long f aligned at the same position (5735 – 5759 bp) on KX240084.1 and KX278421.1 reference sequences. Primers COIL and LCOI were two bases shorter on the 5’ end, ranging from 5737 – 5759 bp. Magnified primer position and polymorphic sites are included in Figure 10. Primer LCO1490 was identical to COI long f, with both having 3 nucleotide polymorphisms on the 5’ end: Two transversions and one transition. Primer COI 14F contained seven degenerate bases and no polymorphism. Primers COIL and LCOI were identical and contained one nucleotide polymorphism on the 5’ end (transition).

Reverse primers were mapped against Polygyridae reference sequences. Primer alignment figures have been provided in Figure 9. Four primers: HCO2198, COIH, COI 698R, and HCOI aligned at positions 6415–6440 bp against reference sequences KX240084.1 and KX278421.1. Primers HCO2198 and HCOI were identical and contained two nucleotide polymorphisms at 6420 bp and 6426 bp, both of which were transitions. Primers COIH and COI 698R contained multiple degenerate bases, resulting in one polymorphism across COIH (transversion) and the absence of polymorphism in COI 698R. Primer COI 839R ranged from 6556–6581 bp, contained many degenerate bases, and displayed three polymorphic sites on the 5' end: Two transitions and one transversion. Primers COI long r and JB4.5 were identical, aligned from 6832 – 6855 bp, and contained four nucleotide polymorphisms each: Three transitions and one transversion.

Custom COI primers were developed using Primer3 (Rozen and Skaletsky 2000) with an optimal length of 20 bp, T_m of 60$_o$C, and a product range of 300-800 bp. Four modified and four custom COI primers were selected to test against *Neohelix* and outgroup tissue for template specificity (Table 5). Modified and custom COI primers from Table 5 were aligned to Polygyridae reference sequences (KX240084.1 and KX27842.1) and mapped in Figure 4.

Table 5. Modified and Custom Primers and PCR Conditions for COI Locus.

Modified Published Primers			
Primer	Sequence (5'-3')	T_m ($_o$C)	Length
polyLCO-F5735	ATTCTACAAATCATAAAGATATTGG	52	25
polyHCO-R6440	TAAACTTCAGGGTGGCCAAAGAATCA	65	26
polyCOI-R6581	ATATGATGGGCCCAAACAATAAAACC	62	26
polyJB4.5-R6855	TATAGATAATACATAATGAAAATG	48	24
Custom Primers			
polyCOI-F99	ATTCTACAAATCATAAAGATATTGG	59	21
polyCOI-R6	TCAGGGTGGCCAAAGAATCA	59	20
polyCOI-R1	AAACTTCAGGGTGGCCAAAG	59	20
polyCOI-R290	GGCCGCTTTGTATTGGGTTT	59	20

Figure 4. Modified and custom COI primers aligned to KX240084.1 (*Praticolella mexicana*) and KX278421.1 (*Polygyra cereolus*) reference sequences.

Note. Lime green = custom forward primer, dark green = modified forward primer, aquamarine = custom reverse primers, and dark blue = modified reverse primers.

Primers were paired in each combination of forward and reverse reads until the most robust COI amplification was uncovered. The pairing of modified Folmer primers polyLCO-F5735 and polyHCO-R6440 resulted in nonspecific amplification of *Neohelix* species (Figure 5A) despite the elimination of primer site polymorphism.

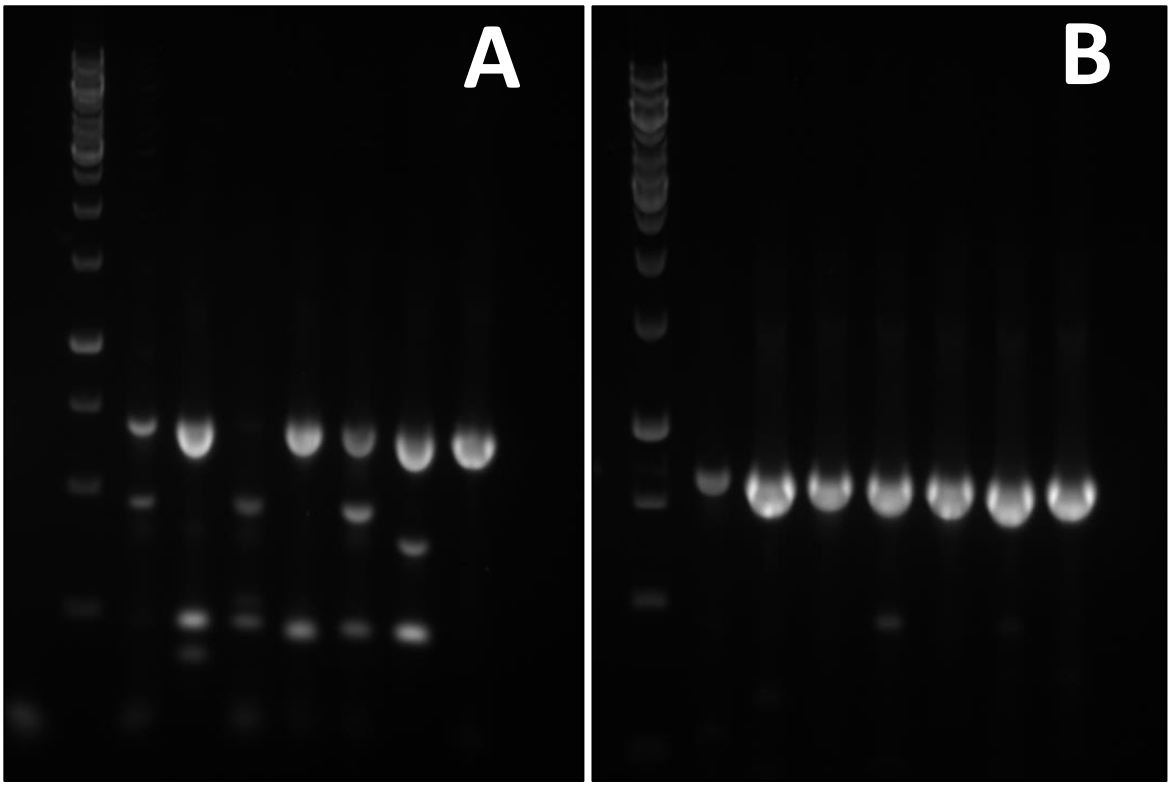

Figure 5. Comparison of amplification success among two modified primer sets.

Part A: Modified primers polyLCO-F5735 and polyHCO-R6440 – Nonspecific amplification of *Neohelix* and outgroup species.
Part B: Modified primers polyLCO-F5735 and polyCOI-R6581 - Specific amplification.

Note. Ladder is 1 kb, identity of lanes (from left to right) are *Neohelix albolabris, N. alleni, N. dentifera, N. divesta, N. major, N. solemi, X. fosteri,* and negative control.

Other combinations of modified and custom primers were equally unsuccessful, including primers: polyHCO-R6440, polyJB4.5-R6855, polyCOI-F99, polyCOI-R6, polyCOI-R1, and polyCOI-R290 (Table 5). Modified primers polyLCO-F5735 and polyCOI-R6581 reliably amplified *Neohelix* and outgroup taxa (Figure 5B). Primer polyCOI-R6581 annealed downstream of the widely used HCO2198 and contributed 141 additional nucleotides toward the COI locus, of which 60 were polymorphic. The incorporation of 141 additional nucleotides resulted in a total length of 846 bp for the COI locus and increased the total number of polymorphic sites to 333. The light band for *N.*

albolabris in lane 1 of Figure 5 was present in Parts A and B, however, its product was still amplified at a concentration of >50 ng/µL.

Sequence Data and Alignments

High quality sequences were generated for four loci including all *Neohelix* species and outgroups sampled. Two loci (cytb and 12S) did not robustly amplify in *Neohelix* taxa and were removed from further analysis. Trimmed alignment data from loci COI, 16S, H3, and 28S are summarized in Table 6 and are available from authors on request.

Table 6. Alignment summary of loci COI, 16S, H3, and 28S.

Locus	# of OTUs	Alignment Length	# of Polymorphic Sites	# of Gaps
COI	18	801 bp	333	0
16S	19	434 bp	199	60
H3	29	327 bp	21	0
28S	28	627 bp	0	0

The number of individuals sequenced for each locus were: COI, 18 individuals; 16S, 19 individuals; H3, 29 individuals; 28S, 28 individuals. Alignment length was calculated using trimmed sequence alignments. Locus COI produced trimmed sequence lengths of 760-801 bp, 16S produced 358-434 bp, H3 produced 270-327 bp, and 28S produced 508-627. The number of polymorphic sites in COI and 16S was large with 16S containing significant gaps. Although H3 and 28S consistently amplified all specimens, H3 produced few informative polymorphic sites and 28S revealed no polymorphism across any *Neohelix* sequence.

Sequence Analysis

The appropriate nucleotide substitution model for each single locus alignment was determined using jModelTest2 (Darriba *et al*. 2012; Guindon and Gascuel 2003) based on a deltaBIC score of zero. The TIM3 model (Posada 2003) with proportion of invariable sites (I) and Gamma distribution (G) was the best supported model for the COI locus with a p-value of 60. The Hasegawa-

Kishino-Yano (HKY) model (Hasegawa *et al*. 1985) with I+G was supported in 16S with a p-value of 46. The HKY model + I was supported in H3 with a p-value of 69. The Jukes-Cantor (JC) model (Jukes and Cantor 1969) was supported in 28S with a p-value of 64. These models were applied to the appropriate partition of concatenated alignments.

Outgroups

Patera perigrapta (Polygyridae: Mesodontini), the most distantly related outgroup and morphologically distinct species in this study, was used to root all single locus and concatenated trees. The other outgroups used for this study were *Mesodon normalis* (Polygyridae: Mesodontini) and *Xolotrema fosteri* (Polygyridae: Triodopsini). *Mesodon normalis* was not used to root trees due to its morphological similarity with *Neohelix* species. *Xolotrema fosteri* was not used to root trees because of its co-residence with *Neohelix* in Tribe Triodopsini.

Total absence of polymorphism in the 28S nuclear locus resulted in a large polytomy (Figure 27); so 28S was eliminated from further phylogenetic inference. Outgroups within the single locus COI trees rooted the intended clades (Figure 12, 13, 14), but one *N. albolabris* from Watauga County, NC MT252628 (Field number ASV 2008-049) and all *N. major* samples clustered with outgroup species. All *Mesodon normalis* samples formed a well-supported (>95 bootstrap [BS], 1.0 posterior probability [PP]) clade separate from Triodopsini. *Xolotrema fosteri* fell within Triodopsini (>71 BS, >0.96 PP) as expected, but was not separated from *Neohelix* with high support (>56 BS, >0.85 PP). The outgroups for single locus 16S trees (Figure 15, 16, 17) followed the same trend as those from COI, but with less support. Mesodontini grouped monophyletically (>77 BS, >1.0 PP), containing all *M. normalis* replicate individuals as well as *N. albolabris* MT254101 (ASV 2008-049), all *N. major* duplicates, and *N. divesta* MT254106 (ACW 2019-001). *Xolotrema fosteri* was not successfully sequenced in locus 16S, therefore, was not displayed on the single locus tree. In nuclear locus H3 (Figure 18, 19, 20), *M. normalis* replicate individuals formed a polytomy with *N. major* duplicates and *N. solemi* MT266965 (ASV 2014-072). Triodopsini was weakly supported (>80 BS, >0.94) with

X. fosteri forming a polytomy composed of *N. albolabris*, *N. solemi*, *N. major*, *N. alleni*, *N. divesta*, *N. lioderma*, and *N. dentifera*.

Comparison of Tree Building Methods

Three informative loci: COI, 16S, and H3 were phylogenetically inferred using three tree building methods: Maximum likelihood (ML), Bayesian inference (BI), and Maximum parsimony (MP). Trees were inferred from single locus alignments which are available from authors on request.

In locus COI (Figure 12, 13, 14), the monophyly of Mesodontini was well supported with ML >95 bootstraps (BS), BI >1.0 posterior probabilities (PP), and MP >96 BS. However, support for Triodopsini differed with BI producing >0.96, but ML producing >71 and MP >61. Similarly, *N. lioderma* and *N. dentifera* duplicate individuals formed a supported clade for BI >0.95, but not for ML >71 or MP. The MP inference for COI grouped *N. dentifera* duplicate taxa with >100 BS support, but *N. lioderma* formed a polytomy. The well supported clade (ML >95, BI >1.0) containing *N. albolabris, N. alleni, N. solemi,* and *N. divesta* duplicate taxa was resolved as a polytomy under the MP COI tree. The reference sequence MG421767.1 (*N. albolabris*) grouped monophyletically with *N. alleni* MT252639 (ASV 2019-039), supported by ML >90 and BI >0.98, but not by MP with a BS support of >059.

Locus 16S (Figure 15, 16, 17) displayed less overall ML support than BI or MP. Mesodontini was monophyletically supported with BI >1.0 and MP >99, but not by ML with a BS of >77. Low ML support continued in Mesodontini with *N. albolabris* MT254101 (ASV 2008-049) and *N. divesta* MT254106 (ACW 2019-001) producing BI support of >0.97 and MP >97, but ML produced a low BS of >74. Likewise, *N. major* MT254104 (ASV 2005-016) and *M. normalis* MT254096 (ASV 2018-037) clustered with a BI of >1.0 and MP of >84, but lost support with a ML of >75. Well-supported nodes in Triodopsini remained constant across ML, BI, and PP inference, except for the clade containing *N. alleni, N. albolabris, N. solemi,* and *N. divesta* replicates. This clade received a support BI of >1.0, but lost support in ML >79 and MP >65.

Locus H3 (Figure 18, 19, 20) produced little support for most nodes, but BI inference was best supported. Individuals *N. albolabris* MT266962 (ASV 2008-049) and *N. divesta* MT266954 (ACW 2019-001), which grouped in Mesodontini for COI and 16S loci, were clustered with BI support of >0.94, but lost support in ML >64 and MP >64 inferences. Individuals *N. major* and *M. normalis*, which previously grouped in Mesodontini under COI and 16S loci, formed a polytomy in H3 with no support. The Triodopsini clade was supported by ML >80 and BI >0.94 PP, but not MP which formed a triodopsid polytomy. Within Triodopsini, all *N. alleni* duplicates and two *N. albolabris* duplicates (MT266963 [ASV 2010-068] and MT266961 [ASV 2007-098]) grouped with BI >1.0 support, but lost support with a ML of >29 and a MP of >50. Duplicates of *N. dentifera* and *N. lioderma* grouped with a ML >82 and BI > 1.0 but were not supported in MP with a BS of > 68.

Comparison of Gene Trees for Incongruence

Each gene tree was compared to identify incongruences in the placement of taxa. Gene trees COI and 16S grouped Mesodontini with moderate support, but lack of support in H3 created a mesodontid polytomy. Loci 16S and H3 grouped *N. albolabris* MT254101; MT266962 (ASV 2008-049) and *N. divesta* MT254106; MT266954 (ACW 2019-001), but absence of the *N. divesta* ACW 2019-001 sequence in COI prevented a unanimous grouping. All replicates of *N. major* and *M. normalis* grouped strongly in COI with ML >98, BI >1.0, and MP >99, but were not supported by 16S with a ML >44, BI >0.88, or MP >71. Triodopsini was supported in 16S with ML >88, BI >1.0, and MP >95. Locus COI weakly supported Triodopsini with a BI of >0.96 but was not supported by a ML of >71 nor a MP of >61. In locus COI, *N. lioderma* grouped with two *N. dentifera* duplicates under a BI of >0.95 but was not supported by a ML of >71 and formed a polytomy during MP. Likewise, locus H3 grouped *N. lioderma* MT266981 (Coles-A188) and four *N. dentifera* duplicates into a well-supported group with a BI of >1.0 and a ML >82, but no MP support. In locus 16S, the sole *N. dentifera* MT254095 (ASV 2018-067) individual did not group with *N. lioderma*. In the COI locus, duplicates of *N. alleni*, *N. albolabris*, *N. solemi*, and *N. divesta* nested within *Neohelix* with ML

> 95 and BI >1.0 but no MP support. The 16S locus produced no support for the aforementioned duplicates with a ML of > 54, BI > 0.87, and MP >53. Additionally, H3 did not support the aforementioned duplicates, producing a ML support of >46 and no BI or MP support. The COI locus grouped all *N. solemi* duplicates into one group as well as one *N. divesta* MT252640 (ASV 2019-046) and one *N. albolabris* MT252631 (ASV 2015-171) with ML >82, BI >1.0, and MP >88. The 16S locus also grouped all *N. solemi* replicates with *N. divesta* MT254091 (ASV 2019-046) and *N. albolabris* MT254099 (ASV 2015-171), supported by a ML of >0.96, BI >1.0, and MP >84. Locus H3 formed a polytomy from *N. solemi*, *N. divesta*, and *N. albolabris* replicates.

Single locus sequences were concatenated into a consensus alignment from which ML, BI, and MP trees were constructed. Polytomies identified in the H3 gene tree partition of the concatenated dataset significantly decreased node support in the consensus tree of COI, 16S, and H3 (Figure 24, 25, 26) Phylogenetic inference using only the COI and 16S concatenated alignment produced the best supported consensus tree (Figure 6, 21, 22, 23).

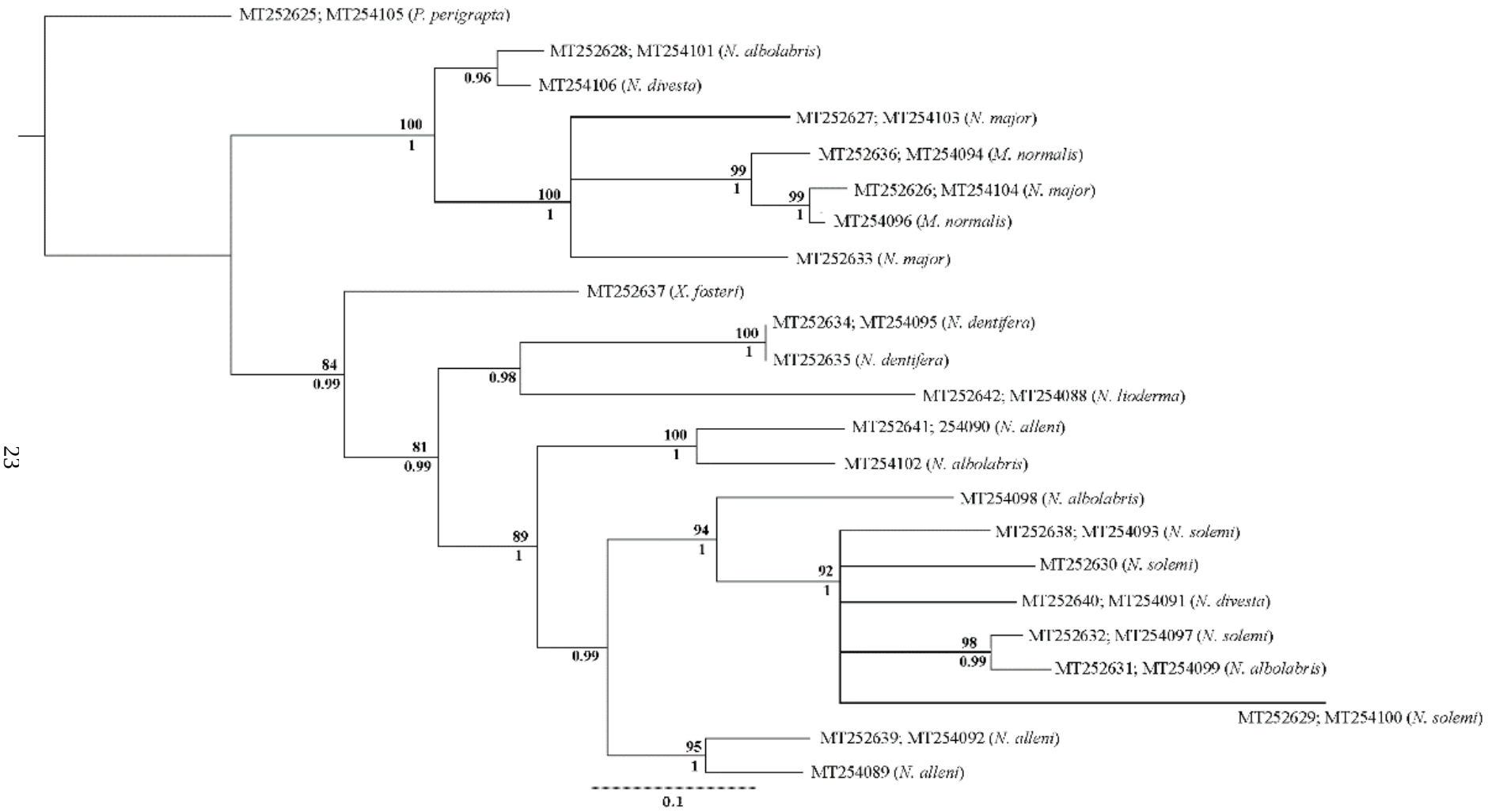

Figure 6. Consensus tree of COI and 16S loci with bootstrap percentages (BS) (above) and posterior probabilities (PP) (below).

DISCUSSION

Morphological identification, especially among convergent taxa, requires verification using either reproductive dissection or genetic data. Phylogenetic analysis is the most reliable method of identifying taxa and can be used to identify and correct misidentifications. However, before phylogenetic implications of interspecific relationships can be applied, identifications must be verified using reproductive dissection. The discussion of data herein is the foundation for further expansion on interspecific *Neohelix* relationships and future studies using distribution, phylogeographic analysis, and reproductive dissection.

DNA Extraction

Mucopolysaccharides and polyphenolic proteins of mollusks may prevent extraction of high-quality DNA (Winnepenninckx *et al*. 1993; Pereira *et al*. 2011). However, extraction of DNA from specimens collected and fixated in >95% EtOH up to ten years prior to extraction (>2010) consistently produced pure DNA with concentrations >50 ng/µL under ND-1000 Nanodrop Spectrophotometer analysis. Specimens fixed prior to 2000 typically produced significantly less DNA product (10-30 ng/µL). In addition to the duration of preservation, fixative type may have impacted extraction success of samples. Fixatives like formalin and paraffin were more commonly used to preserve specimens prior to 2000 and negatively influence DNA quality (Miething *et al.* 2006; Gilbert *et al*. 2007; Ferrer *et al*. 2007; Duval *et al.* 2010; Paireder *et al*. 2013).

The use of ≤5 mg of tissue for DNA extraction under standard Invitrogen PureLink™ Genomic DNA Mini Kit conditions for blood and tissue is recommended, as quantities of tissue between 2-5 mg yielded greater DNA concentration and purity. This finding is concurrent with Pereira *et al.* 2011, which obtained robust molluscan DNA concentrations using tissue from a weight class of ≤5 mg.

PCR Amplification and COI Modification

Clean PCR amplification of loci H3 and 28S in *Neohelix* and outgroup species was correlated with low sequence variation for both loci. Locus 16S initially produced nonspecific amplification (5+ bands); but the Hamstead *et al*. 2015 thermocycler conditions, which included a -0.5°C per cycle touchdown, decreased nonspecific amplification to two bands. Manual agarose gel excision of ideal band size for 16S (~550 bp) produced sequences similar to other molluscan 16S sequences from GenBank. Nonspecific amplification in the 16S locus was mirrored by high rates of polymorphism in the 16S alignment.

Failure of LCO1490 and HCO2198 (Folmer *et al*. 1994) to consistently amplify the COI locus for all *Neohelix* and outgroup species was attributed partially to primer site polymorphism, but also to nonspecific template binding in the reverse oligo. Issues with amplification strength and specificity using Folmer *et al*. 1994 COI primers were also experienced by correspondent Alex C. Cooper from the University of Oklahoma (personal correspondence). As such, amplification issues using Folmer *et al.* 1994 in certain species of Polygyridae do not appear to be an isolated incident.

Custom primers developed for COI amplification: PolyCOI-F99, polyCOI-R6, polyCOI-R1, and polyCOI-R290 (Table 5) did not reliably amplify *Neohelix* and outgroup species. Absence of polymorphism in custom primers as well as identical T_m (59°C) among oligos suggests that lack of amplification was due to nonspecific template binding. However, primer design was limited by the availability of whole mitochondrial genome polygyrid reference sequences. The two references KX240084.1 and KX27842.1 (Polygyridae: Subfamily Polygyrinae) did not share the same subfamily as the species sequenced during this study (Polygyridae: Triodopsinae), and therefore may not be accurate representatives of the species' primer-template binding sites. Additional whole mitochondrial genome sequencing of polygyrids is recommended for further primer design specific to Triodopsinae. Modification of binding sites for LCO1490 and HCO2198 to polyLCO-F5735 and polyHCO-R6440 respectively did not produce clean amplification for all species. Pairing polyLCO-F5735 with the modified reverse primer polyCOI-R6581, which sits 141 bp downstream of

polyHCO-R6440, resulted in consistent and successful COI amplification. This result indicates that the modified forward primer polyLCO-F5735 was specific to the templates of species in this study, but polyHCO-R6440 was not. The rate of polymorphism in the COI alignment using the modified polyLCO-F5735 and polyCOI-R6581 primers was 2.54 nonpolymorphic nucleotides per polymorphic site (846 bp alignment with 333 total polymorphic sites). The rate for the most successful alignment of COI using Folmer *et al.* 1994 primers was 2.58 nonpolymorphic nucleotides per polymorphic site (705 bp alignment with 273 total polymorphic sites). The similar rates between these two primer sets indicates that the 60 additional polymorphic sites added to the COI alignment by modified primers polyLCO-F5735 and polyCOI-R6581 did not greatly deviate from the published rate of polymorphic sites that already composed this locus.

Sequence Alignments and Analysis

Amplification success of loci affected the number of sequenced specimens per locus. Specimens were initially sequenced using non-modified primers (i.e. LCO1490 and HCO2198) to determine whether sequences were similar to published GenBank polygyrid sequences. By performing initial comparisons of sequence data, contamination or inaccuracy of sequences were eliminated. Nuclear locus H3 produced 28 total sequenced specimens with 21 polymorphic sites in a trimmed 327 bp alignment (Table 6). Although informative polymorphic sites in H3 were low compared to the mitochondrial 16S and COI alignments, high amplification specificity and the presence of polymorphic sites makes this locus worth testing during phylogenetic study. While H3 was not phylogenetically informative for delineating *Neohelix* interspecific relationships, it may not provide informative results. Nuclear locus 28S yielded 29 total sequenced specimens, however, it produced zero polymorphic sites with a trimmed alignment length of 627 bp. Due to the absence of informative sites, 28S is not recommended for further phylogenetic use in Polygyridae.

Mitochondrial locus 16S yielded 19 total sequenced specimens with 199 polymorphic sites in a 434 bp trimmed alignment. Significant gaps were common in the 16S alignment using reference

sequences JX839905.1 and JX839894.1, however, hypervariable regions in 16S are well documented (Maly and Brimacombe 1983; Gray *et al*. 1984; Van de Peer *et al*. 1996; Chakravorty *et al*. 2007). Consequently, locus 16S is highly recommended for phylogenetic use in polygyrids. Mitochondrial locus COI was confirmed to be significantly informative for phylogenetic study, yielding 18 total sequenced specimens that produced 333 polymorphic sites in a trimmed 801 bp alignment. Primer modification was necessary for robust amplification in most taxa. Use of modified reverse oligo polyCOI-R6581 instead of polyHCO-R6440 contributed 60 additional polymorphic sites to the COI alignment, which was conducive to greater phylogenetic support. Oligo polymorphism was commonly observed on the 5' end of published COI forward primers, suggesting that the 5' end is more variable for Triodopsinae sequenced during this study.

Outgroups

To determine intraspecific phylogenetic relationships among *Neohelix* species, outgroups were selected based on past phylogenetic and morphology-based studies including Triodopsini (Emberton 1988; Cuezzo 1990; Emberton 1991b; 1994a; 1994a-b; 1995a-b; Perez *et al*. 2014). Previous phylogenetic study of polygyrid relationships found *Neohelix* to be paraphyletic, as *Neohelix alleni* and *Neohelix dentifera* clustered with the *Neohelix* sister clade, *Xolotrema* (Perez *et al*. 2014). Penial dissection by Emberton (1988; 1995b) suggested *Xolotrema* was sister to *Neohelix*. During this study, one *Xolotrema fosteri* individual was used as an outgroup to provide a weak test for *Neohelix* monophyly within Triodopsini. The distant relationship uncovered by Emberton (1995b) between Tribes Mesodontini and Triodopsini based on penial pilaster and allozymic comparisons was not recovered in Perez *et al*. 2014, which described Mesodontini and Triodopsini as sister tribes. During this study, individuals of *Mesodon normalis* and *Patera perigrapta* from Mesodontini were used as outgroups to root the Triodopsini clade.

Many studies have described the remarkable shell convergence among globose polygyrid species, especially in *Mesodon normalis* and *Neohelix major* (Pilsbry 1940; Solem 1976; Emberton

1988; 1991a-b; 1994a-b; 1995a-b; 1996; Perez *et al*. 2014). For this reason, internal anatomy as opposed to shell morphology has been the most reliable mode of identification (Solem 1976; Emberton 1988; 1991a-b; 1994a-b; 1995a-b; Perez *et al*. 2014). However, due to the constraints of this study, internal dissection was not performed. Instead, specimens were selected based on known distribution in addition to morphology-based identification by authors Amy S. Van Devender and R. Wayne Van Devender, both of whom have over 40 years of experience in land snail identification.

Comparison of Tree Building Methods

Constructing a phylogenetic tree from raw sequence data that accurately represents evolutionary relationships requires using the most appropriate analytical model. Phylogenetic inferences Maximum likelihood (ML), Bayesian inference (BI), and Maximum parsimony (MP) are commonly employed (Kim 1996; Alfaro *et al.* 2003; Douady *et al.* 2003; Tamura *et al*. 2011) and were used during this study to compare single locus and consensus trees for consistent topology. Overall similarity among well supported nodes using different analyses suggests that inferences are biologically reasonable interpretations of sequence data (Robinson and Foulds 1981; Choi and Gomez 2009).

Well-supported nodes (>80 BS, 0.95 PP) were identical for ML, BI, and MP topologies. Bayesian inference typically produced higher node support than ML or MP analyses, while MP produced more variable nodes. It was expected that BI posterior probabilities would yield the highest node support, as posterior probabilities are a less conserved method of inference and more likely to support ambiguous hypotheses (Douady *et al*. 2003). Maximum parsimony, while still used, is an older technique proposed by Camin and Sokal (1965) and produces variable phylogenetic support (Felsenstein 1978; Swofford and Olsen 1990; Takezaki and Nei 1994; Kim 1996), therefore, greater topological derivation in MP was expected. For instance, the clade formed by *Neohelix lioderma* and *Neohelix dentifera* in single locus COI trees was well-supported with BI >0.95 PP, not supported with ML >71 BS, and nonexistent in MP with the formation of a polytomy. Likewise, MP analysis of the

H3 locus resulted in a polytomy for Triodopsini whereas ML and BI produced supported nodes for the triodopsine clade (ML >80 BS, BI >0.94 PP). The presence of identical well-supported relationships from the comparison of MP, ML, and BI methodologies suggests that the supported nodes in all three methodologies accurately represent sequence alignments. Polytomies formed by MP did not detract from the viability of these methodologies, as the absence of node support does not reflect conflicting support. Instead, the presence of polytomy rejects the phylogenetic usefulness of its corresponding locus.

Comparison of Gene Trees for Incongruence

Comparing well-supported topologies for incongruence among single locus trees is essential for estimation of deep coalescence or introgression (Delsuc *et al*. 2005; Choi and Gomez 2009). Gene trees strongly supporting alternate nodes could result from incomplete lineage sorting (ILS), i.e. sister species not inheriting sister haplotypes (Maddison 1997; Maddison and Knowles 2006; Kubatko and Degnan 2007; Mirarab *et al*. 2014). In this case, multi-locus concatenation, or combining all loci into a single alignment, would not generate an accurate representation of evolutionary relationships (Kubatko and Degnan 2007; Mirarab *et al*. 2014), and an alternate coalescence method would be required. Therefore, the possibility of ILS must be eliminated.

Mitochondrial loci COI and 16S yielded the greatest support for gene trees, whereas low polymorphism in nuclear locus H3 produced more polytomies. Because polytomy is the lack of support rather than conflicting support, the presence of polytomies are not acknowledged when referencing clade conflict. All gene trees supported a monophyletic Mesodontini containing the same two sister clades. When discussing individuals sequenced in multiple gene trees, the field number instead of GenBank numbers will be referenced for clarity. The first clade incorporated all *Neohelix major* and *Mesodon normalis* individuals while the second clade grouped *N. albolabris* (ASV 2008-049) with *N. divesta* (ACW 2019-001). Triodopsini was a weakly supported clade in all gene trees. Loci COI and H3 formed two sister clades within *Neohelix*. The first clade was composed of *N.*

lioderma and *N. dentifera*, while the second included *N. alleni*, *N. albolabris*, *N. solemi*, and *N. divesta*. However, the absence of *Xolotrema fosteri* in the 16S gene trees resulted in a single *Neohelix* clade with *N. dentifera* plotting as the most basal species. This is likely due to the absence of an outgroup to root Triodopsini. Species distributions within the second clade of *Neohelix* were more variable, primarily due to the disjointed distribution of *N. albolabris* replicates. All *N. alleni* replicates clustered with *N. albolabris* individuals ASV 2008-048, ASV 2010-068, and the MG421767.1 reference sequence, whereas all *N. solemi* replicates were deeply nested in *Neohelix* with *N. divesta* ASV 2019-046 and *N. albolabris* ASV 2015-171. Similar topologies were shared across all single locus trees for COI, 16S, and H3, though variances occurred due to presence or absence of replicates or polytomy. Due to the absence of well-supported conflicting topologies, the presence of ILS or introgression was not supported, therefore, concatenation of single locus alignments was performed to develop a single consensus tree for phylogenetic inference.

Phylogenetic Implications

Phylogenetic analysis in polygyrids is necessary to test morphological identifications and prior relationship hypotheses. This is especially true in Genus *Neohelix*, which contains many convergent species with poorly understood ranges. Due to its frequency and conspicuous size, globose members of *Neohelix* are charismatic fauna of deciduous forests that provide an example of iterated shell evolution, particularly in *Neohelix majo*r and *Mesodon normalis* (Emberton 1988; 1991b; 1994a-b; 1995a-b; 1996). Morphologically similar subspecies of *Neohelix* require further delineation of boundaries, i.e. *N. albolabris albolabris*, *N. albolabris hubrichti*, and *N. albolabris bogani* (Emberton 1988). As reproductive dissections were not performed during this study, subspecific classifications were not targeted for phylogenetic interpretation. If subspecies are present and assigned to the correct species, replicates should form a monophyletic clade. The most robust consensus tree generated by this study (Figure 6) has phylogenetic implications related to past publications focusing on *Neohelix*, Triodopsini, Mesodontini, and Polygyridae (Solem 1976;

Emberton 1988; 1991a-b; 1994a-b; 1995a-b; 1996; Perez *et al*. 2014). Current data tests the monophyly of tribes, genera, and species (Emberton 1988; 1995b; Perez *et al*. 2014). Figure 7 provides a labeled representation of the major clades within this molecular phylogeny: Mesodontini, Triodopsini, and *Neohelix*.

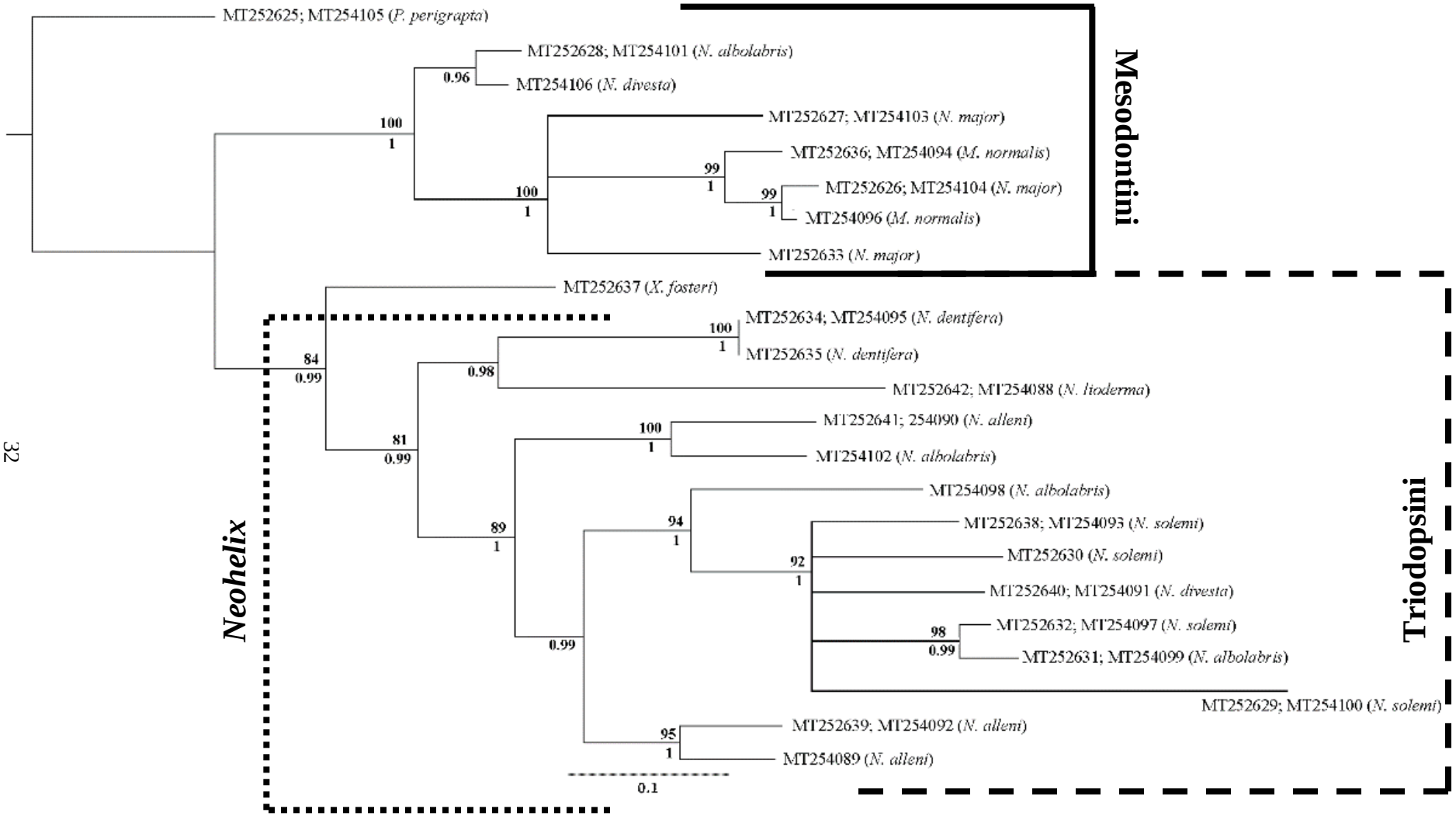

Figure 7. Consensus tree of COI and 16S loci with bootstrap percentages (BS) (above) and posterior probabilities (PP) (below). Three major clades are identified from the top down: Tribe Mesodontini (solid line), Tribe Triodopsini (dashed line), and Genus *Neohelix* (dotted line).

During this study, two outgroups from Mesodontini were used to root the *Neohelix* phylogeny, including *Patera perigrapta* and *Mesodon normalis*. This limited number of mesodontine species does not allow for interpretation of relationships in Mesodontini, though past studies have supported the monophyly this clade (Emberton 1995b; Perez *et al*. 2014). Although *M. normalis* specimens formed a well-supported clade, Mesodontini also incorporated specimens morphologically identified as *Neohelix*. Mesodontini included all specimens identified as *Neohelix major*, *N. albolabris* ASV 2008-049 and *N. divesta* ACW 2019-001.

The latter two *Neohelix*: *N. albolabris* (ASV 2008-049) and *N. divesta* (ACW 2019-001) formed a well-supported sister clade to all *M. normalis* and *N. major* replicates. While the individual identified as *N. albolabris* was collected from Watauga County, NC, the *N. divesta* specimen was collected from Benton County, AR. As population bias could not play a role in the formation of this clade (>75 BS, >0.96 PP), morphological misidentification of one or both specimens could have occurred. Both mature specimens lacked dentition and had closed umbilici (Emberton 1988; 1994a; 1995a; Dourson 2013). The only *Mesodon* local to Watauga Co., N.C. that displays those traits is *M. andrewsae*, however, the known range of *M. andrewsae* does not extend past Tennessee (Perez and Cordeiro 2008) and therefore could not be the identity of the *N. divesta* specimen ACW 2019-001. However, *N. albolabris* (ASV 2008-049) exhibited unique morphology (Figure 8) that could suggest the presence of an unidentified species. Reproductive dissection is required to test prior identification hypotheses before definitive conclusions are drawn.

Figure 8. Photographs of morphologically unique *N. albolabris* specimen from Watauga Co., NC (ASV 2008-049).

The second sister clade within Mesodontini was composed of individuals identified as *Mesodon normalis* and *Neohelix major*. Both species are very similar in shell morphology when found in sympatry (Emberton 1988; 1991a-b; 1994a-b; 1995a-b; 1996). Hypotheses for species ranges were used to supplement morphological identifications in the absence of reproductive dissections. For example, *M. normalis* is typically distributed in northern, more high elevation ranges than *N. major* (Hubricht 1985; Dourson 2013). The well-supported (>100 BS, 1.0 PP) formation of a clade within Mesodontini containing all *N. major* and *M. normalis* replicates suggests that reproductive dissections are required to confirm morphological hypotheses. While the specimen identified as *N. major* (ASV 2005-016) from Jackson Co., NC could be re-identified as *M. normalis*

based on locality, other *N. major* specimens from Stanly Co., NC and Walker Co., GA were collected deep within known localities for *Neohelix major*. Reproductive dissections should be conducted on specimens initially identified as *N. major* to test morphological identifications. Tentative hypotheses for true specimen identity based on phylogenetic and locality data have been provided in Figure 9.

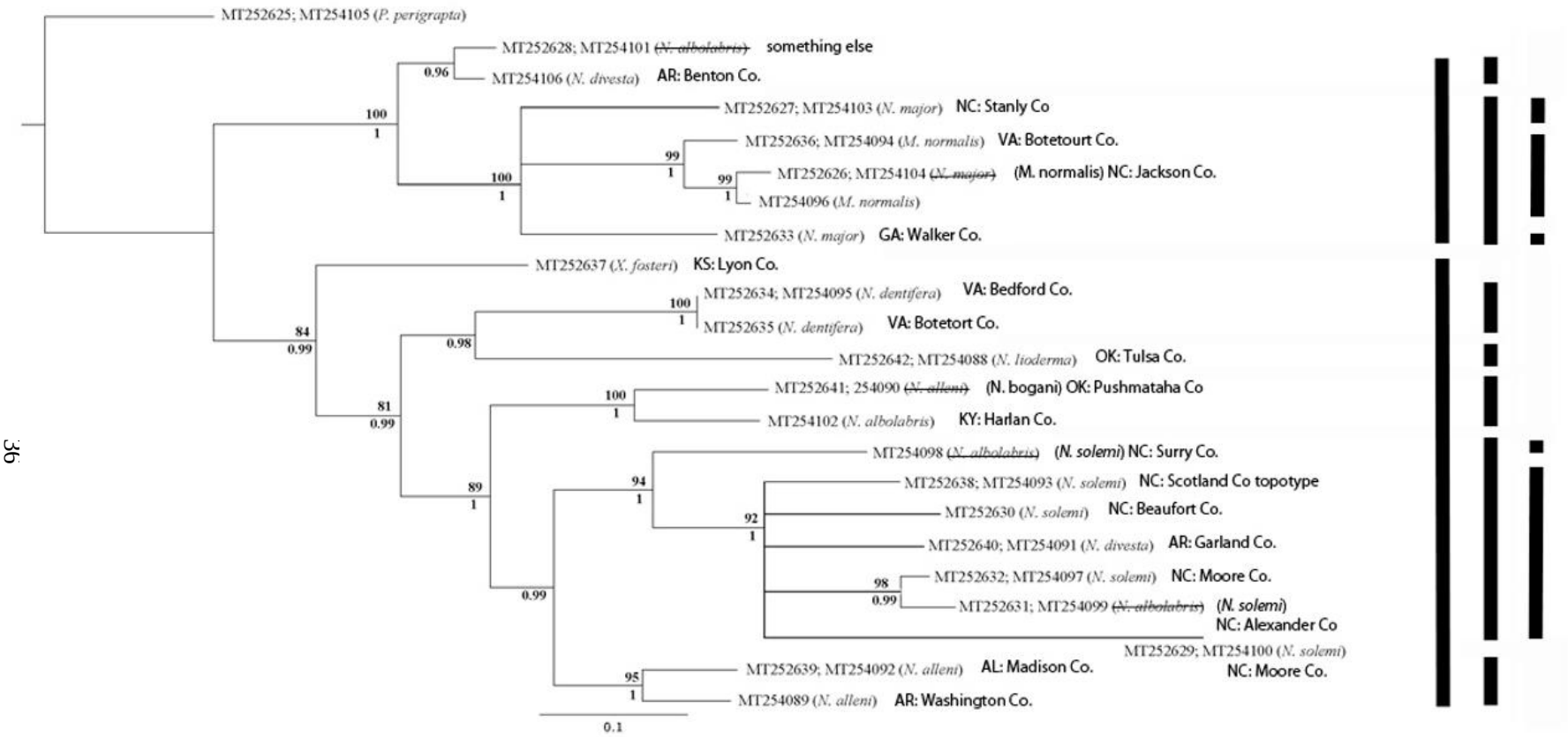

Figure 9. Labeled phylogenetic consensus tree utilizing COI and 16S loci. County data has been included to the right of initial morphological identifications. When appropriate, better approximations of specimen identity based on phylogenetic evidence have been provided. Black bars to right of figure represent clade distribution. Bootstrap (BS) values are displayed above nodes and posterior probabilities (PP) below nodes.

Paraphyly of *Neohelix* was supported in Perez *et al*. 2014 due to *N. alleni* and *N. dentifera* grouping with *Xolotrema*. Although only one *Xolotrema fosteri* was sequenced during this study, no *Neohelix* species formed a clade with this outgroup. *Neohelix* paraphyly in this study was instead attributed to the clustering of *N. albolabris* (ASV 2008-049), *N. divesta* (ACW 2019-001), and *N. major* specimens from Stanly Co., NC and Walker Co., GA within Mesodontini. Before *Neohelix* paraphyly can be confirmed, reproductive dissections must be used to confirm phylogenetic data.

The first species group within the paraphyletic *Neohelix* was a well-supported clade formed by *N. dentifera* and *N. lioderma* specimens. Using reproductive dissection and allozymic characters, Emberton 1988 formed a species group titled the "dentifera group," which included *N. dentifera*, *N. lioderma*, and *N. divesta*. During this study, no *N. divesta* replicates grouped with *N. dentifera* or *N. lioderma*. However, all *N. dentifera* replicates and *N. lioderma* formed the first supported clade (>75 BS, >0.98 PP) in *Neohelix*. Based on the sequence data of this study, inclusion of *N. divesta* within the *N. dentifera* species group was not supported. However, the disjoint clustering of *N. divesta* specimens ACW 2019-001 and ASV 2019-046 as well as the morphological ambiguity of this species supports the addition of reproductive dissections to test identity hypotheses in *N. divesta*.

The second clade in *Neohelix* was strongly supported (>100 BS, 1.0 PP) and composed of *N. alleni* (ASV 2019-078) from Pushmataha Co., OK and *N. albolabris* (ASV 2007-098) from Harlan Co., KY. Population bias between these specimens is highly improbable, suggesting that the morphological hypothesis for one or both of these individuals was incorrect. While the range for *N. albolabris* does not traditionally extend further West than the Mississippi, the subspecies *N. albolabris bogani* named by Emberton 1988 occurs sympatrically with *N. alleni* in Arkansas. The morphological differences between *N. albolabris bogani* and *N. alleni* are very ambiguous and can only be definitively identified using reproductive dissection (Emberton 1988). Nested within *Neohelix* was a well-supported (>95 BS, 1.0 PP) clade of two *N. alleni* specimens from Madison Co., AL (ASV 2019-039) and Washington Co., AR (ASV 2019-086). While the spatial distance between the localities of these individuals suggests that population bias did not affect the formation of this clade, it

also suggests that the *N. alleni* specimen ASV 2019-078 from Pushmataha Co., OK is more likely to be *N. albolabris bogani*. It was hypothesized that the *N. alleni* specimen from Madison Co., AL was *N. alleni fuscolabris* while *N alleni* from Washington Co., AR was *N. alleni alleni*. The subspecies *Neohelix alleni fuscolabris* occupies Northern Alabama and Southern Tennessee (Hubricht 1985; Emberton 1988; Perez *et al*. 2014) whereas *N. alleni alleni* occurs only as far east as Missouri. While the well-supported formation of this clade supports the morphological hypotheses of both individuals as *N. alleni*, subspecific identities cannot be interpreted. Reproductive dissection is required for further testing of these species identities.

The third well-supported clade in *Neohelix* included all *N. solemi* specimens, two *N. albolabris* specimens, and one specimen identified as *N. divesta*. The disjoint clustering of *N. albolabris* individuals reflects the morphological ambiguity of this species when not supported by reproductive dissection. However, both *N. albolabris* individuals from Surry Co., NC and Alexander Co., NC plotted within this clade of primarily *N. solemi* specimens. It is possible that these *N. albolabris* specimens could actually be *N. solemi*, expanding the known range of *N. solemi* into the Piedmont (Emberton 1988). The *N. divesta* individual (ASV 2019-046) from Garland Co., AR is deeply nested within *Neohelix* and requires further reproductive analysis to uncover the appropriate identification for this specimen. It is unlikely that *N. divesta* ASV 2019-046 was actually *N. solemi*, as *N. solemi* is distributed along the Eastern U.S. coastal plains and *N. divesta* ranges through the Midwest (Emberton 1988; Perez and Cordeiro 2008). The clustering of all *N. solemi* replicates suggests that collection localities utilized during this study (Table 1) should be confirmed as part of this species' range. Due to the overwhelming morphological similarity of many *Neohelix* species and absence of reproductive dissection, the hypotheses for species identity presented in this study cannot be tested until reproductive dissections are conducted.

CONCLUSION

Initial goals of this study were to uncover the most effective universal markers for delimiting interspecific *Neohelix* relationships. Although the primers and conditions for mitochondrial loci COI and 16S required modification, the resultant markers produced robust consensus and single gene trees. Authors hope that the COI primers polyLCO-F5735 and polyCOI-R6581 will be useful for future investigations of polygyrids. Markers H3 and 28S were not phylogenetically informative despite the ease of amplification within *Neohelix*, as H3 produced little polymorphism and 28S yielded no polymorphic sites. Future phylogenetic work in polygyrids should test alternate nuclear loci, like locus 18S, which has proved phylogenetically informative in pulmonate gastropods (Klussmann-Kolb *et al*. 2008; Dayrat *et al.* 2011).

The production of phylogenetic markers has allowed for the recognition and correction of morphology-based misidentification, though reproductive dissections should be conducted to confirm specimen identity. This study provides an estimation of interspecific *Neohelix* relationships and has produced molecular tools to improve phylogenetic amplification in mitochondrial locus COI. The use of additional nuclear and mitochondrial markers as well as increased sampling size, type localities, and reproductive dissection would provide more robust future estimates of phylogenetic relationships among *Neohelix* and closely related outgroups. These improvements are especially necessary for future studies involving distribution and phylogeography. Despite utilizing a limited sample size, this work will provide a foundation for future work in *Neohelix* and other polygyrid taxa.

Bibliography

Alfaro, M. E., S. Zoller, and F. Lutzoni. 2003. Bayes or bootstrap? A simulation study comparing the performance of bayesian Markov Chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biological Evolution* **20**:255—266.

Bergsten, J., D. T. Bilton, T. Fujisawa, M. Elliott, M. T. Monaghan, M. Balke, L. Hendrich, J. Geijer, J. Herrmann, G. N. Foster, I. Ribera, A. N. Nilsson, T. G. Barraclough, and A. P. Vogler. 2012. The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology* **61**:851—869.

Bowles, J., D. Blair, and D. P. McManus. 1992. Genetic variants within the genus *Echinococcus* identified mitochondrial DNA sequencing. *Molecular and Biochemical Parasitology* **54**:165—174.

Camin, J. H. and R. R. Sokal. 1965. A method for deducing branching sequences in phylogeny. *Evolution* **19**:311—326.

Campbell, D. C., J. M. Serb, J. E. Buhay, K. J. Roe, R. L. Minton, C. Lydeard. 2005. Phylogeny of North American amblemines (Bivalvia, Unionoida): Prodigious polyphyly proves pervasive across genera. *Invertebrate Biology* **124**:131—164.

Chakravorty, S., D. Helb, M. Burday, N. Connell, and D. Alland. 2007. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of Microbiological Methods* **69**:330—339.

Choi, K. and S. M. Gomez. 2009. Comparison of phylogenetic trees through alignment of embedded evolutionary distances. *BMC Bioinformatics* **10**:e423.

Colgan, D. J., A. McLauchlan, G. D. F. Wilson, S. P. Livingston, G. D. Edgecombe, J. Macaranas, G. Cassis, and M. R. Gray. 1998. Histone H3 and U1 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology* **46**:419—437.

Cuezzo, M. G. 1990. Maturation of the reproductive tract of *Neohelix major* (Binney) (Gastropoda: Polygyridae). *American Malacological Bulletin* **8**:19—24.

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

Dayrat, B., M. Conrad, S. Balayan, T. R. White, C. Albrecht, R. Golding, S. R. Gomes, M. G. Harasewych, A. M. de Frias Martins. 2011. Phylogenetic relationships and evolution of pulmonate gastropods (Mollusca): New insights from increased taxon sampling. *Molecular Phylogenetics and Evolution* **59**:425—437.

Delsuc, F., H. Brinkmann, and H. Philippe. 2005. Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics*, Nature Publishing Group **6**:361—375. Available at: https://hal.archives-ouvertes.fr/halsde-00193293 8 July 2020.

Douady, C., J. F. Delsuc, Y. Boucher, W. F. Doolittle, and E. J. P. Douzery. 2003. Comparison of bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Molecular Biological Evolution* **20**:248—254.

Dourson, D. 2012. Four new land snail species from the southern Appalachian Mountains. *Journal of the North Carolina Academy of Science* **128**:1—10.

Dourson, D. 2013. *Land snails of the Great Smoky Mountains National Park and Southern Appalachians Tennessee & North Carolina*. Goatslug Publications, Bakersville, North Carolina, U.S.A. Pp. 11-335.

Duval, K., R. A. Aubin, J. Elliott, I. Gorn-Hondermann, H. C. Birnboim, D. Jonker, R. M. Fourney, and C. J. Fregau. 2010. Optimized manual and automated recovery of amplifiable DNA from tissues preserved in buffered formalin and alcohol-based fixative. *Forensic Science International: Genetics* **4**:80—88.

Emberton, K. C. 1988. The genitalic, allozymic, and conchological evolution of the eastern North American Triodopsinae (Gastropoda: Pulmonata: Polygyridae). *Malacologia* **28**:159—273.

Emberton, K. C. 1991a. The genitalic, allozymic, and conchological evolution of the tribe Mesodontini (Pulmonata: Stylommatophora: Polygyridae). *Malacologia* **33**:71—178.
Emberton, K. C. 1991b. Ecology of a shell convergence between subfamilies of polygyrid land snails. *Biological Journal of the Linnean Society* **44**:105—120.
Emberton, K. C. 1994a. Polygyrid land-snail phylogeny: External sperm exchange, early North American biogeography, iterative shell evolution. *Biological Journal of the Linnean Society* **52**:241—271.
Emberton, K. C. 1994b. Partitioning a morphology among its controlling factors. *Biological Journal of the Linnean Society* **53**:353—369.
Emberton, K. C. 1995a. When shells do not tell: 145 million years of evolution in North America's polygyrid land snails, with a revision and conservation priorities. *Malacologia* **37**:69—110.
Emberton, K. C. 1995b. Sympatric convergence and environmental correlation between two land-snail species. *Evolution* **49**:469—475.
Emberton, K. C. 1996. Microsculptures of convergent and divergent polygyrid land-snail shells. *Malacologia* **38**:67—85.
Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* **27**:401—410.
Ferrer, I., J. Armstrong, S. Capellari, P. Parchi, T. Arzberger, J. Bell, H. Budka, T. Strobel, G. Giaccone, G. Rossi, N. Bogdanovic, P. Fakai, A. Schmitt, P. Riederers, S. Al-Sarraj, R. Ravid, and H. Kretzschmar. 2007. Effects of formalin fixation, paraffin embedding, and time of storage on DNA preservation in brain tissue: A BrainNet Europe study. *Brain Pathology* **17**:297—303.
Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**:294—299.
Gilbert, M. T. P., T. Haselkorn, M. Bunce, J. J. Sanchez, S. B. Lucas, L. D. Jewell, E. Van Marck, and M. Worobey. 2007. The isolation of nucleic acids from fixed, paraffin-embedded tissues: Which methods are useful when? *PLoS One* **2**:e537.
Gray, M. W., D. Sankoff, and R. J. Cedergren. 1984. On the evolutionary descent of organisms and organelles: A global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. *Nucleic Acids Research* **12**:5837—5852.
Guindon, S. and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**:696—704.
Hamstead, J. W., B. L. Snider, R. Oaks, E. Fitzgerald, J. Woodward, A. Teat, K. M. Hay, M. C. Estep, Z. E. Murrell. 2015. Sixteen polymorphic microsatellite markers for a federally threatened species, *Hexastylis naniflora* (Aristolochiaceae), and co-occurring congeners. *Applications in Plant Sciences* **3**:apps.1500033
Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**:160—174.
Hillis, D. M., C. Mortiz, and B. K. Mable. 1996. *Molecular Systematics*. Sinauer Associates, Inc., Sunderland, Massachusetts, U.S.A.
Hubricht, L. 1985. The distributions of the native land mollusks of the Eastern United States. *Fieldiana, Zoology, New Series* **24**:1—191.
Huelsenbeck, J. P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics Applications Note* **17**:754—755.
Jukes, T. H. and C. R. Cantor. 1969. Evolution of protein molecules. *In*: H. M. Munro, eds., *Mammalian Protein Metabolism*, Academic Press, New York, U.S.A. Pp. 21—132.
Kim, J. 1996. General inconsistency conditions for maximum parsimony: Effects of branch lengths and increasing numbers of taxa. *Systematic Biology* **45**:363—374.

Klussmann-Kolb, A., A. Dinapoli, K. Kuhn, B. Streit, and C. Albrecht. 2008. From sea to land and beyond – New insights into the evolution of euthyneuran Gastropoda (Mollusca). *BMC Evolutionary Biology* **8**:e57.

Kocher, T. D., W. K. Thomas, A. Meyer, S. V. Edwards, S. Paabo, F. X. Villablanca, and A. C. Wilson. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *PNAS* **86**:6196—6200.

Kozlov, A. M., D. Darriba, T. Flouri, B. Morel, and A. Stamatakis. 2019. RAxML-NG: A fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**:4453—4455.

Kubatko, L. S. and J. H. Degnan. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology* **56**:17—24.

Maddison, W. P. 1997. Gene trees in species trees. *Systematic Biology* **46**:523—536.

Maddison, W. P. and L. L. Knowles. 2006. Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* **55**:21—30.

Maly, P. and R. Brimacombe. 1983. Refined secondary structure models for the 16S and 23S ribosomal RNA of *Escherichia coli*. *Nucleic Acids Research* **11**:7263—7286.

Merritt, T. J. S., L. Shi, M. C. Chase, M. A. Rex, R. J. Etter, and J. M. Quattro. 1998. Universal cytochrome *b* primers facilitate intraspecific studies in molluscan taxa. *Molecular Marine Biology and Biotechnology* **7**:7—11.

Miething, F., S. Hering, B. Hanschke, and J. Dressler. 2006. Effect of fixation to the degradation of nuclear and mitochondrial DNA in different tissues. *Journal of Histochemistry and Cytochemistry* **54**:371—374.

Mirarab, S., R. Reaz, Md. S. Bayzid, T. Zimmermann, M. S. Swenson, and T. Warnow. 2014. ASTRAL: Genome-scale coalescent-based species tree estimation. *Bioinformatics* **30**:i541—i548.

Oklahoma Department of Wildlife Conservation (ODWC). 2015. *Oklahoma comprehensive wildlife conservation strategy: A strategic conservation plan for Oklahoma's rare and declining wildlife*. Pp. 30—418.

Paireder, S., B. Werner, J. Bailer, W. Werther, E. Schmid, B. Patzak, M. Cichna-Markl. 2013. Comparison of protocols for DNA extraction from long-term preserved formalin fixed tissues. *Analytical Biochemistry* **439**:152—160.

Pereira, J. C., R. Chaves, E. Bastos, A. Leitao, and H. Guedes-Pinto. 2011. An efficient method for genomic DNA extraction from different molluscs species. *International Journal of Molecular Sciences* **12**:8086—8095.

Perez, K. E. and J. R. Cordeiro. 2008. A guide for terrestrial gastropod identification. *American Malacological Society*, Carbondale, Illinois, U.S.A. Pp. 1—73.

Perez, K. E., N. Defreitas, J. Slapcinsky, R. L. Minton, F. E. Anderson, and T. A. Pearce. 2014. Molecular phylogeny, evolution of shell shape, and DNA barcoding in Polygyridae (Gastropoda: Pulmonata), an endemic North American clade of land snails. *American Malacological Bulletin* **32**:1—31.

Palumbi, S., A. Martin, S. Romano, W. O. McMillan, L. Stice, and G. Grabowski. 1991. *The simple fool's guide to PCR*, Version 2.0. Department of Zoology and Kewalo Marine Laboratory, University of Hawaii, Honolulu, Hawaii, U.S.A. Pp. 1—47.

Pilsbry, H. A. 1940. *Land Mollusca of North America (North of Mexico)*, Monograph 3 (Vol. 1, Part 2). Academy of Natural Sciences of Philadelphia. Pp. 702—912.

Posada, D. 2003. Using ModelTest and PAUP* to select a model of nucleotide substitution. *In*: A.D. Baxevanis, D. B. Davison, R. D. M. Page, G. A. Petsko, L. D. Stein, and G. D. Stormo, eds., *Current Protocols in Bioinformatics*. John Wiley & Sons, New York, U.S.A. Pp. 6.5.1—6.5.14.

Robinson, D. F. and L. R. Foulds. 1981. Comparison of phylogenetic trees. *Mathematical Biosciences* **53**:131—147.

Rozen, S. and H. Skaletsky. 2000. Primer3 on the WWW for General Users and for Biologist Programmers. *Methods in Molecular Biology* **132**:365—386.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87**:651—701.

Solem, A. 1976. Comments on eastern North American Polygyridae. *Nautilus* **90**:25—36.

Solem, A. 1978. Classification of the land Mollusca. *In*: V. Fretter & J. Peake, eds., *Pulmonates. Volume 2A. Systematics, Evolution, and Ecology*. Academic Press, London. Pp.49-98.

Solem, A. 1981. Land snail biogeography: A true snail's pace of change. *In*: G. Nelson & D. E. Rosen, eds., *Vicariance Biogeography: A Critique*. New York, U.S.A. Pp. 197-221.

Solem, A. 1985. Origin and diversification of pulmonate land snails. *In*: E. R. Trueman & M. R. Clarke, eds., *The Mollusca, Volume 10, Evolution*. Academic Press, New York, U.S.A. Pp. 269—293.

Solem, A. and E. L. Yochelson. 1978. North American Paleozoic land snails, with a summary of other Paleozoic nonmarine snails. *Geological Survey Professional Paper 1072*:1—42.

Swofford, D. L. 2002. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland, Massachusetts, U.S.A. Pp. 1—499.

Swofford, D. L. and G. J. Olsen. 1990. *Phylogeny reconstruction*. *In*: D. M. Hillis and C. Moritz, eds., *Molecular systematics*, 1st Edition. Sinauer, Sunderland, Massachusetts U.S.A. Pp. 411—501.

Takezaki, N. and M. Nei. 1994. Inconsistency of the maximum parsimony method when the rate of nucleotide substitution is constant. *Journal of Molecular Evolution* **39**:210—218.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biological Evolution* **28**:2731—2739.

Thaewnon-ngiw, B., S. Klinbunga, K. Phanwichien, N. Sangduen, N. Lauhachinda, and P. Menasveta. 2004. Genetic diversity and molecular markers in introduced and Thai native Apple Snails (*Pomacea* and *Pila*). *Journal of Biochemistry and Molecular Biology* **37**:493—502.

Tongkerd, P., T. Lee, S. Panha, J. B. Burch, and D. O Foighil. 2004. Molecular phylogeny of certain Thai Gastrocoptine micro land snails (Stylommatophora: Pupillidae) inferred from mitochondrial and nuclear ribosomal DNA sequences. *Journal of Molluscan Studies* **70**:139—147.

Van de Peer, Y., S. Chapelle, and R. De Wachter. 1996. A quantitative map of nucleotide substitution rates in bacterial rRNA. *Nucleic Acids Research* **1**:3381—3391.

Winnepenninckx, B., T. Backeljau, and R. de Wachter. 1993. Extraction of high molecular weight DNA from molluscs. *Trends in Genetics* **9**:407.

Appendix

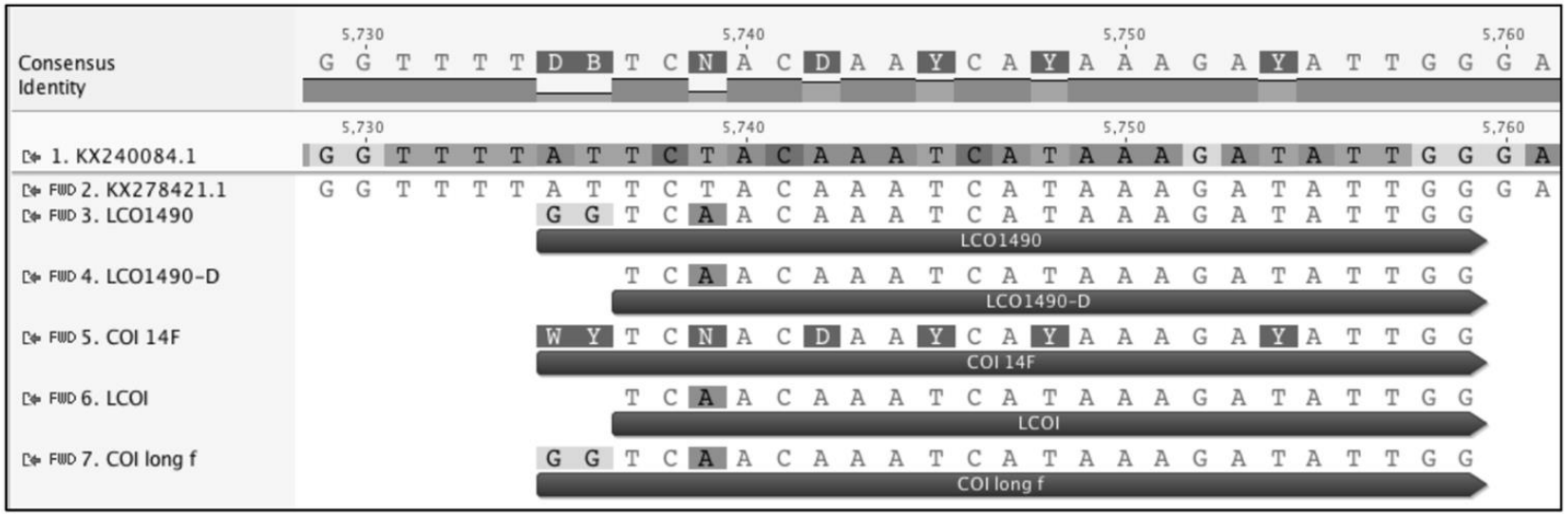

Figure 10. COI published forward primers aligned to KX240084.1 (*Praticolella mexicana*) and KX278421.1 (*Polygyra cereolus*) reference sequences. Annotations represent published forward primers. Nucleotides are highlighted against reference sequences to display polymorphic sites.

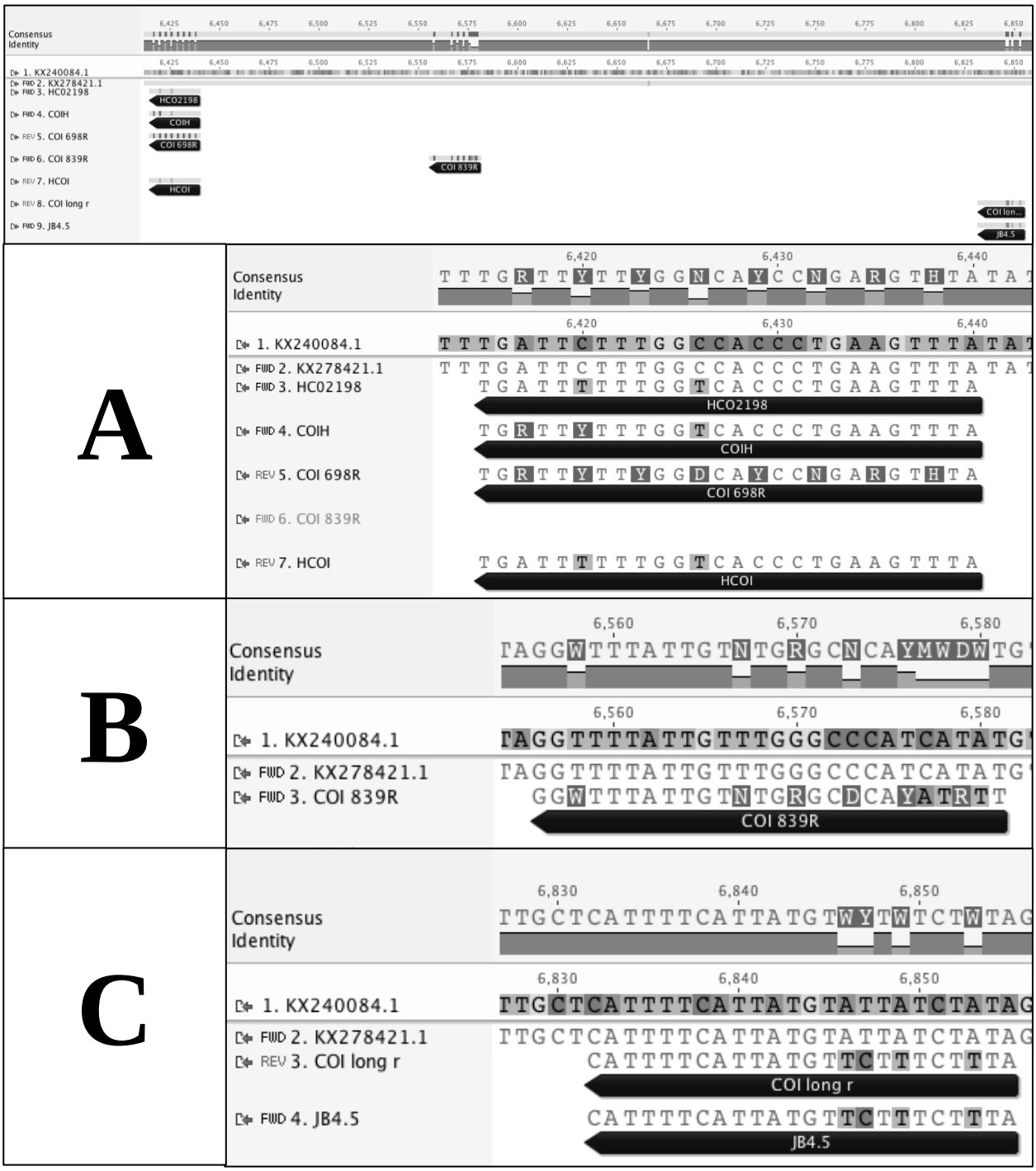

Figure 11. COI published reverse primers aligned to KX240084.1 (*Praticolella mexicana*) and KX278421.1 (*Polygyra cereolus*) reference sequences. Annotations represent published reverse primers. Nucleotides are highlighted against reference sequences to display polymorphic sites.

Note. Part A: HCO2198, COIH, COI 698R, and HCO. Part B: COI 839R. Part C: COI long r, JB4.5.

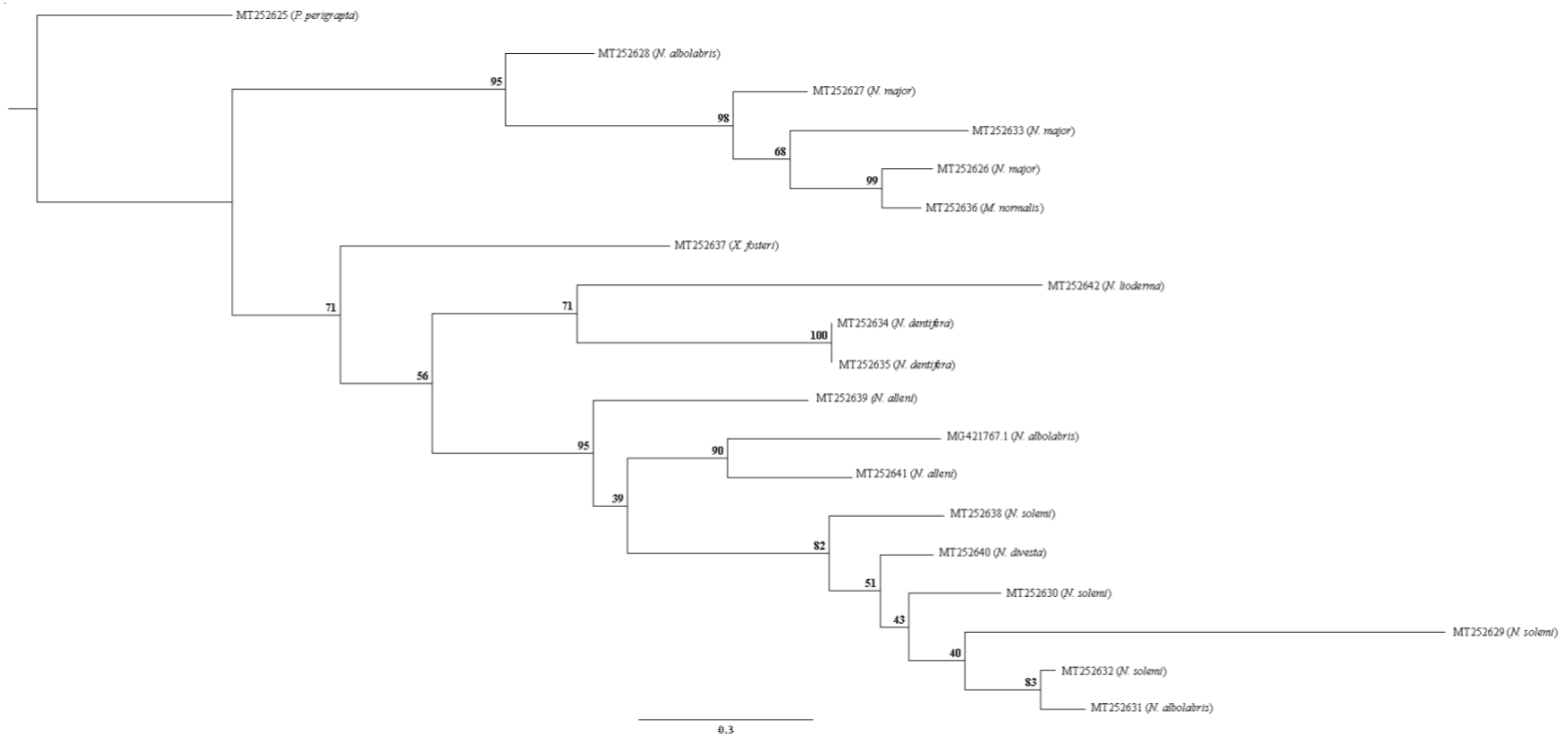

Figure 12. Maximum likelihood molecular phylogeny of locus COI using TIM3+I+G in RAxML-NG. Bootstrap (BS) values displayed above nodes.

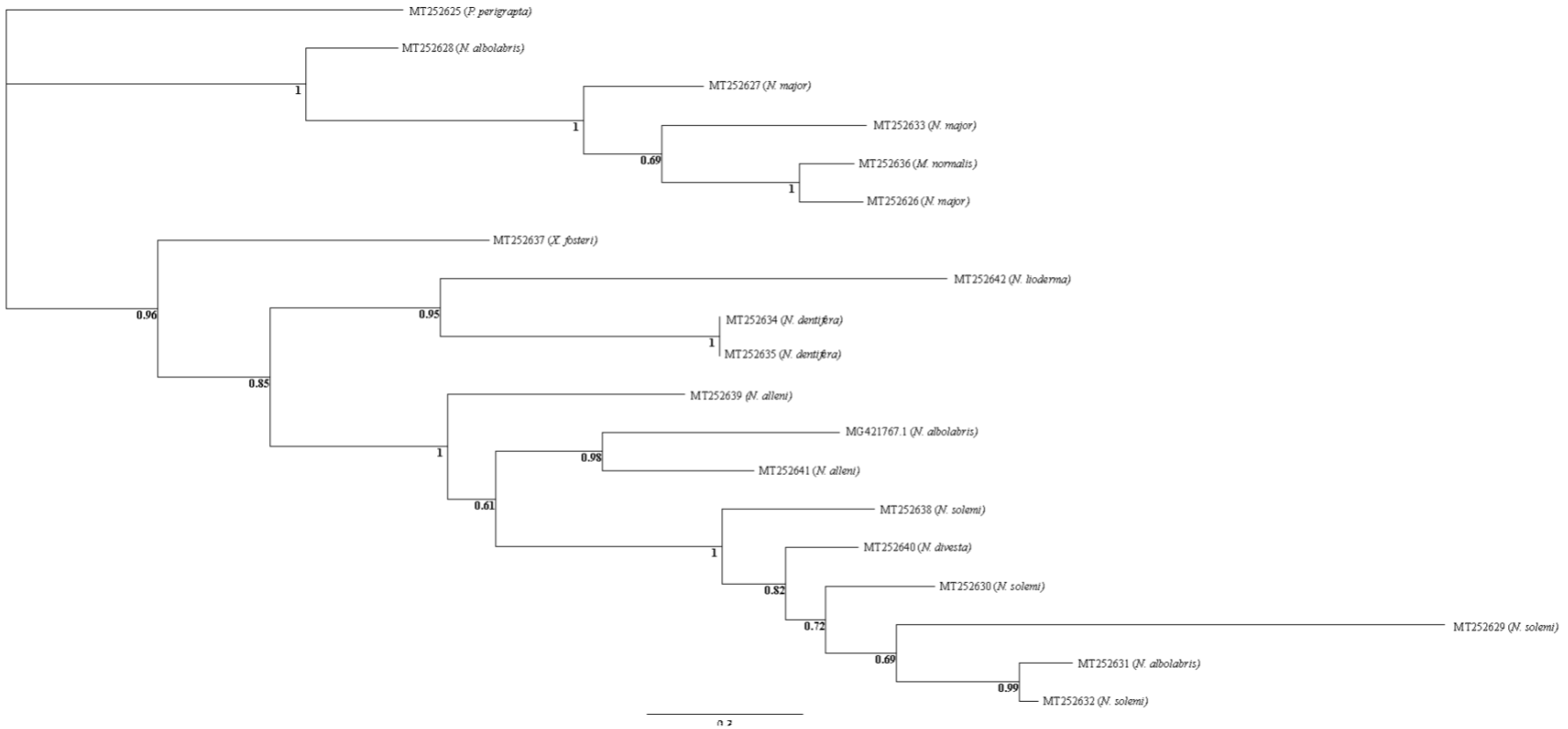

Figure 13. Bayesian inference molecular phylogeny of locus COI in MrBayes. Posterior probabilities (PP) displayed below nodes.

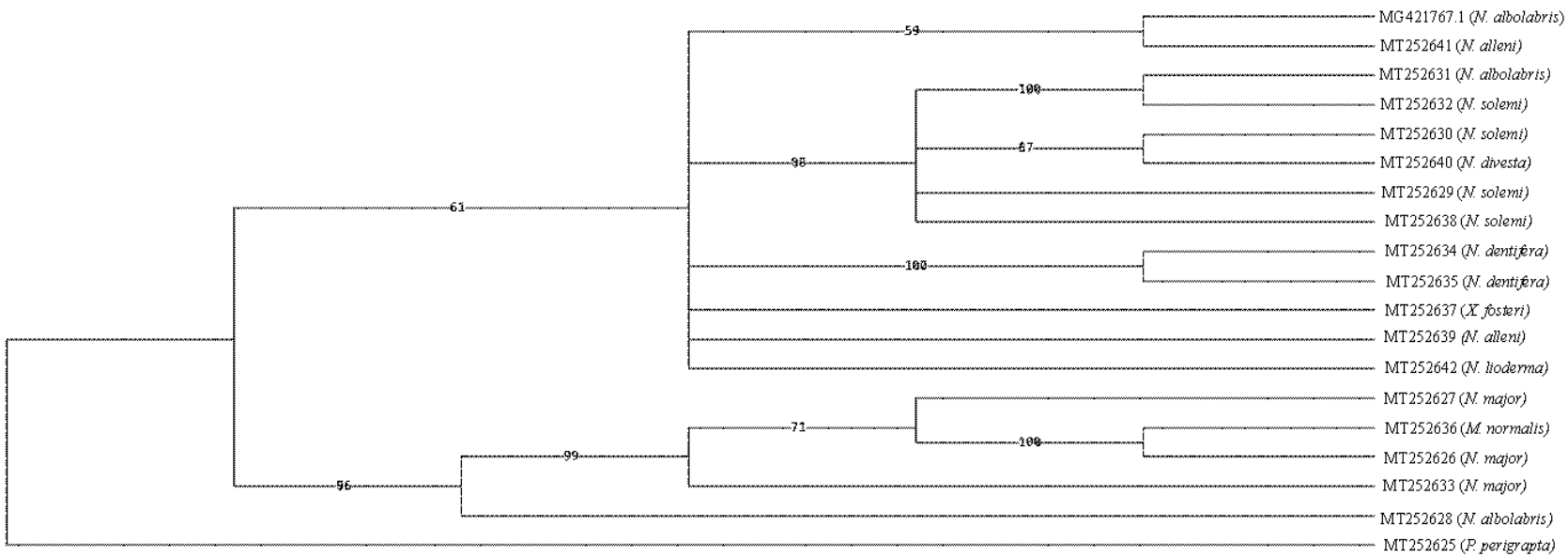

Figure 14. Maximum parsimony molecular phylogeny of locus COI using PAUP*. Bootstrap (BS) values above 50 are displayed in the line of each node.

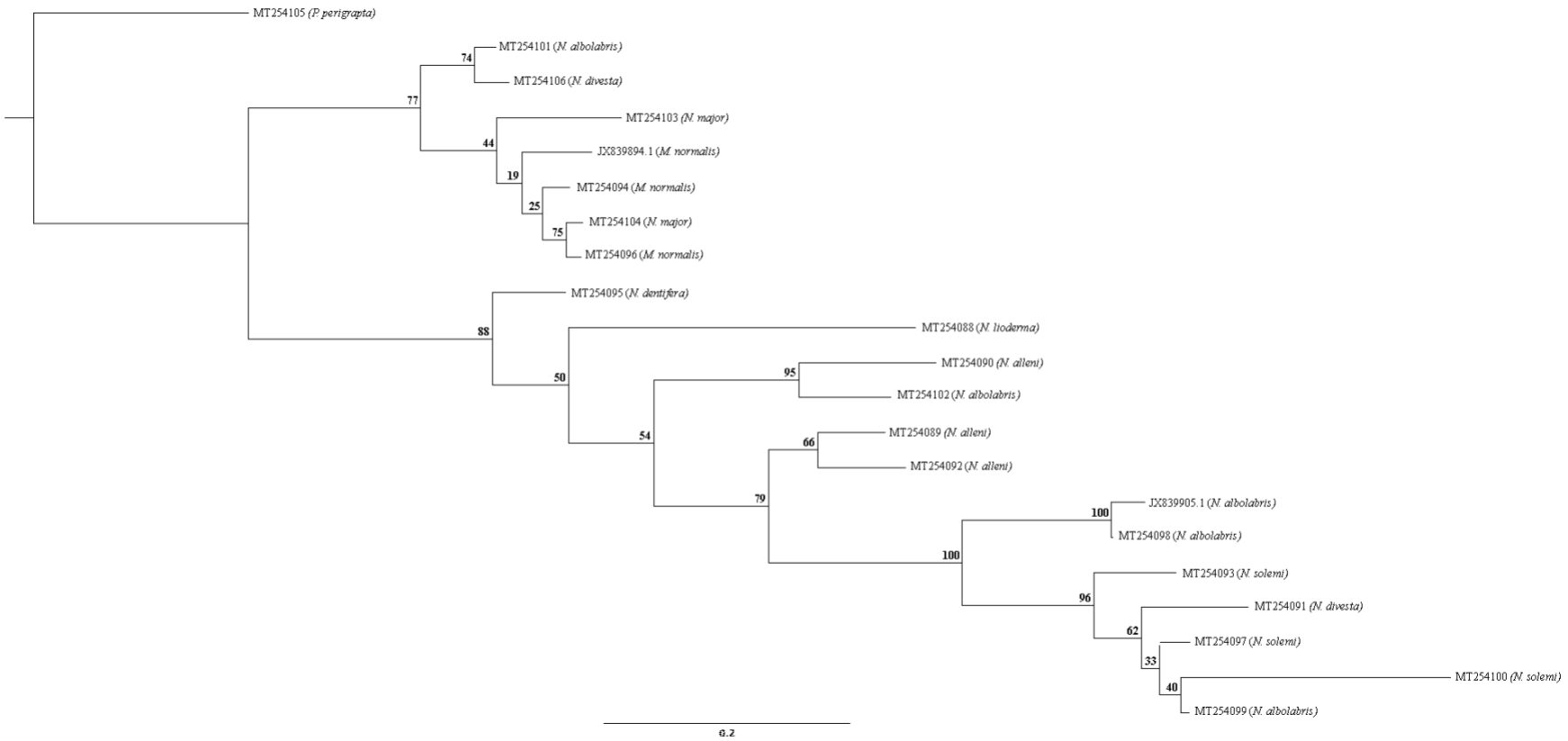

Figure 15. Maximum likelihood molecular phylogeny of locus 16S using HKY+I+G in RAxML-NG. Bootstrap (BS) values are displayed above nodes.

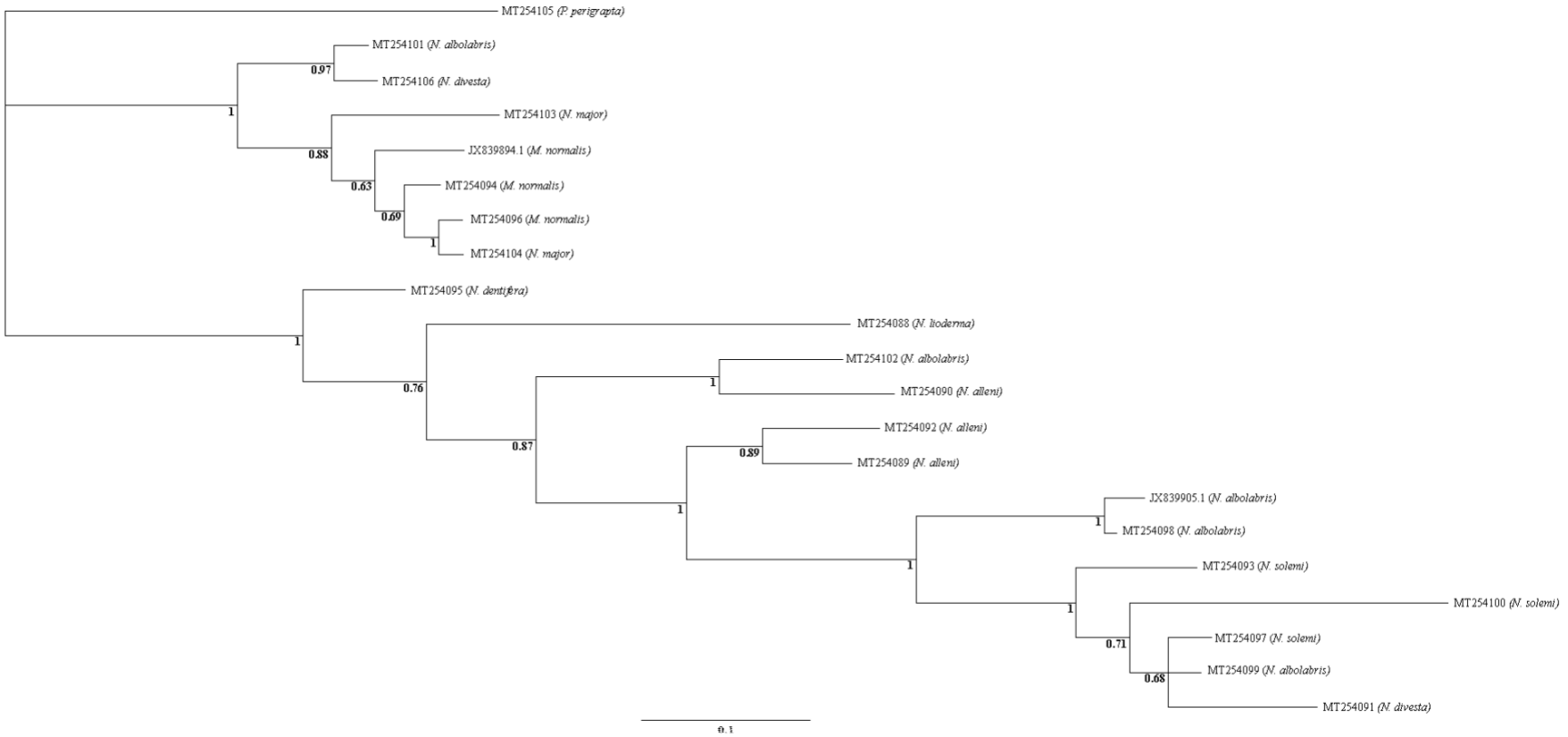

Figure 16. Bayesian inference molecular phylogeny of locus 16S in MrBayes. Posterior probabilities (PP) are displayed below nodes.

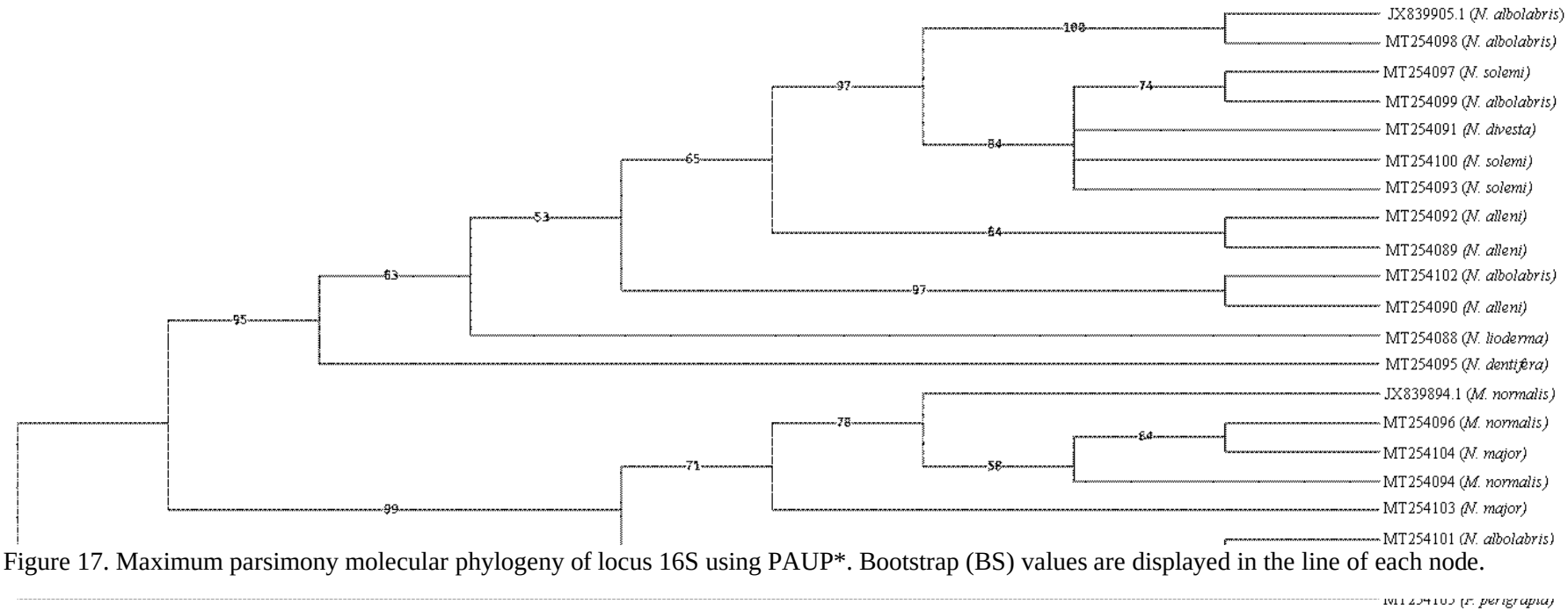

Figure 17. Maximum parsimony molecular phylogeny of locus 16S using PAUP*. Bootstrap (BS) values are displayed in the line of each node.

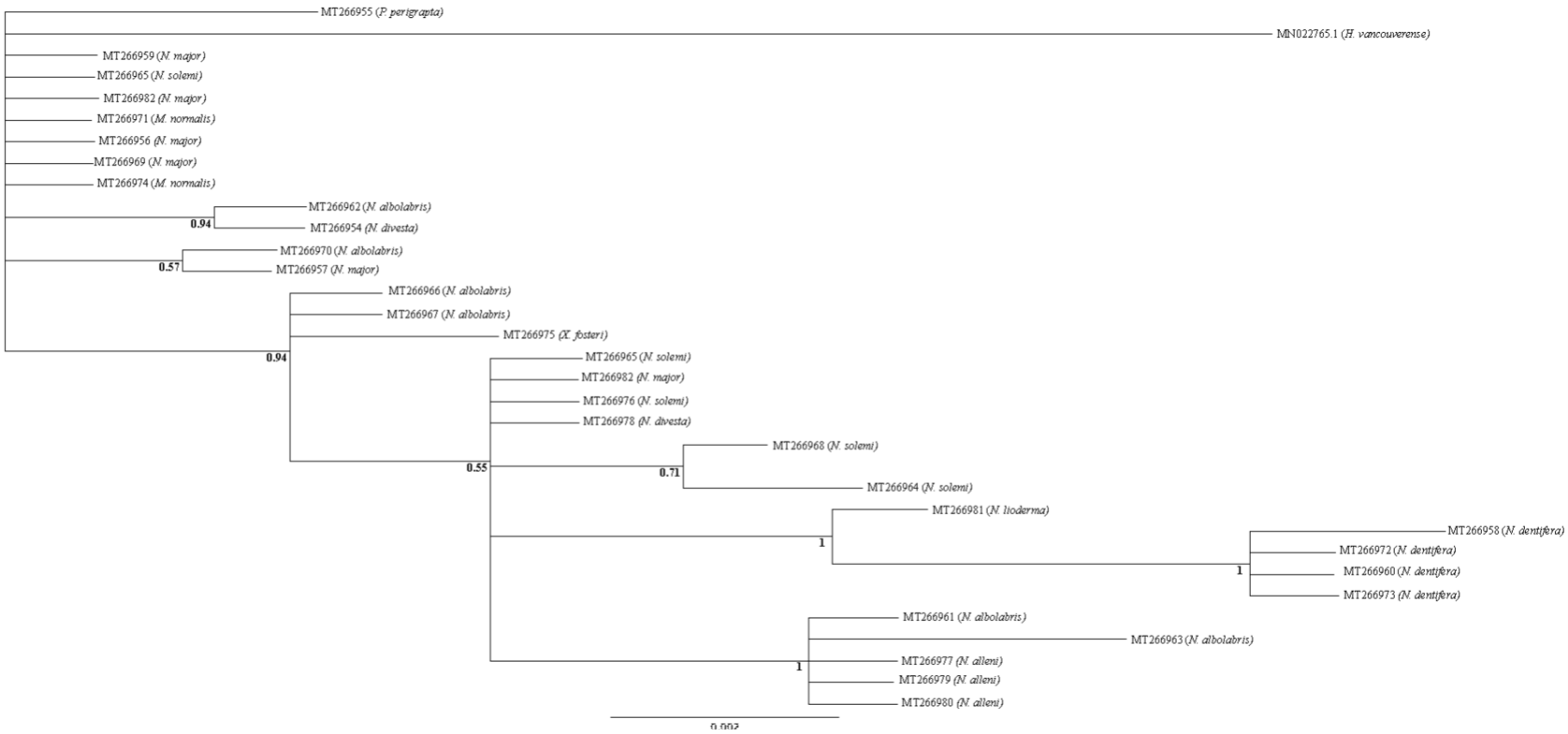

Figure 19. Bayesian inference molecular phylogeny of locus H3 in MrBayes. Posterior probabilities (PP) are displayed below nodes.

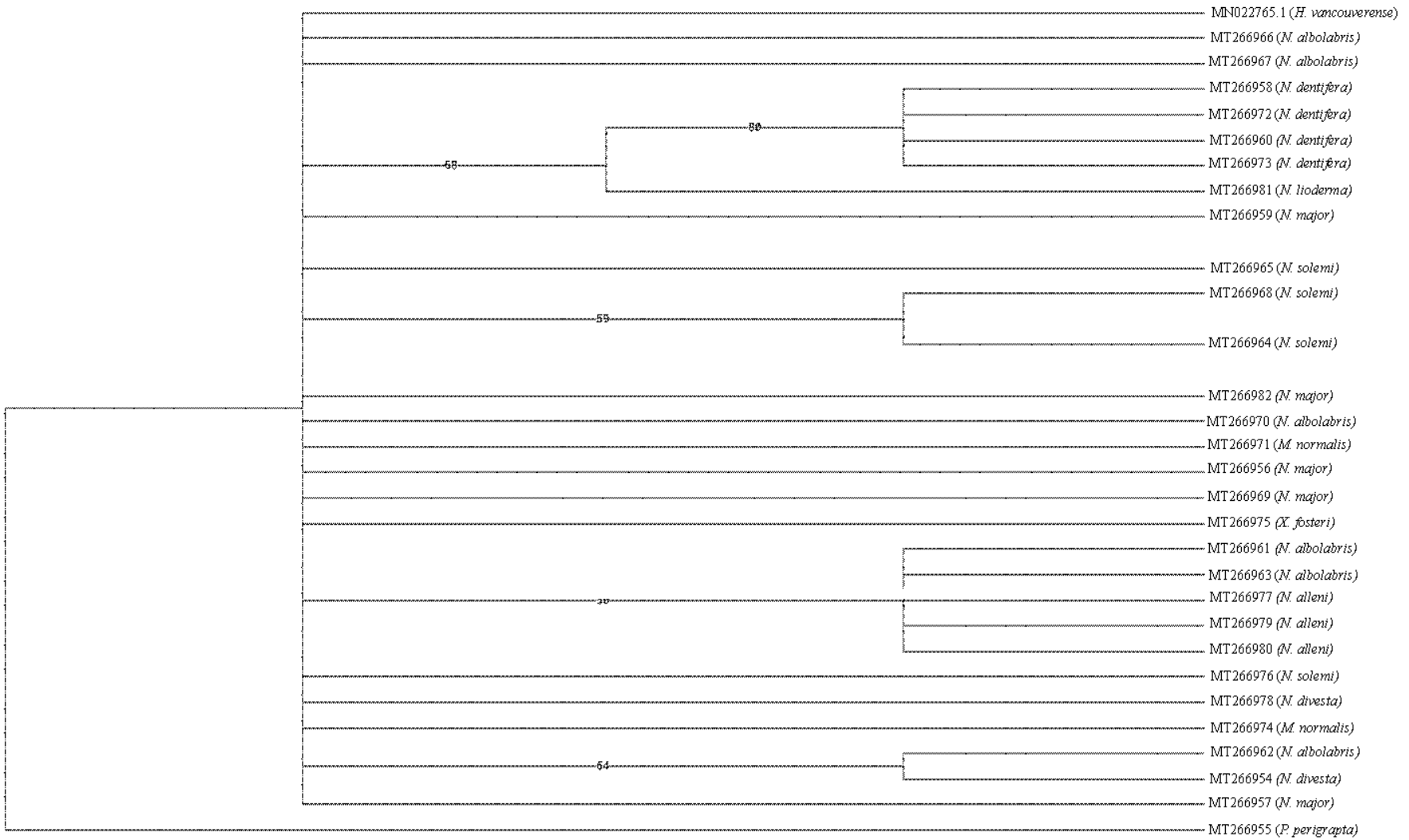

Figure 20. Maximum parsimony molecular phylogeny of locus H3 in PAUP*. Bootstrap (BS) values are displayed in the line of each node.

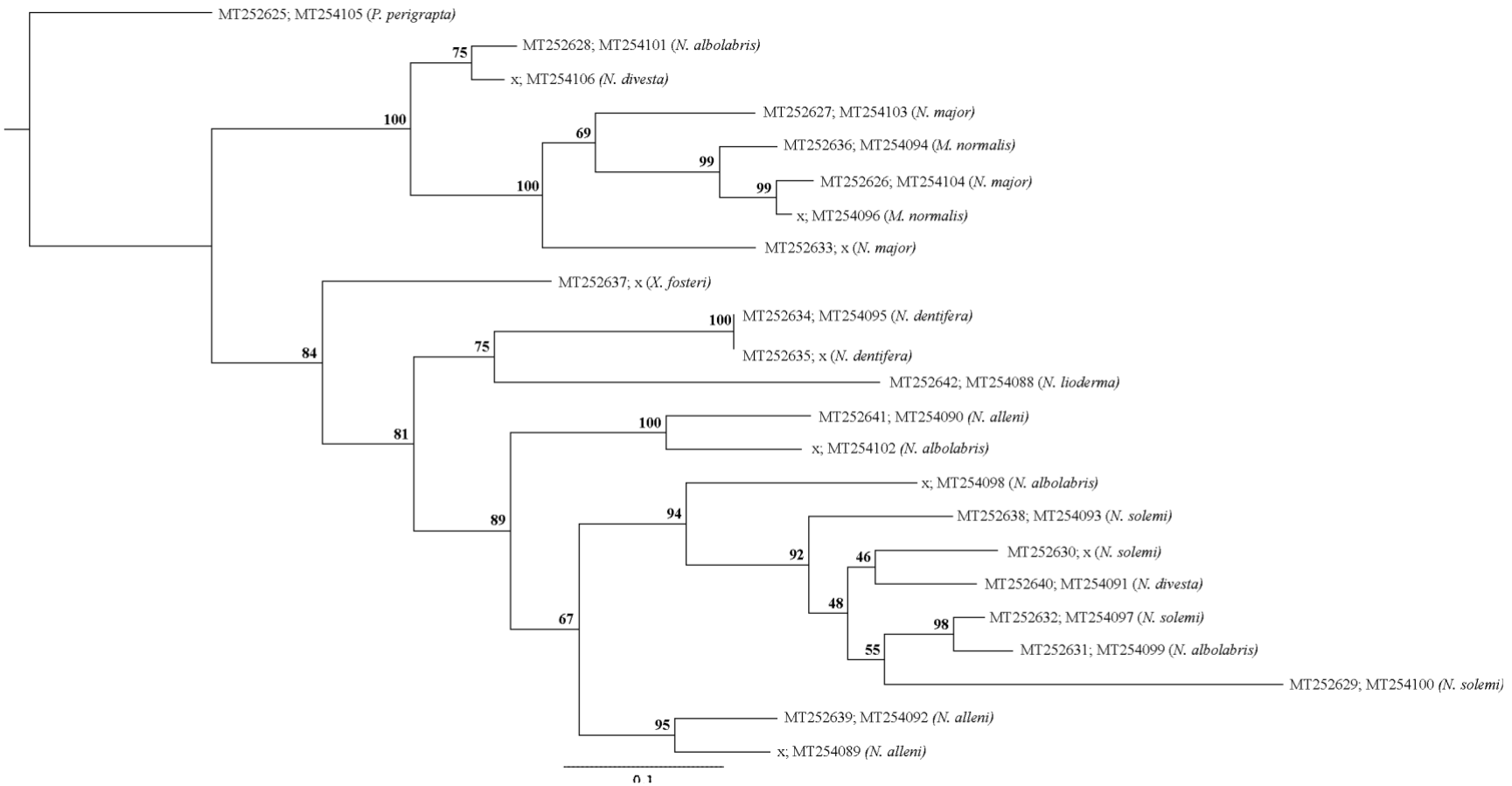

Figure 21. Maximum likelihood consensus phylogeny of the concatenated loci COI and 16 using TIM3+I+G and HKY+I+G, respectively, in RAxML-NG. Bootstrap (BS) values are displayed above nodes.

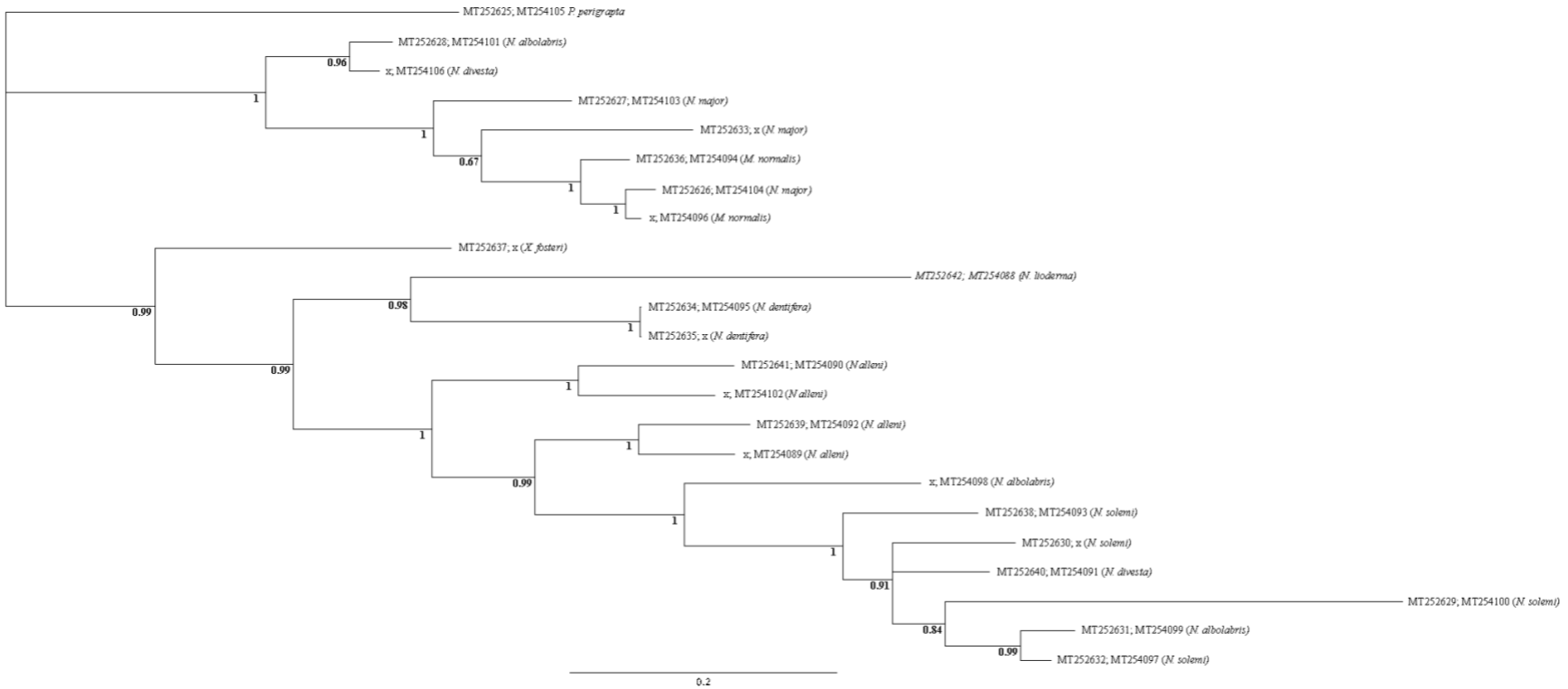

Figure 22. Bayesian inference consensus phylogeny of the concatenated loci COI and 16S in MrBayes. Posterior probabilities (PP) are displayed below nodes.

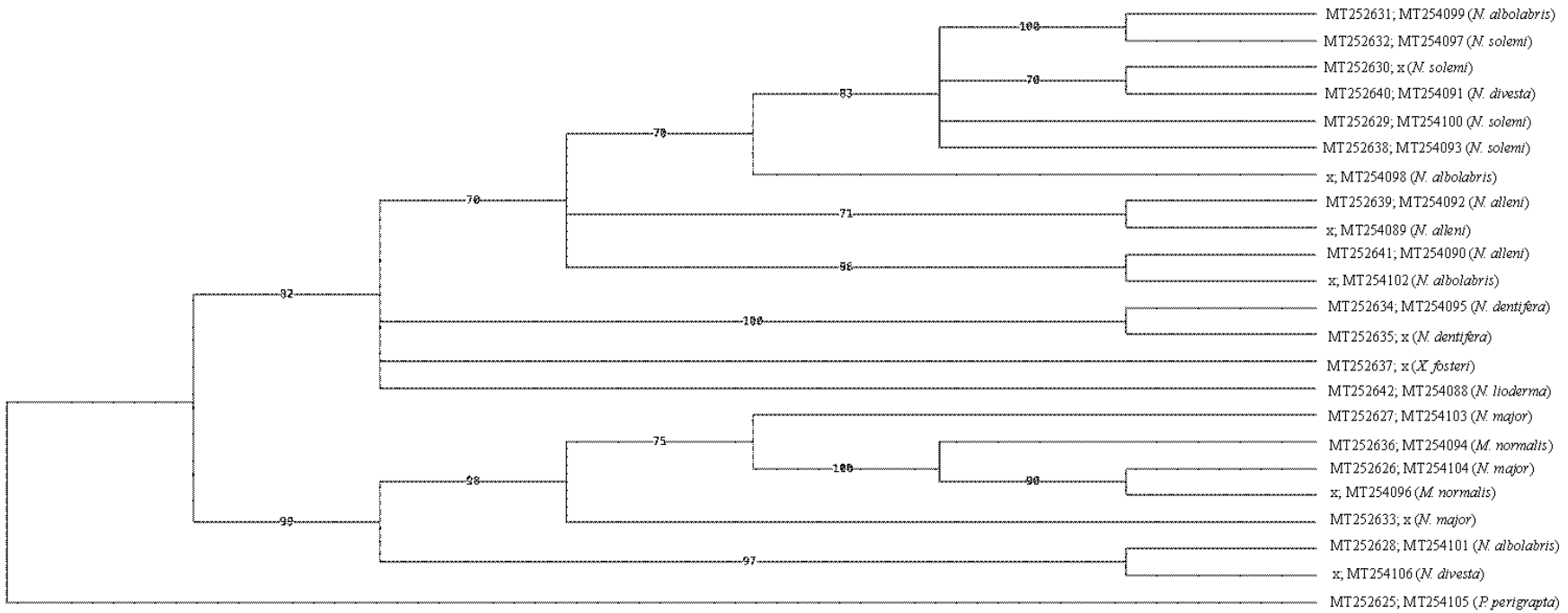

Figure 23. Maximum parsimony consensus phylogeny of the concatenated loci COI and 16S in PAUP*. Bootstrap (BS) values are displayed in the line of each node.

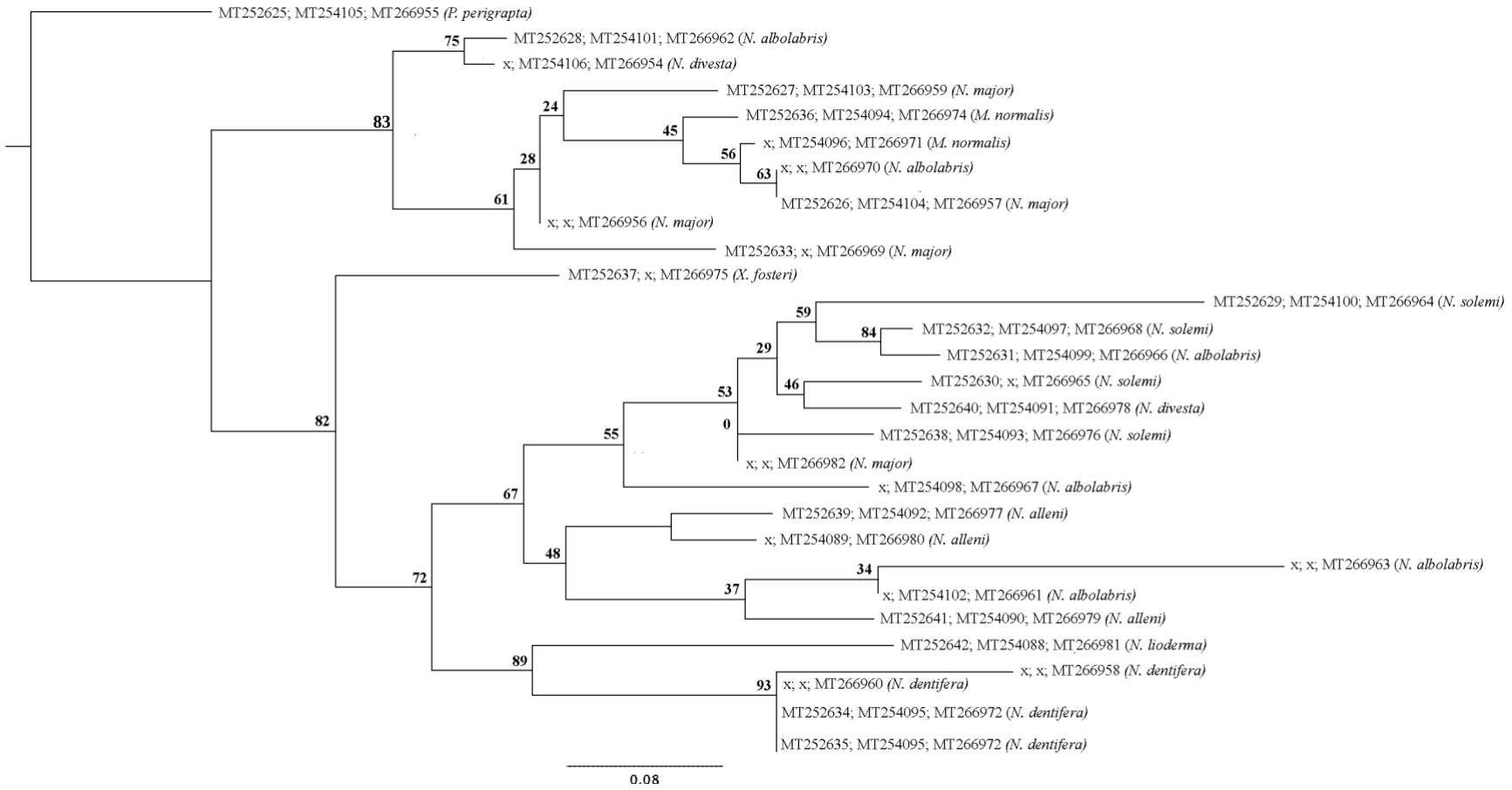

Figure 24. Maximum likelihood consensus phylogeny of the concatenated loci COI, 16S, and H3 using TIM3+I+G, HKY+I+G, and HKY+I, respectively, in RAxML-NG. Bootstrap (BS) values are displayed above nodes.

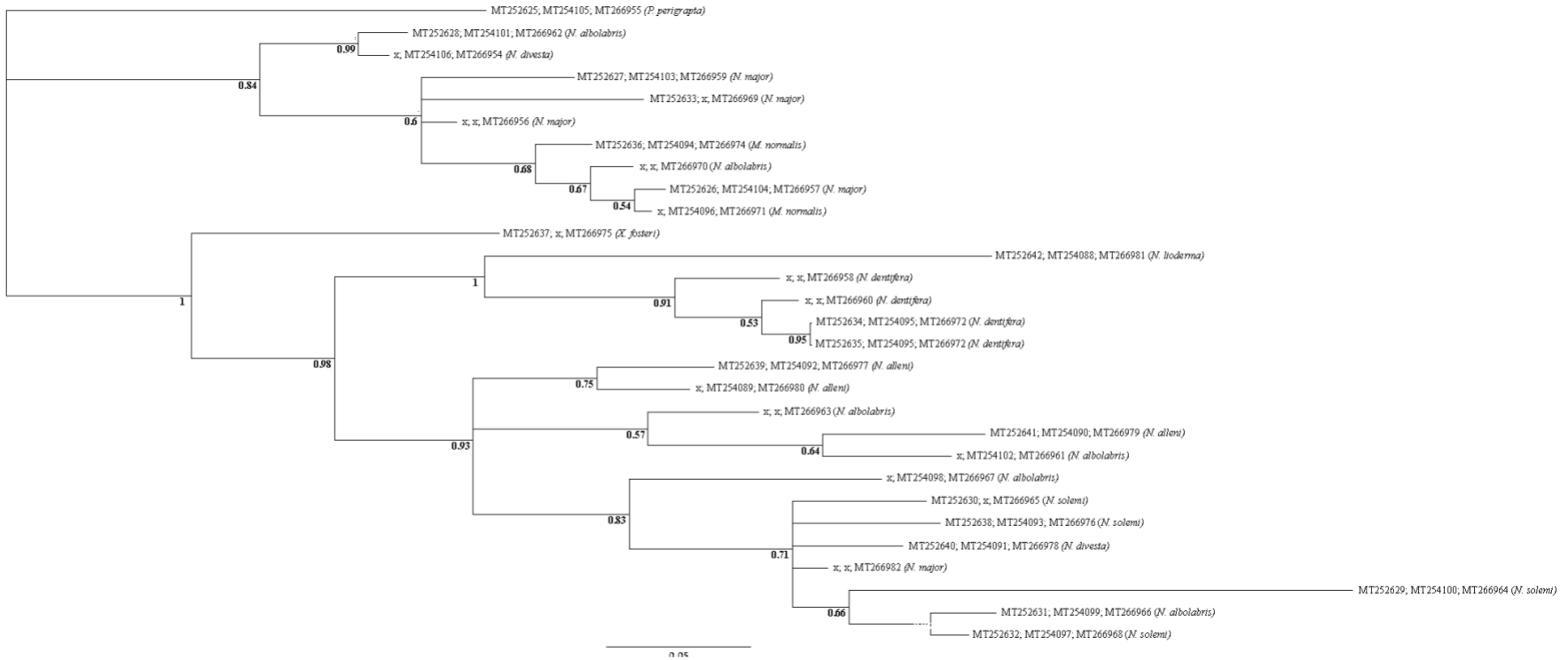

Figure 25. Bayesian inference consensus phylogeny of the concatenated loci COI, 16S, and H3 in MrBayes. Posterior probabilities (PP) are displayed below nodes.

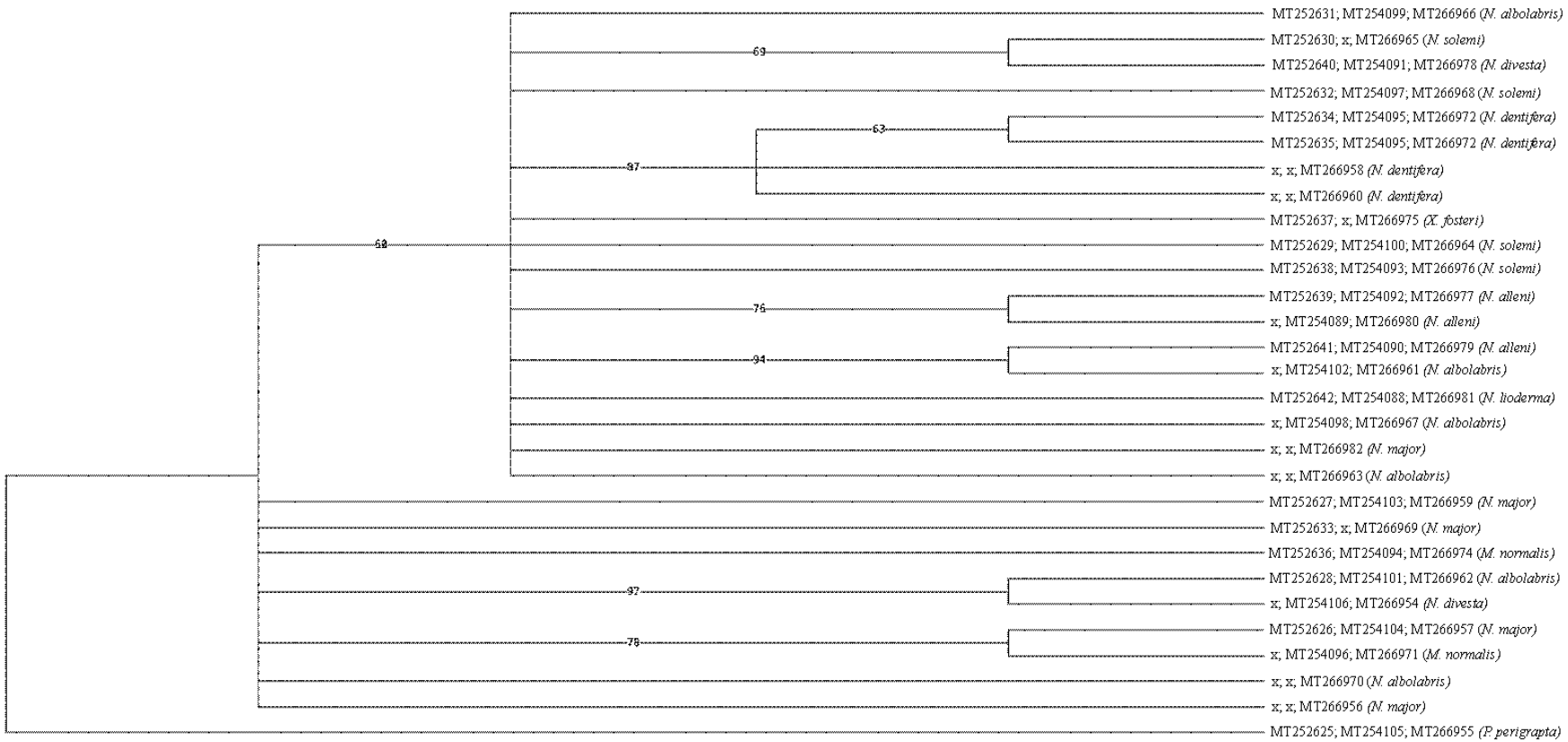

Figure 26. Maximum parsimony consensus phylogeny of the loci COI, 16S, and H3 in PAUP*. Bootstrap (BS) values above 50 are displayed in the line of each node.

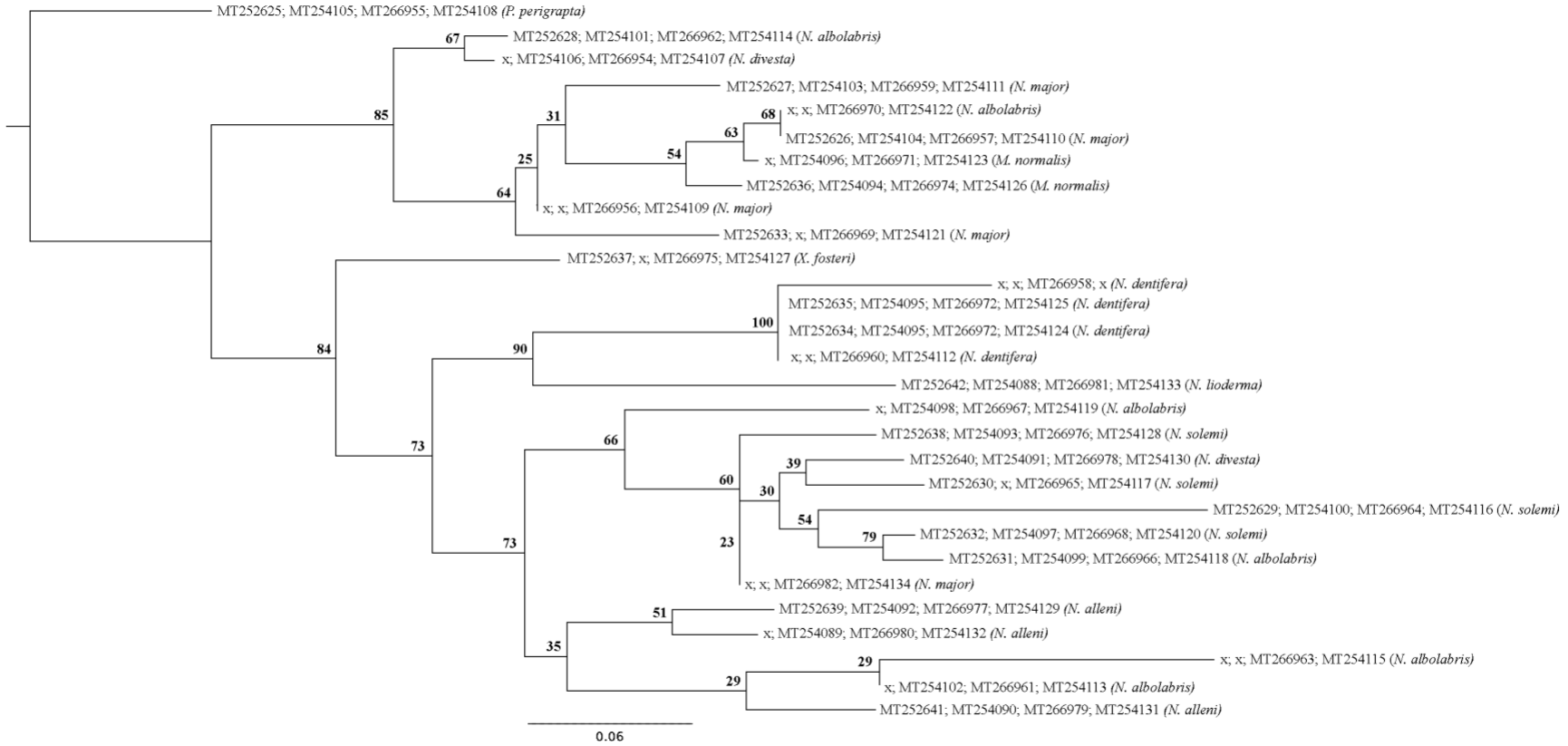

Figure 27. Maximum likelihood consensus phylogeny of the concatenated loci COI, 16S, H3, and 28S using TIM3+I+G, HKY+I+G, HKY+I, and JC, respectively, in RAxML-NG. Bootstrap values (BS) are displayed above nodes.

Vita

Amanda C. Wilkinson was born in Statesville, North Carolina to Mark and Wende Wilkinson. After graduating from East Lincoln Highschool in June 2013, she attended Appalachian State University to study Cell/Molecular Biology. In Spring 2017, Amanda graduated with her Bachelor of Science degree and accepted a position as a Soils Laboratory Technician at Engineering Consulting Services in Charlotte, NC. She volunteered as a collections associate at the Schiele Museum of Natural in Gastonia, NC and worked with land mollusks before returning to Appalachian in Fall 2018 to pursue her Master of Science in Cell/Molecular Biology. Amanda served as the Senator of Biology for the Graduate Student Government Association and sat on the Finance and Campus Concerns committees. She won the People's Choice Award during the 2018 3-Minute Thesis Competition and is an associate member of Sigma Xi, The Scientific Research Honor Society. Amanda C. Wilkinson was awarded her M.S. in August 2020.