INFERRING THE PHYLOGENETIC RELATIONSHIPS WITHIN TWO FISH GENERA: INSIGHTS INTO THE BIOGEOGRAPHY OF THE NORTHERN GULF OF MEXICO

A Thesis

by

ELIZABETH P. HUNT

BS, Bowling Green State University, 2012

Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in

FISHERIES AND MARICULTURE

Texas A&M University-Corpus Christi
Corpus Christi, Texas

August 2018

August 2018

INFERRING THE PHYLOGENETIC RELATIONSHIPS WITHIN TWO FISH GENERA: INSIGHTS INTO THE BIOGEOGRAPHY OF THE NORTHERN GULF OF MEXICO

A Thesis

by

ELIZABETH P. HUNT

This thesis meets the standards for scope and quality of
Texas A&M University-Corpus Christi and is hereby approved.

David S. Portnoy, PhD
Chair

Kevin W. Conway, PhD
Committee Member

Stuart C. Willis, PhD
Committee Member

Frank Pezold, PhD
Committee Member

August 2018

ABSTRACT

A suture zone can be defined as a geographic area where faunal assemblages meet and in which pairs of taxa (species, semi-species, etc.) have come back into secondary contact and may hybridize. While identifying these areas in terrestrial and freshwater environments has been an active area of research, less work has taken place in marine systems, in part, because barriers to dispersal are less apparent. In the northern Gulf of Mexico at least 15 putative sister taxa meet, with some evidence of hybridization, in an area roughly centered south of Mobile Bay (~88°W). This is consistent with the presence of a northern Gulf suture-zone, but data verifying the relationships between putative sister species are lacking. Therefore, in this study, phylogenetic relationships between species in the genera *Ogcocephalus* and *Sphoeroides*, were assessed using massively parallel sequencing of genomic libraries enriched for ultra-conserved DNA elements, with a focus on putative sister species distributed on either side of the hypothesized suture zone. Results of maximum likelihood and species tree analyses reveal multiple instances of speciation resulting in an eastern Pacific and western Atlantic species of Sphoeroides, as opposed to a single diversification event in either region. Further, *S. nephelus* (western Gulf) and *S. parvus* (eastern Gulf) were not resolved in a sister relationship, but instead formed a monophyletic group with *S. maculatus* (Atlantic) and the relationships between the three species and their distributions mirrors that of other taxa in this region. *Ogcocephalus parvus* (eastern Gulf and Atlantic) and *O. declivirostris* (western Gulf) were not resolved as sister taxa as previously described. Instead, *O. declivirostris* was consistently resolved in a group with *O. corniger* (eastern Gulf), although each species is not monophyletic. Relationships between *O. pantostictus*, *O. cubifrons*, and *O. radiatus* are also left unresolved as these three putative species also did not

form exclusive lineages. The discordance seen between species and phylogenetic placement in this genus could be a result of several different factors including erroneous classification, misidentifications, and/or insufficient lineage sorting between taxa, and further study is warranted.

DEDICATION

This thesis is dedicated to my mother, Michelle Hunt, for fostering my fascination with nature as a kid and for her incredible support through all my adventures.

ACKNOWLEDGEMENTS

I would like to thank my committee chair Dr. David Portnoy for all of his assistance and support throughout this project. On top of his professional expertise, his joyful nature and (sometimes) funny jokes made this journey a lot more enjoyable. I would also like to thank my other committee members, Dr. Stuart Willis, Dr. Kevin Conway, and Dr. Frank Pezold, for their input and guidance throughout my research. I feel that I need to give additional thanks to Dr. Stuart Willis for all of his time, effort, and patience while teaching me about phylogenetics and enduring months of trouble shooting protocols with me.

I would also like to thank all the lab mates I have had the pleasure of working with throughout my research. They were always a great resource when I was stuck and provided hours of valuable discussions about research and life. Outside of my research family, a huge thank you goes to all of my friends here at TAMU – CC.

Finally, I would like to thank all of the staff and volunteers that participate in the NOAA Groundfish survey each year for allowing me to sample with them and providing me with the fishes I needed to conduct my research. I would also like to thank José Tavera at the Universidad del Valle for sending me tissues and specimens for my research. A huge thank you to Dr. Tim Mousseau and the Japan Fisheries Education and Research team for supplying tissues. Thanks to Andy Bentley at the University of Kansa, Ben Frable and H.J. Walker at Scripps Institute of Oceanography, and Jeff Williams, Diane Pitassy and Carole Baldwin at the Smithsonian Museum of Natural History for gifting fish tissues to my research from their institutes ichthyology collections.

TABLE OF CONTENTS

CONTENTS PAGE

LIST OF FIGURES

FIGURE PAGE

LIST OF TABLES

TABLES PAGE

CHAPTER I: INTRODUCTION

Biogeography is an integrative science that utilizes data from various disciplines, such as ecology, genetics, and geology, to elucidate the underlying processes responsible for shaping species diversity and distribution (Parenti & Ebach 2009). Species distributions are shaped by dispersal, extirpation, and vicariance, while processes such as divergence and extinction drive changes in species diversity (Rosen 1978; Coyne & Orr 2004; Parenti & Ebach 2009). Dispersal describes the movement of lineages into new areas, thereby expanding or shifting their ranges (Myers & Giller 1988). After dispersing into new areas, individuals may occupy different habitats, feed on different resources, or through varying means become distinct from other conspecifics. Physical (distance) and ecological separation (species occupying different niches) can lead to cessation of gene flow and isolation, which with time may allow populations to accumulate genetic differences leading to speciation. By contrast, vicariant events occur when a barrier impermeable to individual dispersal arises within the range of a species, resulting in an immediate cessation of gene flow. Over time accumulated genetic differences may lead to new species (Myers & Giller 1988). The relative rates of speciation and dispersal (immigration) versus extirpation or extinction drive levels of biodiversity in a particular area (Parenti & Ebach 2009) and extirpation may explain discordant distributions of taxa which were once widespread (Parenti & Ebach 2009).

Vicariance zones are created by disruptive geological or environmental events creating barriers to gene flow that impact multiple taxa, examples include the rising of mountain ranges, the closing of seaways, or the presence of a thermo-haline barrier, among others (Rosen 1975; Hobbs, Frisch & Allen 2008; Grummer *et al.* 2015). These barriers may facilitate allopatric speciation because they immediately eliminate gene flow (Ronquist 1997). Allopatry, thought to

be the most common mode of speciation, is a direct result of geographic isolation between populations of a species and subsequent reproductive isolation may occur due to processes such as genetic drift, mutation, and selection that shape local genetic variation (Coyne & Orr 2004).

Occasionally, physical barriers that previously isolated biotic assemblages, are eliminated, creating areas of secondary contact or 'suture zones' (Remington 1968). Several outcomes are possible for populations meeting in suture zones, including reinforced speciation, the formation of stable clines, exclusion, or homogenization, (Remington 1968; Taylor *et al.* 2005; Quenouille *et al.* 2011; Moore, Merges & Kadereit 2013). Reinforcement describes a phenomenon wherein populations in secondary contact develop stronger and/or earlier-acting reproductive isolating barriers as a result of natural selection (Coyne & Orr 1998; 2004). Clines are commonly observed where secondary contact has occurred, and species experience spatially-restricted patterns of admixture. In a cline species gradually transition from one into the other across their combined range, with or without hybridization in areas of overlap (Remington 1968). Exclusion between populations, generally due to strong competition between species with effective reproductive isolation, could lead to static distributions, reduction of the range of one species and expansion of the other, or complete extinction/extirpation of one species (Bulgarella *et al.* 2013). Finally, genetic homogenization of populations or incipient species after secondary contact may occur because reproductive isolation is insufficiently developed and gene flow between groups is high enough to erase any differences that developed in isolation. Therefore, without additional reinforcement, formerly allopatric distributions do not guarantee speciation (Coyne & Orr 2004).

One of the signatures of historical vicariant events effecting multiple lineages and the presence of contemporary suture zones reflecting shared vicariance is genealogical concordance across taxa (Avise, 2000; Coyne & Orr 2004; Whinnett *et al.* 2005). Genealogical concordance

can be defined as an agreement in phylogenetic relationships across co-distributed lineages, implying that a shared historical biogeographic event is responsible for shaping the observed phylogenies (Avise 2000). These phylogenetic hypotheses can then be used to compare speciation events with known historical environmental/geologic events to better understand processes that generate current biogeographic patterns (Rosen 1978).

Suture zones are well documented in terrestrial environments because barriers to gene flow (current or historical) are easily observed and species distributions are well characterized (Remington 1968; Hewitt 2000). While identifying these areas in terrestrial and freshwater environments has been an active area of research, barriers to dispersal in marine systems are less apparent (Hobbs *et al.* 2008) and marine suture zones are understudied. Nonetheless, recent investigations have documented several suture zones in marine systems. Observations of hybridization between sister-taxa in the Indo-Pacific (Hobbs *et al.* 2008; Hobbs & Allen 2014; DiBattista *et al.* 2016) has led to the description of a suture zone around the Christmas and Cocos Islands, where 15 hybrid crosses, eight of which have been confirmed using genetics, have been observed among chaetodontid and acanthurid fishes (Hobbs & Salmond 2008; Hobbs & Allen 2014).

In the northern Gulf of Mexico (hereafter Gulf) another putative suture zone has been described (Figure 1; Dahlberg 1970; McClure & McEachran 1992). This area ranges longitudinally from Apalachee Bay, Florida (~84°W) to the Chandeleur Sound, Louisiana (~89°W), centered around Mobile Bay, Alabama (~88°W; Portnoy & Gold 2012). Portnoy & Gold (2012) list 15 pairs of putatively closely related taxa of fishes and invertebrates that meet in this zone, with several of the pairs, including searobins in the genus *Prionotus*, exhibiting apparent genetic admixture. Several vicariance events have been proposed to explain the

biogeographic patterns observed in the northern Gulf. The first is the drainage of the Tennessee River basin directly into the Gulf, ~2.4 million years ago (Figure 2; Simpson 1900; Ginsburg 1952), that would have caused large amounts of freshwater to be drained through Mobile Bay and perhaps creating a thermohaline barrier to dispersal of sensitive marine species. Another barrier may have been created by the strong unidirectional currents of the Suwanee Straights, flowing from the Gulf into the Atlantic Ocean through the Okeefenokee Trough which separated the continental United States from the Florida Peninsula (Figure 3; Pinet & Popenoe 1985). This feature would have closed ~1.75 million years ago (Bert & Harrison 1988). Finally, this putative vicariance zone may have been caused by periods of extended cooling during the early Pleistocene glaciations (~700-135 thousand years ago), which could have forced marine species into glacial refugia (Figure 4; Dahlberg 1970).

Portnoy and Gold (2012) suggested that while any of these vicariance events could be responsible for the divergence patterns observed in the northern Gulf, estimated divergence times of some species are not consistent with any of these events. For example, the estimated time since divergence between populations of lane snapper (*Lutjanus synagris)* in the eastern and western Gulf is between 20-50 thousand years ago. An alternative hypothesis proposed by Portnoy and Gold (2012) is the recession of the Laurentide Ice Sheet, which would have sent floods of cold, fresh water down the Mississippi River and into the northern Gulf, lasting thousands of years, and may have been an important vicariance event (Figure 5).

While physical barriers may no longer be present, current habitat differences in the northern Gulf may still be reinforcing separation (Hollenbeck 2016). For example, two sediment types characterize the northern Gulf, with terrigenous sediment dominant in the west and biogenic carbonate sediments dominant in the east (Balsam & Beeson 2003). Taxa that are

better suited to a specific habitat type may not successfully disperse into habitats in which they are less fit (Coyne & Orr 2004). Other ecological factors such as depth preference, salinity tolerance, migration patterns, and timing of reproduction may be important reinforcers of reproductive isolation and historical distributions (Coyne & Orr 1989; Palumbi 1994; Rundle & Schluter 1998).

To understand processes creating the patterns of biodiversity in the northern Gulf, it is necessary to test for genealogical concordance between other putative sister taxa in this zone. Portnoy and Gold (2012) list 15 sister species pairs distributed on either side of the suture zone; this study aims to investigate relationships between three of those pairs by generating phylogenomic hypotheses for two genera of fishes, the *Ogcocephalus* batfishes and the *Sphoeroides* pufferfishes, that show their greatest diversity in the western North Atlantic. For these genera relationships at the species level are not well understood and the relationships of these putative sister taxa unknown.

Classification of the Genus Ogcocephalus

Ogcocephalidae is a family of lophiiform fishes commonly referred to as batfishes and is comprised of ten genera (*Coelophrys, Halieutopsis, Dibranchus, Halieutaea, Halicmetus, Malthopsis, Halieutichthys, Ogcocephalus, Solocisquama*, and *Zalieutes)* with 78 described species (Bradbury 1999; 2003; Ho, Roberts & Shoa 2013; Nelson, Wilson & Grande 2016). These fishes inhabit benthic, marine habitats on the outer continental shelf and continental slope and can be found circumglobally in tropical and subtropical climates, with most species occurring at depths 1,500-3,000 meters (Nelson 2006), although, there are also species inhabiting shallow coastal waters (Ho, Chakrabarty & Sparks 2010; Derouen *et al.* 2015). This family has several synapomorphies including a dorsoventrally flattened body, the ability to "walk" on leg-

like pectoral and pelvic fins, the presence of an illicium and esca which are retracted into the illicial cavity, gill openings in or above the pectoral fin base and the presence of conical tubercles and bucklers (spines arranged in a radiating pattern) covering the body (Bradbury 1967; Nelson 2006). Classifying genera and species within Ogcocephalidae is difficult because meristic and morphological features (e.g. rostrum length) have proven to be problematic (Bradbury 1967). The genus *Ogcocephalus* is no exception and problematic diagnostic characteristics have led to debates of species validity and misidentifications (Bradbury 1980).

The genus *Ogcocephalus* contains 13 named species, most of which are found in the western Atlantic and Gulf, with two island endemic species found in the eastern Pacific (Garman 1899; Hubbs 1958). Relationships between the species in the genus *Ogcocephalus* have only been examined using meristic and morphological characteristics, with little molecular data available for most species in the genus (e.g. Derouen *et al.* 2015). A single study investigated the molecular relationships of species in Ogcocephalidae, and this study was conducted across genera and only focused on mitochondrial DNA (mtDNA), with only one species from the genus *Ogcocephalus, O. declivirostris*, included. (Derouen *et al.* 2015). Further investigation into the relationships within batfishes will greatly benefit from additional molecular analysis, especially within genera.

Many questions about the genus *Ogcocephalus* could be addressed using a phylogenomic approach. First, phylogenomic data may better resolve the relationships between species that have been difficult to discern from morphological data alone. Importantly, these data can be used to elucidate the relationship between two pairs of putative sister taxa in the northern Gulf (*O. cubifrons* and *O. pantostictus*, *O. parvus* and *O. declivirostris)* distributed on either side of the hypothesized suture zone (Bradbury 1980; 2003; Portnoy & Gold 2012). In addition,

phylogenomic data may clarify the status of *O. radiatus. Ogcocephalus radiatus* was first described by Mitchill (1818), but the description was vague and many of the characteristics described are shared among other *Ogcocephalus* species. Unfortunately, the holotype has been lost over time, precluding direct morphological comparison, and it has been repeatedly speculated that *O. radiatus* may be a synonym of *O. cubifrons* (Longley 1941; Bradbury 1980). Although *O. radiatus* is not included in Bradbury's (2003) list of *Ogcocephalus*, the species name is still used in some circles (Brooks, *et al.* 2005) and may add to the confusion over the identity and distribution of species in this genus. Finally, this study will add to our understanding of patterns of diversification of fishes in the western Atlantic. It has been documented that events such as the rising of the Isthmus of Panama (Wellington & Robertson 2001; Floeter *et al.* 2008; Thacker 2017) and the closing of the Tethys seaway (Hrbek & Meyer 2003; Floeter *et al.* 2008) have shaped the diversity of fishes throughout the Atlantic, but past mechanisms shaping diversification in the northern Gulf are less well understood (Portnoy & Gold 2012).

Classification of the Genus Sphoeroides

Sphoeroides is a genus of pufferfishes in the family Tetraodontidae (subfamily Tetraodontinae). The family includes two subfamilies and 196 species in 26 genera (Carpenter & De Angelis 2002; Matsuura 2015). Characteristics that unite the family include: fusion of the teeth into four tooth plates, lack of fusion between the premaxillae and dentaries at the midline, absence of ribs and epineurals, scaleless body with small lappets (a small skin flap) or prickles, and an ability to inflate the body (Nelson 2006). The subfamily Tetraodontinae includes most of the species in the family Tetraodontidae, encompassing 159 species and 25 genera (Roberts 1998; Britz & Kottelat 1999; Saenjundaeng, Vidthayanon & Grudpun 2013; Nelson *et al.* 2016). This subfamily is united by a conspicuous lateral line (usually), a round body in cross section, one or

two conspicuous nostrils on each side of the head, and a gill opening usually proceeding below the midpoint of the pectoral fin (Nelson 2006).

The distributions of *Sphoeroides* species are mostly in the western Atlantic or eastern Pacific, with the exception of *S. pachygaster*, which is distributed circumglobally in tropical and subtropical waters, and *S. marmoratus*, which is endemic to the eastern Atlantic (Walker & Bussing 1996; Carpenter & De Angelis 2002). Due to their unique color patterns, most species of *Sphoeroides* are easily identifiable. Most species in the genus inhabit shallow waters in bays, estuaries, and reef areas. Once again, the exception lies with *S. pachygaster*, which is known to be found at depths as great as ~400 meters (Carpenter & De Angelis 2002).

Relationships between species in this genus have been assessed primarily through morphometric and meristic data (Shipp 1970; Carpenter & De Angelis 2002), with limited molecular data available (but see Yamanoue *et al.* 2008; Amaral *et al.* 2013; Santini *et al*. 2013). The few molecular studies have focused primarily on family relationships within the order Tetraodontiformes, with little emphasis on relationships. Although Amaral *et al.* (2013) primarily investigated species in the genus *Colomesus*, their study includes six species (of 22) in the genus *Sphoeroides,* providing the most comprehensive molecular data for this genus currently. Interestingly, Amaral *et al.* (2013) found that *Colomesus* is nested within *Sphoeroides*, making *Sphoeroides* paraphyletic.

The current study focuses on the relationships between putative sister species *S. parvus* and *S. nephelus*. *Sphoeroides parvus* is found primarily in the northwestern and southern Gulf but is present as far east as northern Florida (Carpenter & De Angelis 2002). *Sphoeroides nephelus* is also found in northern Gulf but from the Gulf coast of Florida, off the Atlantic coast of Florida, and throughout the Caribbean, though it is absent in the western Gulf (Carpenter &

De Angelis 2002). These taxa overlap in the area of the hypothesized suture zone. We also investigate the relationship of *S. pachygaster* to other *Sphoeroides* species. Since *S. pachygaster* is found circumglobally and in more oceanic habitat than other *Sphoeroides* species, samples from across the species range may provide insight into population structure of the species.

Phylogenetic Inference

Traditionally, molecular phylogenetic studies have typically been conducted using mitochondrial DNA (mtDNA) and/or a small number of nuclear loci, and those data were then concatenated to infer the organismal phylogeny or "species tree" (Degnan & Rosenberg 2009a; Lemmon, Emme & Lemmon 2012; Eytan *et al.* 2015; Gilbert *et al.* 2015). However, using only a few markers to resolve species trees can prove problematic, as discordance between gene trees can lead to poorly supported species trees (Degnan & Rosenberg 2009a). New genomic techniques have allowed for the observation and analysis of hundreds of molecular markers, and these "phylogenomic" approaches have been employed to more accurately resolve difficult phylogenies (Degnan & Rosenberg 2009a; Lemmon *et al.* 2012; Eytan *et al.* 2015). In particular, genomic techniques utilizing ultraconserved elements (UCEs) may allow for more thorough phylogenetic investigations of closely related groups of organisms without the need to develop taxon-specific genomic markers. UCEs are genomic loci with core regions that are conserved across evolutionary distant taxa (Bejerano et al. 2004). Ultra-conserved elements were originally described from mammalian genomes (Bejerano et al. 2004), and have since been utilized with great utility for phylogenetic reconstruction in fishes (Faircloth et al. 2013; Gilbert et al. 2015). Ultra-conserved elements can be divided into two sections, the highly conserved core region and the flanking regions, which often contain more variable sites (Faircloth et al. 2012; Gilbert et al. 2015). The core regions have shown to be informative for resolving deeper divergences, while

flanking regions may be informative for recent divergences, which makes them an attractive option for generating shallow and sometimes difficult topologies (Lemmon *et al.* 2012; Gilbert *et al.* 2015).

The goal of this study was to generate robust phylogenetic hypotheses for both *Ogcocephalus* batfishes and *Sphoeroides* pufferfishes using UCE data with concatenated, maximum likelihood and species tree analyses under different thresholds for missing data. The work provides information on relationships within two genera of fishes that diversified in the western Atlantic and eastern Pacific elucidates the relationships of putative sister species distributed across the northern Gulf suture zone to test patterns of genealogical concordance with similarly distributed taxa.

CHAPTER I: METHODS

Sampling

Phylogenetic analyses included eight of the 13 species in the genus *Ogcocephalu*s (72% taxon coverage), as well as one outgroup, the blackfin goosefish, *Lophius gastrophysus* (Table 1), and 14 of the 23 species in the genus *Sphoeroides* species (61% taxon coverage), as well as one outgroup, the smooth pufferfish, *Lagocephalus laevigatus* (Table 1). Samples from *S. pachygaster* were collected from the western Atlantic as well as from the Pacific, near Japan. All tissues were collected from vouchered specimens, either from specimens collected and identified in the field from various surveys for this project and placed in the Biodiversity and Research Collection at Texas A&M University in College Station, or from specimens housed at other institutions (Table 1).

Library preparation

Samples were processed using a modified in-solution hybrid enrichment protocol (Faircloth et al. 2012) targeting 500 UCE loci. Fin clips or muscle tissue from each specimen were stored in 95% non-denatured EtOH or DMSO-EDTA-NaCl buffer (Seutin *et al.* 1991) before DNA extraction, and total genomic DNA extracted using a phenol-chloroform extraction protocol (Sambrook *et al*. 1989). After extraction, DNA was standardized to 20 ng/µl in 100 µl of 1X TE buffer before random shearing to a length of ~800 bp (+/- 200bp) using a M220 Focused-ultrasonicator (Covaris). After sonication, samples were purified using Ampure XP (Agencourt) in a 0.6X concentration to eliminate DNA fragments less than ~400bp in length. Extractions were quantified with the AccuBlue High Sensitivity dsDNA Quantitation Kit (Biotium) to calculate concentrations and electrophoresed on a 2% agarose gel to visually confirm the size distribution of DNA fragments. Samples were subsequently end repaired and universal double-stranded oligonucleotide adapters were blunt-end ligated using the NEB Ultra II DNA Library Prep Kit for Illumina (New England Biolabs). Samples were amplified using 15 cycles of polymerase chain reaction (PCR) using the 'Adapterama' iTru5 and iTru7 indexed PCR primers (Glenn *et al.* 2016). Fifty microliter PCR reactions contained water, 15 uL ligated DNA, NEBNext Ultra II Q5 Master Mix, and 50 pmol of both iTru5 and iTru7 primers. PCR cycle conditions were as follows: initial denaturation at 98°C for 45 seconds followed by 18 cycles of 98°C for 15 seconds, 60°C for 30 seconds, and 72°C for 1 minute and 15 seconds, followed by a final extension at 72°C for 5 minutes. Indexed libraries underwent target enrichment of UCEs using a set of 2,001 baits, designed to capture 500 UCE loci in actinopterygian fishes (MYbaits_Actinopts-UCE-0.5Kv1, MYcroarray). Enrichment was followed by post-hybridization amplification using 16 cycles of PCR. Fifty microliter PCR

reactions contained 1X Phusion High Fidelity buffer, 1.5 mM $MgCl_2$, 0.25 mM of each dNTP, 25 pmol of both P5 and P7 primers, 2.0 U/uL of Phusion Taq, and 15 uL DNA. PCR cycle conditions were as follows: initial denaturation at 95°C for 5 minutes, then 30 seconds at 98°C, followed by 16 cycles of 98°C for 15 seconds, 65°C for 30 seconds, and 72°C for 1 minute and 30 seconds, followed by a final extension at 72°C for 5 minutes. To ensure successful UCE enrichment, libraries underwent quantitative PCR (qPCR) using GoTaq qPCR master mix (Promega) and a StepOnePlus Real-Time PCR System (Applied Biosystems, inc.). Twenty-five microliter qPCR reactions contained water, GoTaq qPCR Master Mix, 25 pmol of qPCR primer, and 2 uL of DNA standardized to 1 ng/uL, for a total reaction volume of 25 uL. qPCR conditions were as follows: initial denaturation at 95°C for 5 minutes, followed by 40 cycles of 95°C for 10 seconds, 60°C for 20 seconds, and 72°C for 30 seconds, followed by a final a cooling step at 40°C for 10 seconds. Enrichment was considered successful when there was at least a 50X enrichment of target product as compared to pre-enriched enriched libraries (Faircloth *et al.* 2012). Any libraries which did not meet the 50X enrichment threshold were re-enriched following the aforementioned protocol. A first group of 22 libraries were quantified, standardized, and pooled for paired-end (2 x 300 bp) sequencing on an Illumina MiSeq (Illumina, San Diego, CA). The resulting MiSeq sequences were used to create a reference of UCE loci and subsequent sequences were mapped to this reference. This first set of 22 libraries, plus 46 additional libraries, were pooled together and sequenced (2 x 150 bp) on an Illumina HiSeq 4000 (Illumina, San Diego, CA) by Genewiz, Inc.

Reference Assembly and UCE Identification

Demultiplexed sequences from the MiSeq were trimmed for quality and adapter contamination using Trimmomatic (Bolger *et al*. 2014) and contigs assembled using Trinity

(Grabherr *et al.* 2011). Contigs were matched to UCE loci using a modified pipeline including PHYLUCE (Faircloth 2015). The PHYLUCE pipeline utilizes LASTZ to align assembled contigs to UCE probes, normally discarding all contigs that are duplicate assemblies of the same locus in an effort to avoid paralogs (Faircloth *et al.* 2012), but this results in significant missing data due to the use of assemblers designed for transcriptomic data (i.e. Trinity). Here assembled contigs were clustered at an 80% similarity threshold using CD-hit to reduce redundancy (Li & Godzik 2006), then re-introduced to PHYLUCE for identification by probe similarity. For each genus, a database was created that contained records of UCE loci present across individuals. FASTA sequences for each locus were extracted using the PHYLUCE probe.matches.sqlite database. For each locus, a multiple sequence alignment was created using MAFFT (Katoh & Standley 2013), alignments checked by eye, and a consensus sequence of the alignment created using EMBOSS (Rice, Longden & Bleasby 2000). The final consensus sequences for each locus were concatenated into a single reference for read mapping using Bowtie2.

Demultiplexed HiSeq reads, as well as the previous MiSeq reads were mapped directly to the MiSeq references for both *Ogcocephalus* and *Sphoeroides*, respectively, to identify UCE loci. After UCE loci were identified in each individual, multiple sequence alignments per locus were generating using MAFFT and using the L-INS-i algorithm. Alignments for each locus were trimmed using the program trimAl with a 30% gap threshold to eliminate poor-quality and information-deficient bases at the ends of alignments (Capella-Gutierrez *et al.* 2009). Final datasets were filtered for different thresholds of missing data using the phyluce_align_get_only_loci_with_min_taxa script in PHYLUCE, resulting in five different datasets: complete (only includes loci shared among all individuals), ≤10% missing data allowed, ≤ 25% missing data allowed, and ≤ 50% missing data allowed, and incomplete). The incomplete

dataset includes any locus present across three or more of 41 individuals in the *Sphoeroides* analysis and three or more of 24 individuals in the *Ogcocephalus* analysis.

Analysis

Maximum-likelihood (ML) analysis was used to infer individual gene trees for each data set (complete, ≤ 10% missing, ≤ 25% missing, ≤ 50% missing and incomplete) for each genus. An estimate of phylogeny was estimated for a concatenated supermatrix partitioned by locus, as well as for each locus separately, using RAxML v8.02 (Stamatakis 2014). Trees were generated using the GTR + GAMMA substitution model in RAxML, as the program only employs variations on the GTR model, even with partitioned data. To determine node support, bootstopping in RAxML was used. This method computes bootstrap thresholds at the time of analysis and determines that enough bootstrap replicates have been generated for sufficient support when the difference in branch support is smaller than 3% for 99% of the permutations (Pattengale *et al.* 2010; Stamatakis 2016).

To generate species trees, tree files from the best ML gene trees for each locus were compiled into a single tree file and analyzed in a "summary coalescent" framework using ASTRAL-III (Zhang *et al.* 2017). Each individual was assigned a species designation for the ASTRAL inference. Simultaneous computation of gene trees and species tree is impractically computationally demanding for high numbers of loci, so summary methods, such as that implemented ASTRAL, are statistically consistent and appropriate for phylogenomic data when gene tree heterogeneity is due to incomplete lineage sorting rather than insufficient signal and phylogenetic uncertainty (Mirarab *et al.* 2014; Sayyari & Mirarab 2016; Nute *et al*. 2018). ASTRAL infers an unrooted species tree from a set of unrooted gene trees using a quartet-based method. Quartet-based summary methods partition each gene tree into unrooted trees containing

4 taxa, and ASTRAL then finds the species tree that is congruent with the largest number of quartet trees produced by the gene trees (Mirarab *et al.* 2014). ASTRAL-III tends to resolve the same tree topology and branch lengths as trees computed with Bayesian inference under the multispecies coalescent model (MSC) with both simulated data and empirical datasets (Mirarab *et al.* 2014). ASTRAL-III has become a popular method for analyzing genomic data due to its ability to accurately infer species trees in the presence of incomplete lineage sorting (ILS) and manage gene tree error and uncertainty by collapsing low-support branches (Zhang *et al.* 2017). Unlike other Bayesian based species tree estimations, support in ASTRAL is represented as the *local posterior probability* (LPP), calculated by the frequency of gene tree quartets (Sayyari & Mirarab 2016). Tree topology and node support were compared quantitatively across the concatenated and species trees for each dataset to evaluate the effects of missing data as well as patterns of concordance between genera.

CHAPTER I: RESULTS

After sequence capture there were 436 and 459 loci shared among a minimum of three individuals for *Sphoeroides* and *Ogcocephalus,* respectively, which were used for the incomplete dataset. A total of 20 and 205 loci were shared across all individuals of *Sphoeroides* and *Ogcocephalus*, respectively. Final alignment information, including number of loci, mean alignment length, number of variable sites, and number of parsimony informative sites (PI) for each dataset are reported in Table 2. Overall, *Ogcocephalus* datasets contained a higher number of loci with shorter average alignment lengths and fewer variable sites as compared to the *Sphoeroides* datasets (Tables 2a and 2b). The frequency of variable sites locations also follows the expected UCE distribution: the core region of the UCE shows lower variability compared to the flanking regions, where more variable sites are found (Figure 5).

Phylogenetic Analysis of Sphoeroides

The topology of phylogenetic hypotheses generated using ML analysis on the concatenated datasets was almost completely consistent across all levels of missing data (ranging from complete to incomplete) and recovered the same six clades. Most nodes were well resolved (ML bootstrap > 85%), with several nodes near the tips having moderate to low support (Figure 7 A-E). Node support for branches that were not fully resolved in the complete dataset varied across phylogenies inferred from different levels of missing data. In most cases, node support was highest when there was only 10% of data missing and then decreased as the amount of missing data increased. This may be indicative of some underlying issues with the loci contained in datasets characterized by more missing data and warrants further investigation. Within the *Sphoeroides*, the *S. pachygaster* clade (Clade A) was always fully supported (ML bootstrap = 100%) and is the first to diverge among sampled species. The next node in the tree divided the species into two groups. In the complete dataset the division fell between a clade containing *S. greeleyi* and *S. tyleri* (Clade B) and its sister clade containing *S. lispus*, *S. annulatus* and *S. testudineus* (Clade C) and a clade containing the rest of the species. In the complete dataset this node was only moderately supported (ML bootstrap = 69%). In all other trees Clade B was sister to a clade containing all other species, including Clade C, and the clades and their relationships were also constant. While Clade B was consistently resolved, *S. greeleyi* and *S. tyleri* were not reciprocally monophyletic. Clade C contained three species: *S. lispus, S. annulatus*, and *S. testudineus*. Clade D contained three species: *S. dorsalis*, *S. maromotus* and *S. spengleri*. Clade E only contained *S. trichocephalus* and Clade F, contained *S. lobatus*, *S. parvus, S. maculatus* and *S. nephelus*. Bootstrap support for these clades was 100% through all analyses, though support within them varied across differing levels of missing data.

The species tree analyses yielded the same topologies as the concatenated ML analysis of the same data, with most of the nodes in the trees being fully supported. The one exception was that in clade F of the complete dataset the arrangement of *S. parvus* and *S. nephelus* differed compared to all of the incomplete data sets and ML trees. There was an increase in node support across datasets, with the complete dataset having fewer fully resolved nodes than incomplete datasets, and with the tree hypothesized from 10% missing data having the most fully resolved and well supported nodes, particularly towards the tips of the tree. (Figure 7 F-J). In the complete ML analysis, the relationship within *S. pachygaster* (western Pacific versus western Atlantic samples) did not for two distinct lineages, whereas in the other datasets western Pacific samples and western Atlantic samples formed separate monophyletic units.

Phylogenetic Analysis of Ogcocephalus

The topology of the phylogenetic hypotheses generated using ML analyses on the concatenated datasets were very similar across all levels of missing data (ranging from complete to incomplete) and recovered the same major clades, while relationships within clades varied dependent on the amount of missing data. All internal nodes were fully resolved (ML bootstrap = 100%), with nodes towards the tips having moderate to poor support (Figure 8A-E). The same fully supported nodes were recovered across all ML analyses of missing data. Node support for branches that were not fully resolved in the complete dataset varied across phylogenies inferred from different levels of missing data. For poorly resolved nodes towards the tips of the trees, there was not a single dataset that exhibited the highest support across all branches. Within *Ogcocephalus, O. darwini* (Clade A) was always resolved as sister to a clade containing all other *Ogcocephalus* species, and similarly, Clade B contained all *O. parvus* individuals and was sister to the remaining species. *Ogcocephalus vespertillio* was always recovered as sister to Clade C,

which was made up of two sub-clades: Clade C1, containing *O. declivirostris* and *O. corniger O. cubifrons, and Clade C2, containing O. pantostictus, and O. radiatus.* Relationships among taxa in both C1 and C2 were unclear and none of the species were resolved as monophyletic. Trees generated with a missing data consistently showed the same topology but were different from the tree from the complete dataset in the relationships within clades C1 and C2.

In the species tree analyses, the same clades and inter-clade relationships were recovered as the concatenated ML analysis (Figure 8F-I) with most nodes being fully supported (LPP = 1.0). Node support towards the tips was lowest in the complete analysis and increased with missing data. However, the species tree analysis was not any more successful than the ML analysis when trying to resolve species relationships within clades C1 or C2, likely because branch lengths between the species were extremely small.

CHAPTER I: DISCUSSION

Phylogenomics of Sphoeroides

Phylogenomic hypotheses generated using both concatenated ML and species tree approaches provide strong support for the inferred relationships, with several species relationships congruent with previous work using mtDNA (Amaral *et al.* 2013). Similar to Amaral *et al.* (2013), this study recovered *S. pachygaster* as sister to all other species in the genus. Individuals in the *S. pachygaster* clade form two lineages, one from the eastern Pacific and the other from the western Atlantic. This may indicate that these two lineages represent different species. Further range-wide sampling and molecular analyses within this species needs to be conducted.

Sphoeroides parvus and *S. nephelus* (Shipp 1970), the putative sister taxa distributed on either side of the hypothesized northern Gulf suture zone, were not resolved strictly as sisters

here. Instead, *S. parvus* (western Gulf) shares a common ancestor with the clade containing *S. nephelus* (Peninsular Florida) and *S. maculatus* (Atlantic coast excluding Florida). While *S. parvus* and *S. nephelus* were not sister species, the relationship between *S. parvus* and the other two species is congruent with other taxa in the northern Gulf and Atlantic (Carpenter & De Angelis 2002), including *Prionotus scitulus*, *P. martis,* and *P. carolinus (*Portnoy *et al.* 2017*)*. This suggests that divergence of Atlantic coast populations following a Gulf vicariance event may also be a common pattern among marine organisms along the U.S. coast. However, confirmation of which events were responsible for the separation of the linages leading to *S. parvus* and *S. nephelus/S. maculatus* will require time calibration of the inferred phylogeny.

Phylogenomics of Ogcocephalus

In contrast to the hypotheses generated for *Sphoeroides*, the relationships within *Ogcocephalus* were not as well resolved. In both the concatenated ML and species tree analyses only three species were fully resolved: *O. darwini, O. parvus,* and *O. vespertilio*. In each analysis the Galapagos island endemic *O. darwini* was fully supported as being sister to the rest of the *Ogcocephalus* species included in this study. With the exception of *O. porrectus* (not sampled in this study), all other *Ogcocephalus* species are distributed in the western Atlantic. Outside of the genus *Ogcocephalus*, most of the diversity of the family Ogcocephalidae is found in the Pacific, with nine of ten genera occurring there (Hubbs 1958; Bradbury 1999; Nelson 2006; Cruz-Acevedo, Salas-Singh & Aguirre-Villaseñor 2017; Ho *et al.* 2013). Previous studies estimated that *Ogcocephalus* diverged from other eastern Pacific and western Atlantic genera between 45 million and 20 million years ago, well before the closing of the Isthmus of Panama (Derouen *et al.* 2015; Thacker 2017). This study included one species of *Ogcocephalus* and as such did not examine divergence time between eastern Pacific and western Atlantic

Ogcocephalus species specifically. Sampling of *O. porrectus*, along with the remaining members of *Ogcocephalus* and other species in the family, could further elucidate the biogeographic history of this group.

All analyses in the current study recovered *O. parvus* as sister to a large group containing all other species besides *O. darwini*. *Ogcocephalus vespertilio* was also supported as being sister to clades B and C. The molecular data examined here disagree with previous results using morphometric data, which suggested that *O. parvus* and *O. declivirostris* were sister species (Bradbury 1980). In the concatenated ML analyses of the complete dataset, movement of one individual of *O. declivirostris* into *O. parvus* was most likely artifactual and may have been driven by an underlying problem with those loci, which warrants further investigation; it was not observed in any of the other analyses. Aside from this one sample and analysis, all other analyses strongly suggest that *O. declivirostris* is related to *O. corniger* (clade C1), though neither species was recovered as monophyletic, clearly dismissing the hypothesis that *O. declivirostris* and *O. parvus* as geminate species in the putative northern Gulf vicariant event. We observed a similar pattern of non-monophyly between species in clade C2, although, while neither *Ogcocephalus pantostictus* and *O. radiatus* form monophyletic groups, the species tree recovered *O. pantostictus* and *O. cubifrons* as sister species. Nevertheless, the nodes within both clades C1 and C2 are poorly supported and branch lengths very short, and this poses a variety of possibilities for species resolution.

One potential reason for the lack of monophyly in *O. pantostictus* and *O. radiatus* is the questionable discrimination of these species. Previous studies have concluded that *O. radiatus* may not be valid, stating that the description provided by Mitchill (1818) overlaps greatly with *O. cubifrons* and therefore they should be considered the same species (Longley 1941; Bradbury

1980; 2003). The specimen described by Mitchill was collected off the coast of Florida, where *O. cubifrons* is regularly encountered (Bradbury 1980). With the original holotype of *O. radiatus* being lost, it is impossible to effectively discern differences between *O. radiatus* and *O. cubifrons*. Molecular data collected here do not support a strong distinction between the two species, in agreement with previous studies. However, in this study *O. pantostictus* is also present in this complex and individuals from *O. radiatus and O. pantostictus* are found throughout clade B, further clouding the situation.

Similarly, there is slim support for the distinction of *O. corniger* and *O. declivirostris*. In general, identification of species in this genus is difficult due to problematic diagnostic characteristics that are either unclear or overlapping (e.g. rostrum length, rostrum orientation, overlapping ray counts, etc. see Bradbury 1967; 1980). The incongruence between species names and phylogenetic position could indicate that there are in actuality fewer evolutionary lineages than are currently recognized as species, or that some species have been misidentified. Alpha taxonomy (and diagnostics) for species in this genus may need to be revised. Specimens used in this study will undergo re-identification informed by the molecular data to try to clear up confusion in the literature and possible identification issues in this study.

Another reason for the incongruent patterns found in clades C1 and C2 may be derived from conflicting signal in the data due to gene tree incongruence. The short branches in these clades in the concatenated ML analysis suggest that very few mutations have occurred since divergence of these species. In the species tree analysis, short branches are also observed indicating that the time interval between divergence events was short on a coalescent scale or relative to population size and the rate of lineage sorting. Rapid divergence can lead to problems with phylogenetic inferences, even when many genes are used, because there is a high

probability of incomplete lineage sorting (Pamilo & Nei 1988; Maddison & Knowles 2006). Conflicting gene tree and species tree relationships are more likely when species have large effective population sizes and/or long generation times because not enough time since divergence has passed to allow for genes to sort between species (Maddison & Knowles 2006; Degnan & Rosenberg 2009b; Linkem *et al.* 2016). We note that a pattern of incongruence between gene trees and species identity is also seen in the morphologically distinct *Prionotus paralatus* (western Gulf) and *Prionotus alatus* (eastern Gulf) (Portnoy *et al.* 2017). Preliminary population genetic analysis, however, indicate difference in allele frequencies between the morphotypes to suggest they are independent evolutionary units (Portnoy unpublished data). Gene tree incongruence may also be indicative of admixture between species within Clade C1 and C2 after secondary contact (Templeton 2001; Yang & Rannala 2010). To further investigate either possibility, sampling across the range of each species would be needed to properly test for evolutionary distinctness.

CHAPTER I: CONCLUSION

Here we show that using UCEs to infer phylogenies of recently diverged fishes generates robust phylogenetic hypotheses for *Sphoeroides,* while the genus *Ogcocephalus* was more problematic. Broadly, the topologies of each respective genus did not change with respect to the type of analysis or degree of missing data, although branch support values did change modestly. Allowing missing data so that more loci can be included in the analysis increased support, although we also found that some partitions with higher missing data showed poorer support. This contrasts with recent studies that showed that missing data had little effect on phylogenetic accuracy, as long as there was no difference among loci aside from completeness. This may

suggest that, in this case, there are systematic biases among loci with greater amounts of missing data that render them less useful for phylogenetic analyses.

This study however, did recover close, if not sister, relationships between one pair of species each of *Sphoeroides* and *Ogcocephalus* distributed on either side of the hypothesized northern Gulf suture zone, while another pair of *Ogcocephalus* were rejected as a geminate species pair for this feature. Further investigation into *Ogcocephalus* is warranted to resolve incongruence between species identity and phylogenetic position which may be caused by misidentification of specimens, incorrect alpha taxonomy or the presence of incomplete lineage sorting and gene tree incongruence. Finally, this study would benefit from additional analyses investigating divergence times of both *Sphoeroides* and *Ogcocephalus* lineages to compare speciation events with geological events and additional taxon sampling for each group might further strengthen the phylogenetic hypotheses generated here.

REFERENCES

Amaral, C.R.L., Brito, P.M., Silva, D.A. & Carvalho, E.F. (2013) A New Cryptic Species of South American Freshwater Pufferfish of the Genus Colomesus (Tetraodontidae), Based on Both Morphology and DNA Data (ed D Huchon). *PLoS ONE*, 8, e74397–15.

Avise, J.C. (2000) Cytonuclear genetic signatures of hybridization phenomena: Rationale, utility, and empirical examples from fishes and other aquatic animals. *Reviews in Fish Biology and Fisheries*, 10, 253–263.

Balsam, W.L. & Beeson, J.P. (2003) Sea-floor sediment distribution in the Gulf of Mexico. *Deep Sea Research Part I: Oceanographic Research Papers*, 50, 1421–1444.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S. & Haussler, D. (2004) Ultraconserved elements in the human genome. *Science*, 304, 1321–1325.

Bert, T.M. & Harrison, R.G. (1988) Hybridization in western Atlantic stone crabs (genus Menippe): evolutionary history and ecological context influence species interactions. *Evolution*, 42, 528–544.

Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120.

Bradbury, M.G. (1967) The Genera of Batfishes (Family Ogcocephalidae). *Copeia*, 1967, 399–422.

Bradbury, M.G. (1980) A revision of the fish genus Ogcocephalus with descriptions of new species from the western Atlantic Ocean (Ogcocephalidae: Lophiformes). *Proceedings of the California Academy of Sciences*, 42, 229–285.

Bradbury, M.G. (1999) A Review of the Fish Genus Dibranchus with Descriptions of New Species and a New Genus, Solocisquama (Lophiiformes, Ogcocephalidae). *Proceedings of the California Academy of Sciences*, 51, 259–310.

Bradbury, M.G. (2003) Annotated checklist of fishes. *California Academy of Sciences Annotated Checklists of Fishes,* 17, 1–17.

Britz, R. & Kottelat, M. (1999) Carinotetraodon imitator, a new freshwater pufferfish from India (Teleostei: Tetraodontiformes). *Journal of South Asian Natural History*, 4, 39–47.

Brooks, R.A., Keitzer, S.C. & Sulak, K.J. (2005) *Taxonomic Composition and Relative Frequency of the Benthic Fish Community Found on Natural Sand Banks and Shoals in the Northwestern Gulf of Mexico*. USGS.

Bulgarella, M., Trewick, S. A., Minards, N. A., Jacobson, M. J., & Morgan-Richards, M. (2014). Shifting ranges of two tree weta species (H emideina spp.): competitive exclusion and changing climate. *Journal of Biogeography*, *41*(3), 524-535.

Capella-Gutierrez, S., Silla-Martinez, J.M. & Gabaldon, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25, 1972–1973.

Carpenter, K.E. & De Angelis, N. (2002) *The Living Marine Resources of the Western Central Atlantic*. Rome: Food and Agriculture Organization of the United Nations.

Coyne, J.A. & Orr, H.A. (1989) Patterns of Speciation in Drosophila. *Evolution*, 43, 362–381.

Coyne, J.A. & Orr, H.A. (1998) The evolutionary genetics of speciation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 353, 287–305.

Coyne, J.A. & Orr, H.A. (2004) *Speciation*. Sinauer Associates, Sunderland, MA.

Cruz-Acevedo, E., Salas-Singh, C. & Aguirre-Villaseñor, H. (2017) Distribution of Dibranchus species (Pisces: Ogcocephalidae) from the Eastern Central Pacific and their relationship with environmental factors. *Marine Biodiversity,* 1–11.

Degnan, J.H. & Rosenberg, N.A. (2009a) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24, 332–340.

Degnan, J.H. & Rosenberg, N.A. (2009b) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24, 332–340.

Derouen, V., Ludt, W.B., Ho, H.-C. & Chakrabarty, P. (2015) Examining evolutionary relationships and shifts in depth preferences in batfishes (Lophiiformes: Ogcocephalidae). *Molecular Phylogenetics and Evolution*, 84, 27–33.

DiBattista, J.D., Whitney, J., Craig, M.T., Hobbs, J.-P.A., Rocha, L.A., Feldheim, K.A., Berumen, M.L. & Bowen, B.W. (2016) Surgeons and suture zones: Hybridization among four surgeonfish species in the Indo-Pacific with variable evolutionary outcomes. *Molecular Phylogenetics and Evolution*, 101, 203–215.

Eytan, R.I., Evans, B.R., Dornburg, A., Lemmon, A.R., Lemmon, E.M., Wainwright, P.C. & Near, T.J. (2015) Are 100 enough? Inferring acanthomorph teleost phylogeny using Anchored Hybrid Enrichment. *BMC Evolutionary Biology*, 1–20.

Faircloth, B.C. (2015) *PHYLUCE Is a Software Package for the Analysis of Conserved Genomic Loci*. Cold Spring Harbor Labs Journals.

Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T. & Glenn, T.C. (2012) Ultraconserved Elements Anchor Thousands of Genetic Markers Spanning Multiple Evolutionary Timescales. *Systematic Biology*, 61, 717–726.

Faircloth, B.C., Sorenson, L., Santini, F. & Alfaro, M.E. (2013) A Phylogenomic Perspective on the Radiation of Ray-Finned Fishes Based upon Targeted Sequencing of Ultraconserved Elements (UCEs) (ed CS Moreau). *PLoS ONE*, 8, e65923–7.

Floeter, S.R., Rocha, L.A., Robertson, D.R., Joyeux, J.C., Smith-Vaniz, W.F., Wirtz, P., Edwards, A.J., Barreiros, J.P., Ferreira, C.E.L., Gasparini, J.L., Brito, A., Falcón, J.M., Bowen, B.W. & Bernardi, G. (2008) Atlantic reef fish biogeography and evolution. *Journal of Biogeography*, 35, 22–47.

Garman, S. (1899) *The Fishes*.

Gilbert, P.S., Chang, J., Pan, C., Sobel, E.M., Sinsheimer, J.S., Faircloth, B.C. & Alfaro, M.E. (2015) Genome-wide ultraconserved elements exhibit higher phylogenetic informativeness than traditional gene markers in percomorph fishes. *Molecular Phylogenetics and Evolution*, 92, 140–146.

Ginsburg, I. (1952) Eight new fishes from the Gulf coast of the United States, with two new genera and notes on geographic distribution. *Washington Academy of Sciences*, 42, 81–101.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. & Regev, A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, 29, 644–652.

Grummer, J.A., Calderón-Espinosa, M.L., Nieto-Montes de Oca, A., Smith, E.N., Méndez-de la Cruz, F.R. & Leaché, A.D. (2015) Estimating the temporal and spatial extent of gene flow among sympatric lizard populations (genus Sceloporus) in the southern Mexican highlands. *Systematic Biology*, 24, 1523–1542.

Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907–913.

Ho, H.C., Chakrabarty, P. & Sparks, J.S. (2010) Review of the Halieutichthys aculeatus species complex (Lophiiformes: Ogcocephalidae), with descriptions of two new species. *Journal of Fish Biology*, 77, 841–869.

Ho, H.C., Endo, H., Studies, K.S.Z.2008. A new species of Halicmetus (Lophiiformes: Ogcocephalidae) from the Western Pacific, with comments on congeners. *Zoological Studies*.

Ho, H.C., Prokofiev, A.M., Studies, K.S.Z.2009. A new species of the batfish genus Malthopsis (Lophiiformes: Ogcocephalidae) from the northwestern Indian Ocean. *zoolstud.sinica.edu.tw*.

Ho, H.-C., Roberts, C.D. & Shao, K.-T. (2013) Revision of batfishes (Lophiiformes: Ogcocephalidae) of New Zealand and adjacent waters, with description of two new species of the genus Malthopsis. *Zootaxa*, 3626, 188–13.

Hobbs, J. & Allen, G.R. (2014) Hybridization among coral reef fishes at Christmas Island and the Cocos (Keeling) Islands. *Raffles Bulletin of Zoology*.

Hobbs, J.P.A. & Salmond, J.K. (2008) Cohabitation of Indian and Pacific Ocean species at Christmas and Cocos (Keeling) Islands. *Coral Reefs*, 27, 933–933.

Hobbs, J., Frisch, A.J. & Allen, G.R. (2008) Marine hybrid hotspot at Indo-Pacific biogeographic border. *Biology Letters*, 258–261.

Hollenbeck, C.M. (2016) Genomic Studies of Red Drum (Sciaenops ocellatus) in US Waters (Doctoral dissertation)., 1–161.

Hrbek, T. & Meyer, A. (2003) Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). *Journal of Evolutionary Biology*, 16, 17–36.

Hubbs, C.L. (1958) Ogcocephalus darwini, a new batfish endemic at the Galapagos Islands. *Copeia*, 1958, 161.

Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780.

Lanier, H.C., Huang, H. & Knowles, L.L. (2014) How low can you go? The effects of mutation rate on the accuracy of species-tree estimation. *Molecular Phylogenetics and Evolution*, 70, 112–119.

Lemmon, A.R., Emme, S.A. & Lemmon, E.M. (2012) Anchored Hybrid Enrichment for Massively High-Throughput Phylogenomics. *Systematic Biology*, 61, 727–744.

Li, W. & Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22, 1658–1659.

Linkem, C.W., Minin, V.N. & Leaché, A.D. (2016) Detecting the Anomaly Zone in Species Trees and Evidence for a Misleading Signal in Higher-Level Skink Phylogeny (Squamata: Scincidae). *Systematic Biology*, 65, 465–477.

Longley, W.H. (1941) Systematic Catalogue of the Fishes of Tortugas, Florida. *Carnegie Institution of Washington Publication*, 535, 1-311.

Maddison, W.P. & Knowles, L.L. (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology*, 55, 21–30.

Marko, P.B. (2002) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the isthmus of panama. *Molecular Biology and Evolution*, 19, 2005–2021.

Matsuura, K. (2015) Taxonomy and systematics of Tetraodontiform fishes: a review focusing primarily on progress in the period from 1980 to 2014. *Ichthyological Research*, 62, 72–113.

Mayr, E. (1970). *Populations, species, and evolution: an abridgment of animal species and evolution*. Harvard University Press.

McClure, M.R. & McEachran, J.D. (1992) Hybridization between Prionotus alatus and P. paralatus in the Northern Gulf of Mexico (Pisces: Triglidae). *Copeia*, 1992, 1039.

Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S. & Warnow, T. (2014) ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30, 541–548.

Mitchill, S. L. (1818). Memoir on ichthyology. *The fishes of New York described and arranged. American Monthly Magazine Critical Review*, *2*, 241-248.

Moore, A.J., Merges, D. & Kadereit, J.W. (2013) The origin of the serpentine endemic Minuartia laricifoliasubsp. ophioliticaby vicariance and competitive exclusion. *Molecular Ecology*, 22, 2218–2231.

Nelson, J.S. (2006) *Fishes of the World 4th Edition*. Hoboken: John Wiley & Sons, Inc.

Nelson, J.S., Wilson, M.V.H. & grande, T.C. (2016) *Fishes of the World*, Fifth. Wiley.

Nute, M., Chou, J., Molloy, E. K., & Warnow, T. (2018). The performance of coalescent-based species tree estimation methods under missing data. *BMC Genomics*, *19*(5), 41-57.

Palumbi, S.R. (1994) Genetic Divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, 25, 547–572.

Pamilo, P. & Nei, M. (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution*, 5, 568–583.

Parenti, L.R. & Ebach, M.C. (2009) *Comparative Biogeography*. University of Califoria Press, Berkeley, CA.

Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R.P., Moret, B.M.E. & Stamatakis, A. (2010) How many bootstrap replicates are necessary? *Journal of computational biology*, 17(3), 337–354.

Petit, R.J. & Excoffier, L. (2009) Gene flow and species delimitation. *Trends in Ecology & Evolution*, 24, 386–393.

Pinet, P.R. & Popenoe, P. (1985) A scenario of Mesozoic-Cenozoic ocean circulation over the Blake Plateau and its environs. *Geological Society of American Bulletin*, 96, 618–626.

Portnoy, D.S. & Gold, J.R. (2012) Evidence of multiple vicariance in a marine suture-zone in the Gulf of Mexico. *Journal of Biogeography*, 39, 1499–1507.

Portnoy, D. S., Willis, S. C., Hunt, E., Swift, D. G., Gold, J. R., & Conway, K. W. (2016) Molecular phylogenetics of New World searobins (Triglidae; Prionotinae). *Molecular Phylogenetics and Evolution*, *107*, 382-387.

Quenouille, B., Hubert, N., Bermingham, E. & Planes, S. (2011) Speciation in tropical seas: Allopatry followed by range change. *Molecular Phylogenetics and Evolution*, 58, 546–552.

Randall, J.E. (1998) Zoogeography of Shore Fishes of the Indo-Pacific Region. *Zoological Studies*, 37(4), 227-268.

Remington, C.L. (1968) Suture-Zones of Hybrid Interaction Between Recently Joined Biotas. *Evolutionary Biology*, pp. 321–428. Springer US, Boston, MA.

Rice, P., Longden, I. & Bleasby, A. (2000) EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics*, 16, 276–277.

Roberts, T.R. (1998) Freshwater fugu or pufferfishes of the genus Tetraodon from the Mekong basin, with descriptions of two new species. *Ichthyological Research*, 45, 225–234.

Ronquist, F. (1997) Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Systematic Biology*, 46, 195–203.

Rosen, D.E. (1975) A Vicariance Model of Caribbean Biogeography. *Systematic Biology*, 24, 431–464.

Rosen, D.E. (1978) Vicariant patterns and historical explanation in biogeography. *Systematic Biology*, 27, 159–188.

Rundle, H.D. & Schluter, D. (1998) Reinforcement of Stickleback Mate Preferences: Sympatry Breeds Contempt. *Evolution*, 52, 200–208.

Saenjundaeng, P., Vidthayanon, C. & Grudpun, C. (2013) Tetraodon palustris, a new freshwater pufferfish (Tetraodontiformes: Tetraodontidae) from the Mekong Basin of Thailand. *Zootaxa*, 3686, 77–84.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular Cloning*. Cold spring harbor laboratory press, New York.

Santini, F., Sorenson, L. & Alfaro, M.E. (2013) A new phylogeny of Tetraodontiform fishes (Tetraodontiformes, Acanthomorpha) based on 22 loci. *Molecular Phylogenetics and Evolution*, 69, 177–187.

Sayyari, E. & Mirarab, S. (2016) Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. *Molecular Biology and Evolution*, 33, 1654–1668.

Seutin, G., White, B.N. & Boag, P.T. (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, 69, 82–90.

Shipp, R.L. (1970) *The Pufferfishes (Osteichthyes: Tetraodontidae) of the Atlantic Ocean and Adjacent Waters.*

Simpson, C.T. (1900) On the Evidence of the Unionidae Regarding the Forber Courses of the Tennessee and Other Southern Rivers. *Science*, 12, 133–136.

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

Stamatakis, A. (2016) The RAxML v8. 2. X Manual. *Heidelberg Institute for Theoretical Studies. Available from: http://sco. h-its. org/exelixis/web/software/raxml/# documentation*

Steininger, F.F. & Rögl, F. (1984) Paleogeography and palinspastic reconstruction of the Neogene of the Mediterranean and Paratethys. *Geological Society, London, Special Publications*, 17, 659–668.

Taylor, E.B., Boughman, J.W., Groenenboom, M., Sniatynski, M., Schluter, D. & Gow, J.L. (2005) Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (Gasterosteus aculeatus) species pair. *Molecular Ecology*, 15, 343–355.

Templeton, A.R. (2001) Using phylogeographic analyses of gene trees to test species status and processes. *Molecular Ecology*, 10, 779–791.

Thacker, C.E. (2017) Patterns of divergence in fish species separated by the Isthmus of Panama. *Evolutionary Biology*, 17, 1–14.

Walker, H.J. & Bussing, W.A. (1996) Two New Pufferfishes of the Genus Sphoeroides from the Eastern Pacific. *Copeia*, 1996, 677–684.

Wellington, G.M. & Robertson, D.R. (2001) Variation in larval life-history traits among reef fishes across the Isthmus of Panama. *Marine Biology*, 138, 11–22.

Whinnett, A., Zimmermann, M., Willmott, K.R., Herrera, N., Mallarino, R., Simpson, F., Joron, M., Lamas, G. & Mallet, J. (2005) Strikingly variable divergence times inferred across an

Amazonian butterfly 'suture zone'. *Proceedings of the Royal Society B: Biological Sciences*, 272, 2525–2533.

Yamanoue, Y., Miya, M., Matsuura, K., Katoh, M., Sakai, H. & Nishida, M. (2008) A new perspective on phylogeny and evolution of tetraodontiform fishes (Pisces: Acanthopterygii) based on whole mitochondrial genome sequences: Basal ecological diversification? *BMC Evolutionary Biology*, 8, 212–15.

Yang, Z. & Rannala, B. (2010) Bayesian species delimitation using multilocus sequence data (eds Z Yang and B Rannala). *PNAS*, 107, 9264–9269.

Zhang, C., Sayyari, E. & Mirarab, S. (2017) ASTRAL-III: Increased Scalability and Impacts of Contracting Low Support Branches. In *RECOMB International Workshop on Comparative Genomics* 53–75.

Table 1: List of tissues, their sample IDs, loaning institution and voucher numbers. (BRTC: Biodiversity Research and Teaching Collections – Texas A&M University, JFRA: Japan Fisheries Research and Education Agency, KU: The University of Kansas, MGL: Marine Genomics Laboratory, SCRIPPS: Scripps institution of Oceanography, UVTP: Universidad del Valle

Species	Sample ID	Institute	Voucher
Balistes vetula	Bvet1	MGL	NA
Balistes vetula	Bvet2	MGL	NA
Lophius gastrophysus	Lgas1	KU	KU 4961
Lophius gastrophysus	Lgas2	KU	KU 8203
Lophius gastrophysus	Lgas3	KU	KU 8204
Ogcocephalus corniger	Ocor1	BRTC	15896.01 H
Ogcocephalus corniger	Ocor2	BRTC	15896.01 F
Ogcocephalus corniger	Ocor3	BRTC	15896.01 G
Ogcocephalus cubifrons	Ocub2	BRTC	16979.01 B
Ogcocephalus cubifrons	Ocub1	BRTC	16979.01 C
Ogcocephalus darwini	Odar1	SCRIPPS	SCRIPPS 00-154-1
Ogcocephalus darwini	Odar2	SCRIPPS	SCRIPPS 00-154-2
Ogcocephalus darwini	Odar3	SCRIPPS	SCRIPPS 00-154-3
Ogcocephalus declivirostris	Odec1	BRTC	15914.01
Ogcocephalus declivirostris	Odec3	BRTC	15915.02
Ogcocephalus declivirostris	Odec4	BRTC	15916.01
Ogcocephalus declivirostris	Opar3	BRTC	17342.01
Ogcocephalus declivirostris	Odec2	BRTC	17512.01 A
Ogcocephalus pantostictus	Opan1	BRTC	17481.01
Ogcocephalus pantostictus	Opan2	BRTC	17530.02
Ogcocephalus parvus	Opar1	BRTC	16078.05
Ogcocephalus parvus	Opar2	BRTC	17531.01
Ogcocephalus radiatus	Orad2	BRTC	17485.02 A
Ogcocephalus radiatus	Orad3	BRTC	17485.02 B
Ogcocephalus radiatus	Orad1	BRTC	16114.01
Ogcocephalus vespertillio	Oves1	BRTC	17589.01
Sphoeroides annulatus	Sann3	UVTP	UVPT 551
Sphoeroides annulatus	Sann1	SCRIPPS	SCRIPPS 004-85
Sphoeroides annulatus	Sann2	SCRIPPS	SCRIPPS 007-91
Sphoeroides dorsalis	Sdor1	BRTC	17377.01
Sphoeroides dorsalis	Sdor2	BRTC	17378.02 A
Sphoeroides dorsalis	Sdor3	BRTC	17378.02 C
Sphoeroides greeleyi	Sgre2	BRTC	17587.01

Sphoeroides greeleyi	Sgre3	BRTC	17588.01
Sphoeroides greeleyi	Sgre1	BRTC	17584.01
Sphoeroides lispus	Slis1	SCRIPPS	SCRIPPS 09-355
Sphoeroides lobatus	Slob2	SCRIPPS	SCRIPPS 007-17
Sphoeroides lobatus	Slob1	SCRIPPS	SCRIPPS 007-57
Sphoeroides lobatus	Slob3	SCRIPPS	SCRIPPS 007-98
Sphoeroides maculatus	Smac1	KU	KU 1136
Sphoeroides maculatus	Smac2	KU	KU 1225
Sphoeroides maculatus	Smac3	KU	KU 1226
Sphoeroides maromotus	Smar1	USMN	USMN 093
Sphoeroides maromotus	Smar2	USMN	USMN 094
Sphoeroides nephelus	Snep1	KU	KU 5432
Sphoeroides pachygaster	SpachB1	BRTC	19329.01
Sphoeroides pachygaster	SpachB2	BRTC	19330.01
Sphoeroides pachygaster	SpachB3	Belize	NA
Sphoeroides pachygaster	SpacJ1	JFRA	NA
Sphoeroides pachygaster	SpacJ2	JFRA	NA
Sphoeroides pachygaster	SpacJ3	JFRA	NA
Sphoeroides parvus	Spar1	BRTC	17325.01 A
Sphoeroides parvus	Spar2	BRTC	17307.02 A
Sphoeroides parvus	Spar4	BRTC	17458.03 A
Sphoeroides parvus	Spar5	BRTC	17458.03 B
Sphoeroides parvus	Spar3	BRTC	17458.03 C
Sphoeroides spengleri	Sspe2	BRTC	17340.01 B
Sphoeroides spengleri	Sspe3	BRTC	17340.01 C
Sphoeroides testudineus	Stes1	BRTC	17573.01
Sphoeroides testudineus	Stes2	BRTC	17574.01
Sphoeroides testudineus	Stes3	BRTC	17575.01
Sphoeroides trichocephalus	Stri1	UVTP	UVTP 690
Sphoeroides trichocephalus	Stri2	UVTP	UVTP 691
Sphoeroides trichocephalus	Stri3	UVTP	UVTP 692
Sphoeroides tyleri	Styl1	BRTC	17580.01

Table 2: Composition of each dataset for a) Sphoeroides and b) Ogcocephalus

a)

Sphoeroides				
Dataset	Loci	Mean Length	Variable Sites	PI Sites
Complete	20	1114.15	2035	1675
10% Missing Data	429	1015.31	44783	37792
25% Missing Data	432	1010.91	44895	37889
50% missing Data	435	1006.00	44989	37933
Incomplete	436	974.21	44996	37939

b)

Ogcocephalus				
Dataset	Loci	Mean Length	Variable Sites	PI Sites
Complete	205	974.58	27711	9243
10% Missing Data	452	981.59	42323	18868
25% Missing Data	454	977.72	42381	18910
50% missing Data	458	971.97	42427	18933
Incomplete	459	970.34	42458	18937

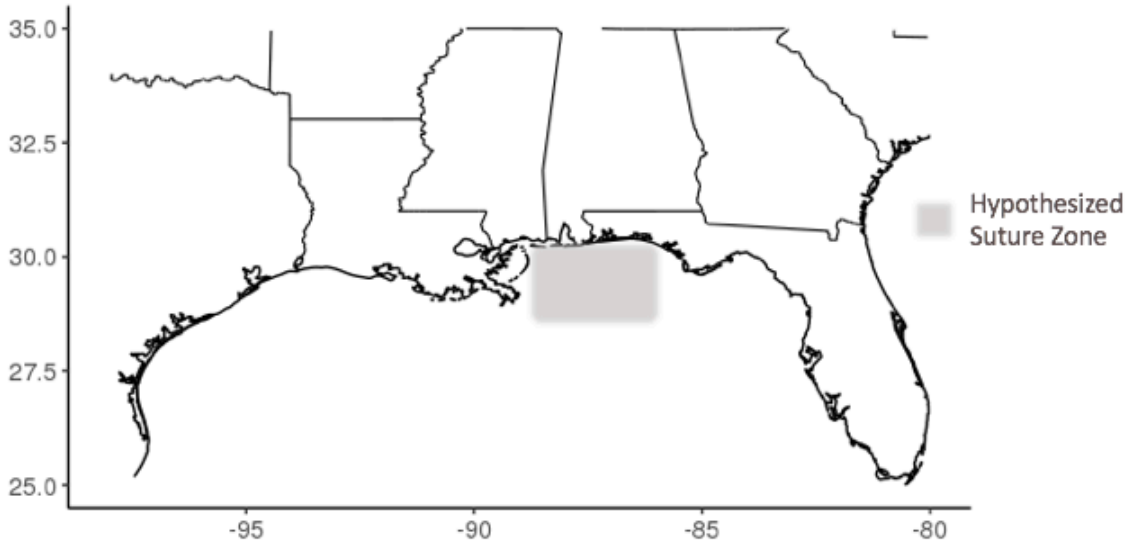


Figure 1*:* Map of the northern Gulf of Mexico with the hypothesized suture zone shown in the grey box.

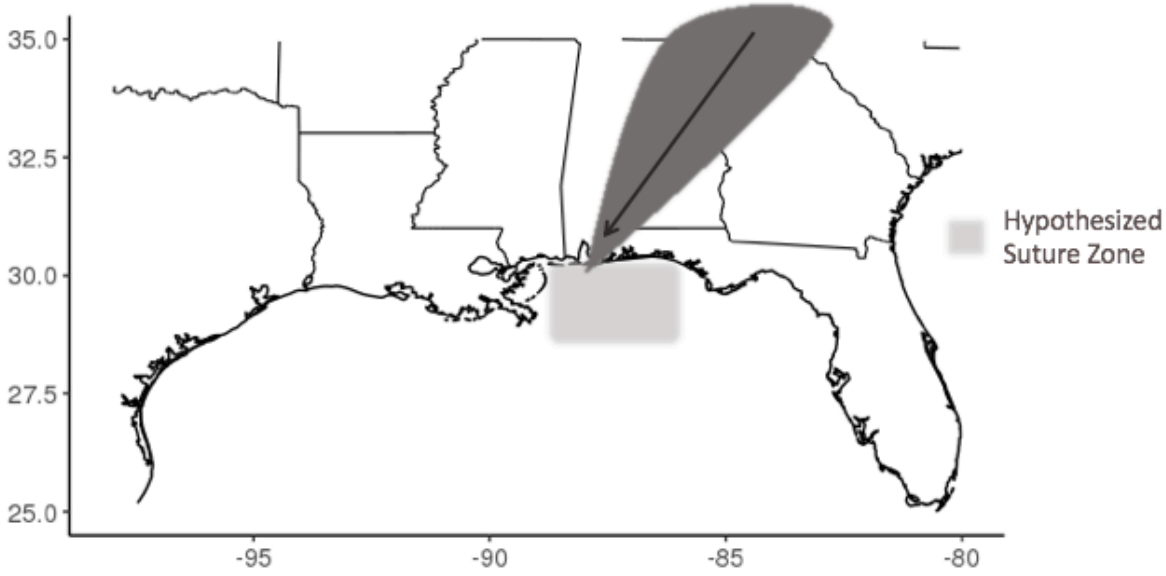


Figure 2*:* Map of historic drainage of the Tennessee River basin directly into the Gulf, ~2.4 million years ago as proposed by Ginsburg (1952) shown in dark grey in relation to the current hypothesized norther Gulf of Mexico suture zone (light grey).

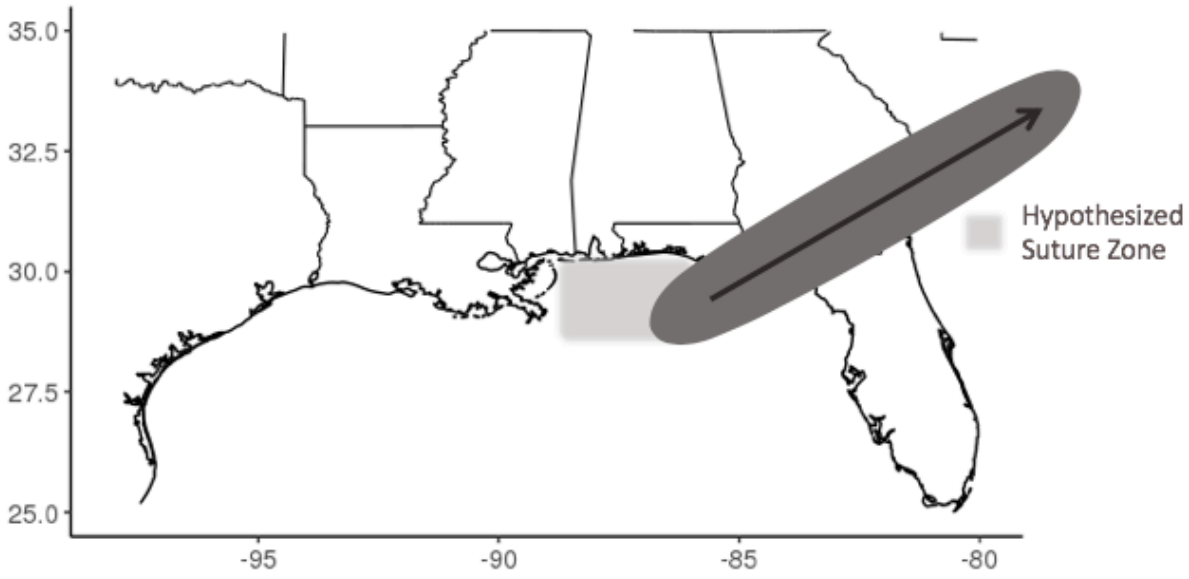


Figure 3: Map of the Suwanee Straights (dark grey) which separated the continental United States and Florida (closed ~1.75 million years ago) in relation to the current hypothesized norther Gulf of Mexico suture zone (light grey).

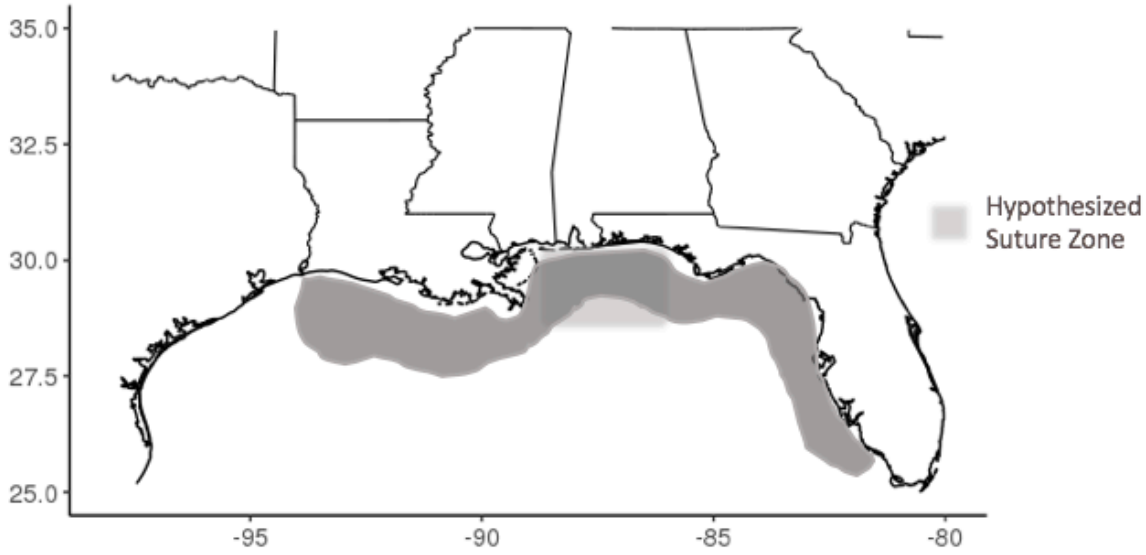


Figure 4: Map of past glacial refugia (dark grey) present during the early Pleistocene glaciations (~700-135 thousand years ago) n relation to the current hypothesized norther Gulf of Mexico suture zone (light grey).

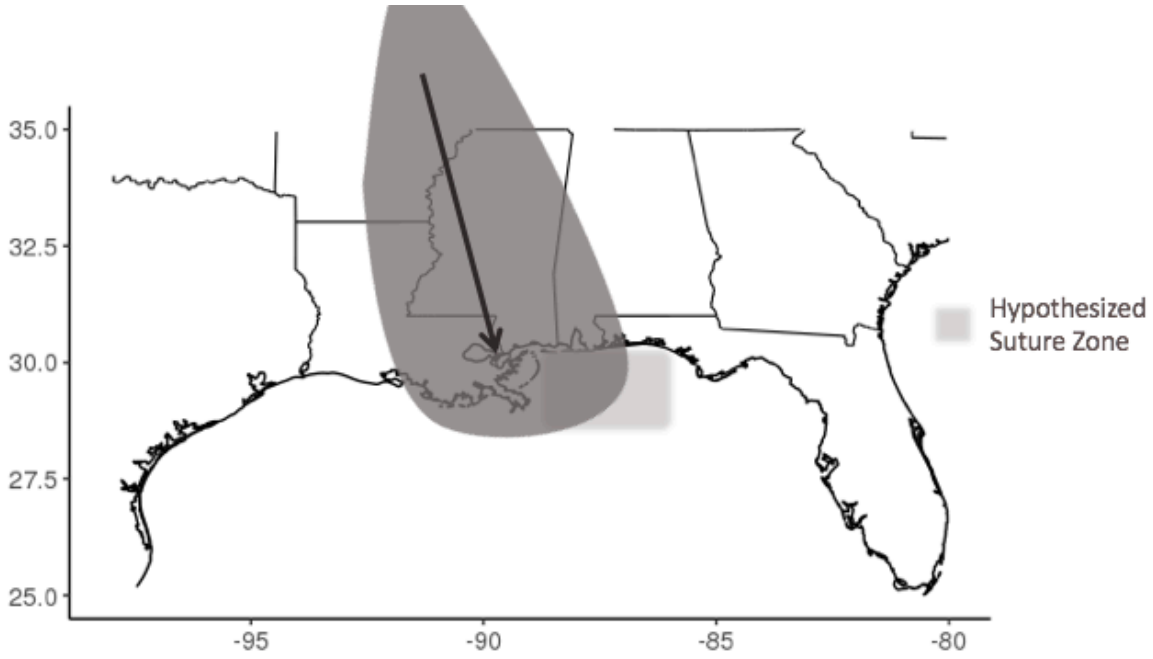


Figure 5*:* Map of historic outflow of the Mississippi (~16-9 kya; dark grey) present in the northern Gulf of Mexico in relation to the current hypothesized norther Gulf of Mexico suture zone (light grey).

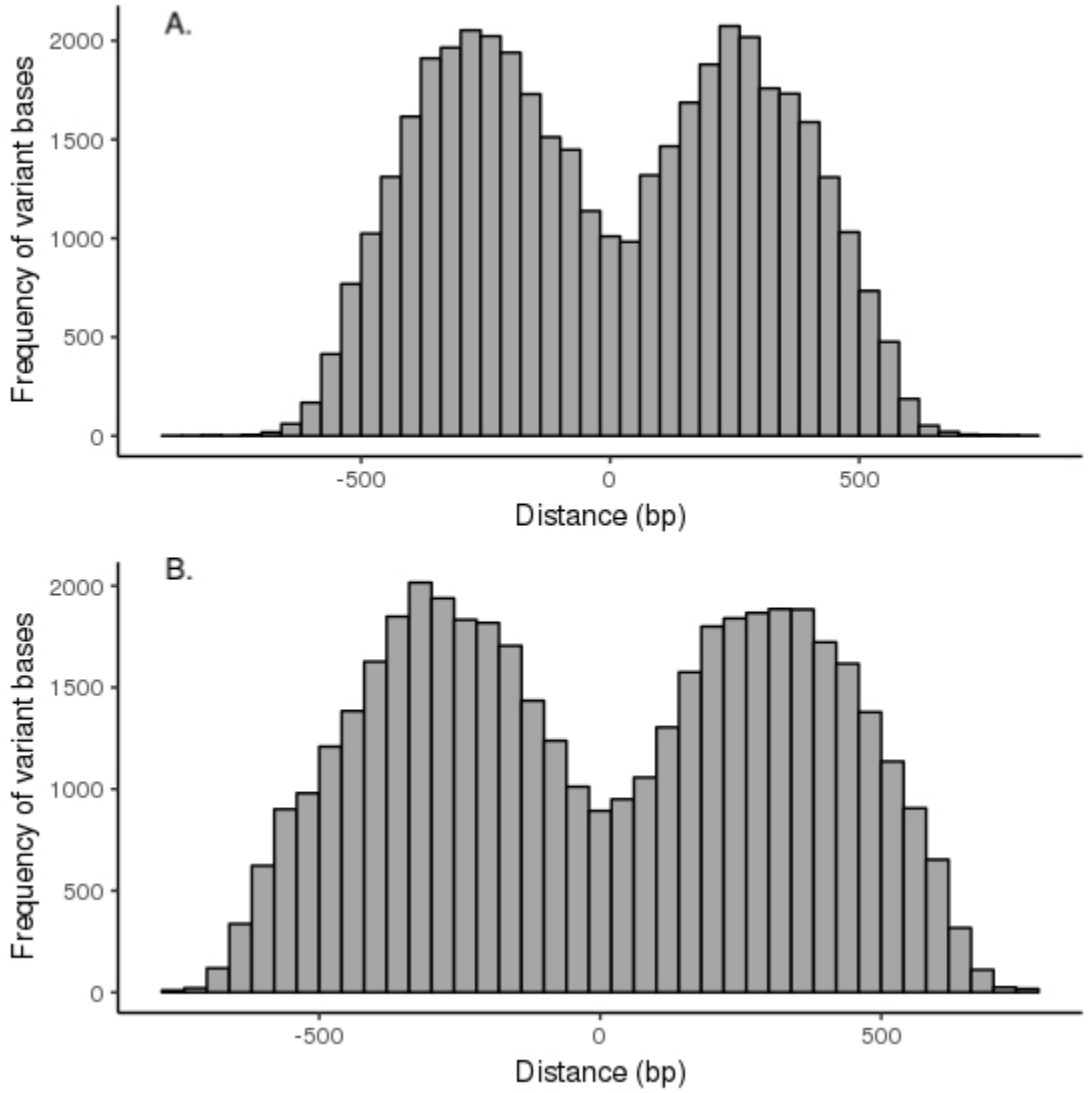
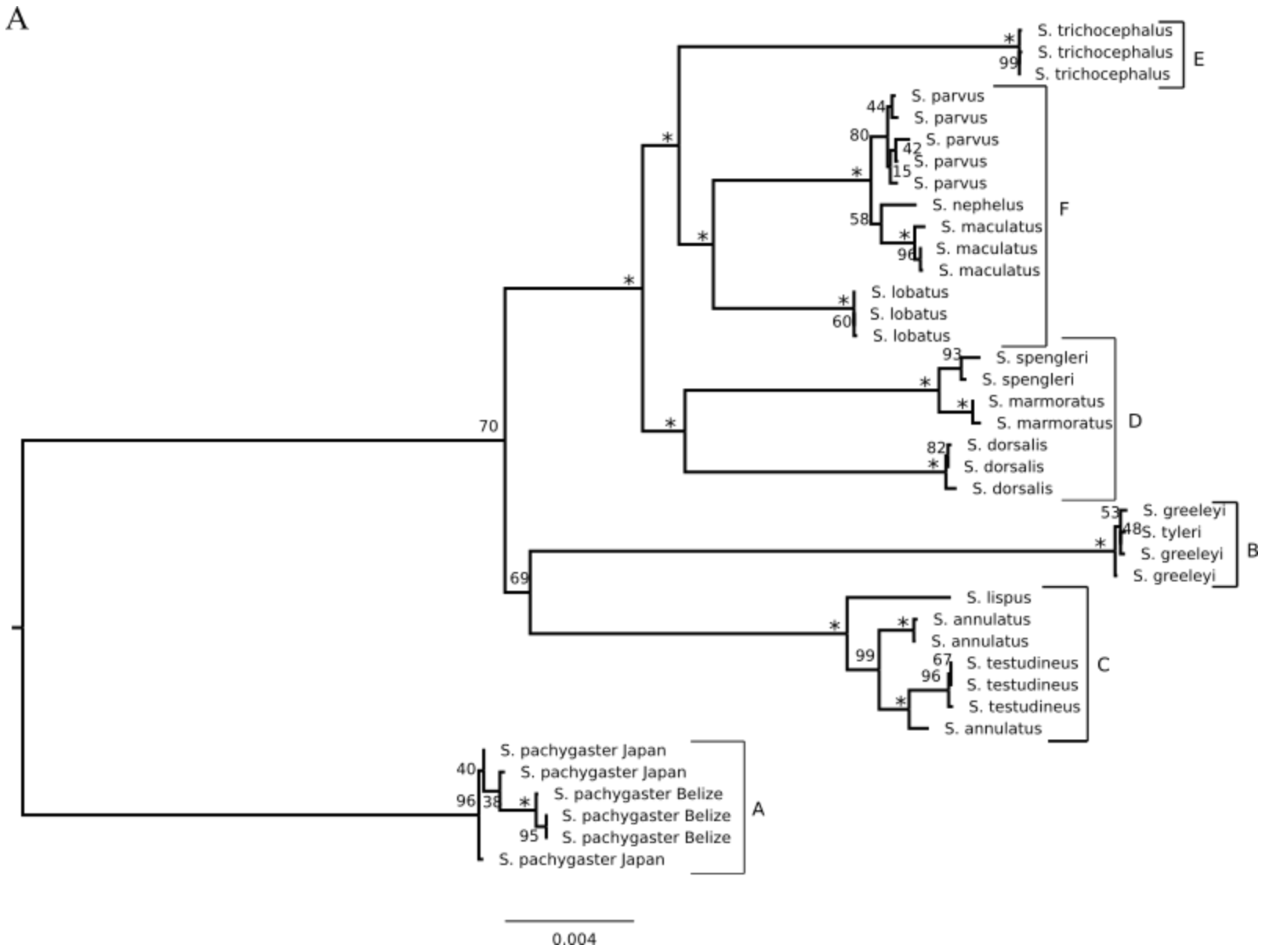
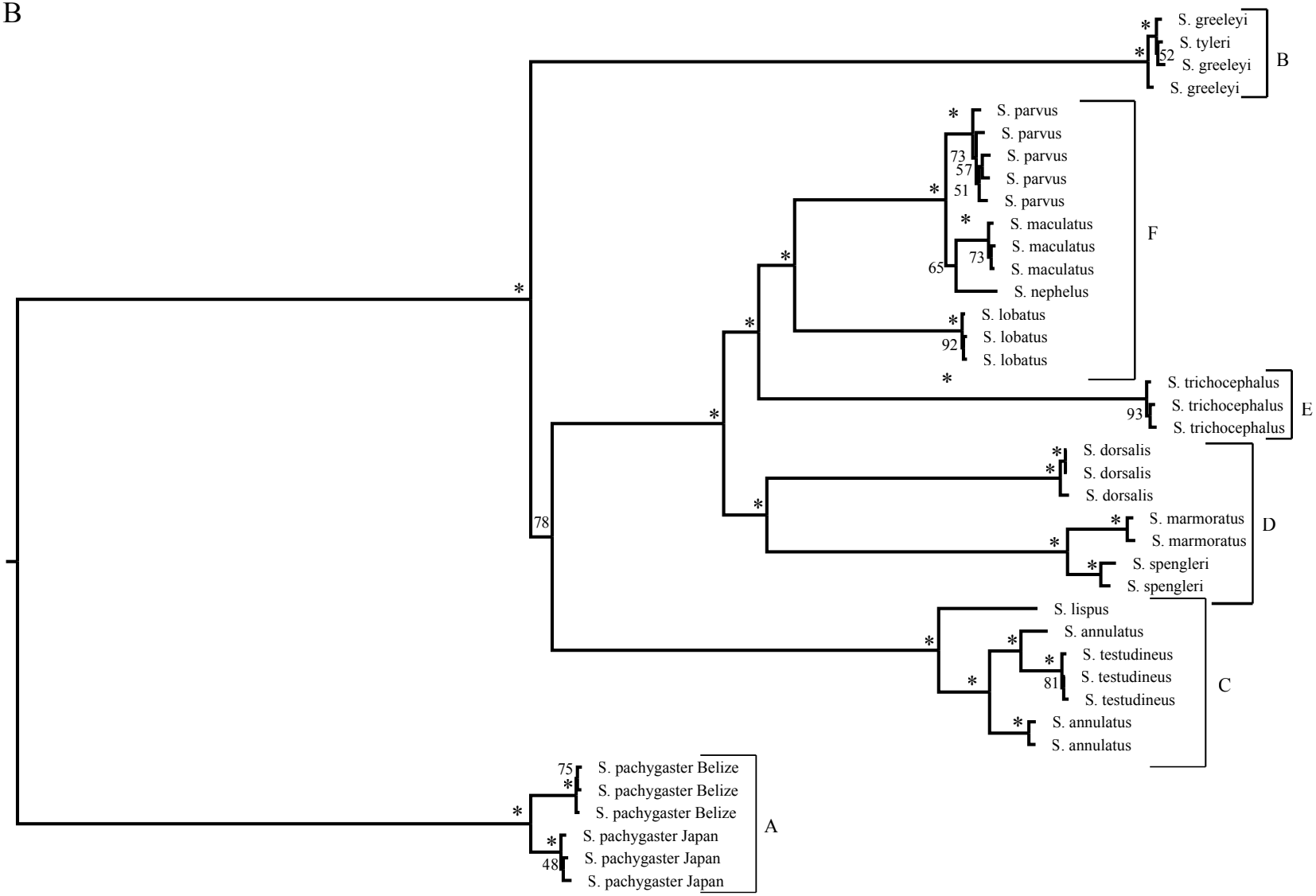


Figure 6: Frequency distribution of variable positions across UCE loci for all loci sequenced in a) Sphoeroides and b) Ogcocephalus not including variable positions for heterozygotes. Number of variable sites are summed in 10 base pair bins. Distance is measured in base pairs from the midpoint of the alignment outward.

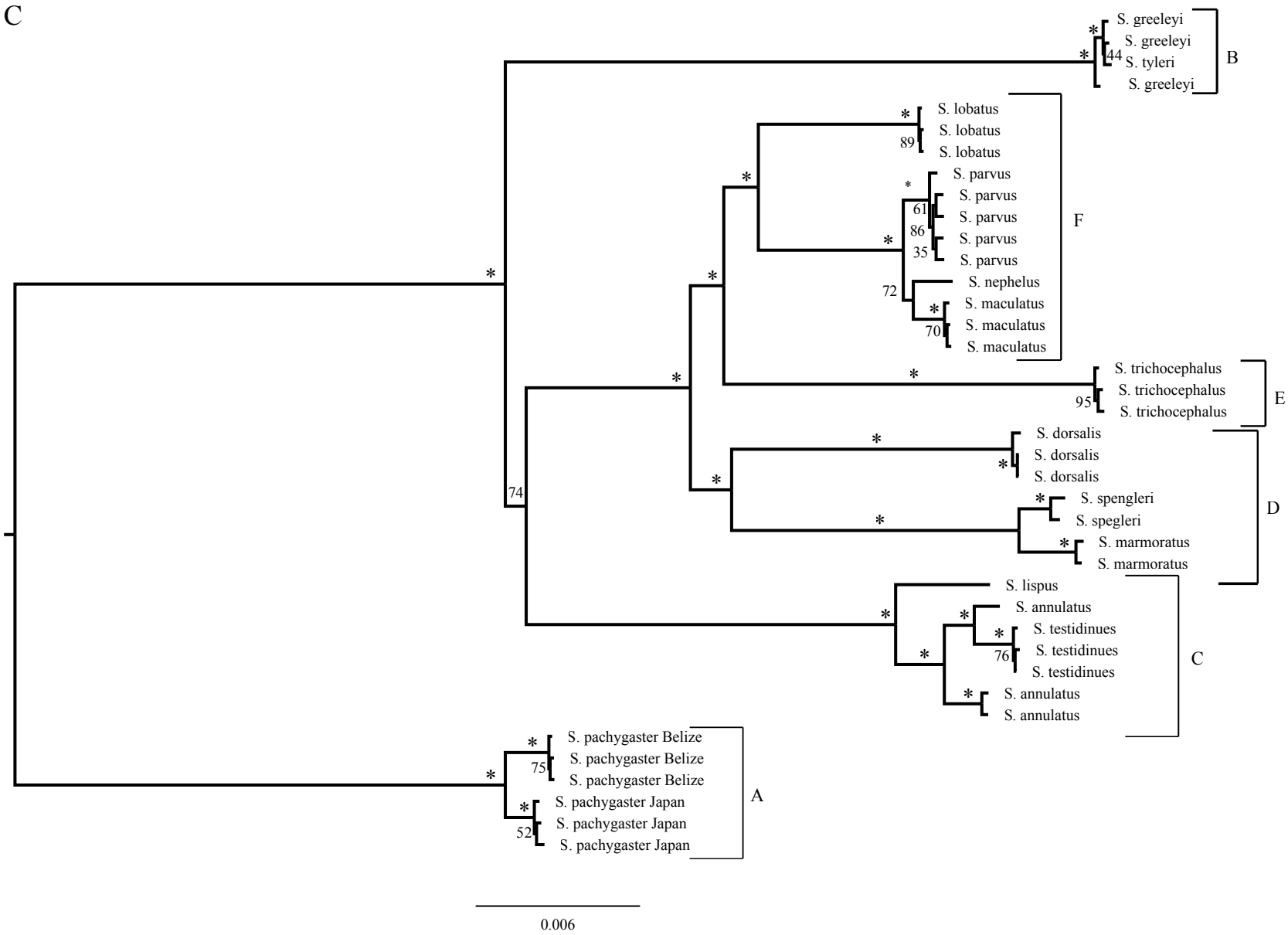
Figure 7*: Sphoeroides* phylogenetic hypotheses generated using ML with the full concatenated datasets and the species trees generated using Bayesian analysis. Full support ML bootstrap % = 100 and LPP = 1.0 denoted by the *. A) Concatenated ML analysis of the complete dataset B) Concatenated ML analysis of the dataset with 10% missing data C) Concatenated ML analysis of the dataset with 25% missing data D) Concatenated ML analysis of the dataset with 50% missing data E) Concatenated ML analysis of the incomplete dataset. ML trees have had the outgroup pruned out for ease of viewing. F) Species tree analysis of the complete dataset G) Species tree analysis of the dataset with 10% missing data H) Species tree analysis of the dataset with 25% missing data I) Species tree analysis of the dataset with 50% missing data J) Species tree analysis of the incomplete dataset. Branch length for ML analysis is scaled at 0.02 mutations per site and branch lengths in the coalescent analysis are measure in coalescent units.

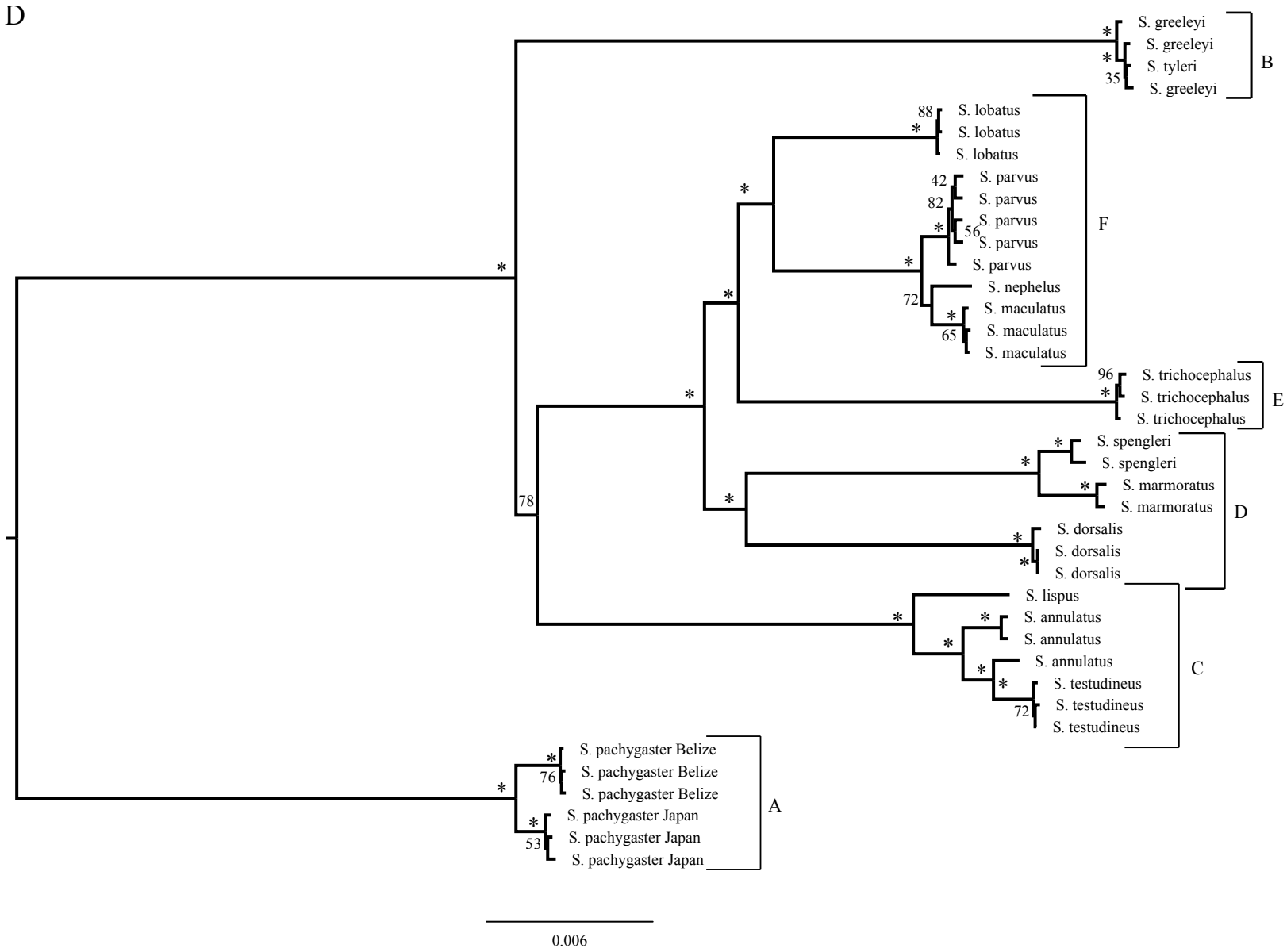

A
S. trichocephalus
S. trichocephalus
S. trichocephalus
E
S. parvus
S. parvus
S. parvus
S. parvus
S. parvus
S. nephelus
S. maculatus
S. maculatus
S. maculatus
F
S. lobatus
S. lobatus
S. lobatus
S. spengleri
S. spengleri
S. marmoratus
S. marmoratus
D
S. dorsalis
S. dorsalis
S. dorsalis
S. greeleyi
S. tyleri
S. greeleyi
S. greeleyi
B
S. lispus
S. annulatus
S. annulatus
S. testudineus
S. testudineus
S. testudineus
S. annulatus
C
S. pachygaster Japan
S. pachygaster Japan
S. pachygaster Belize
S. pachygaster Belize
S. pachygaster Belize
S. pachygaster Japan
A
0.004

B

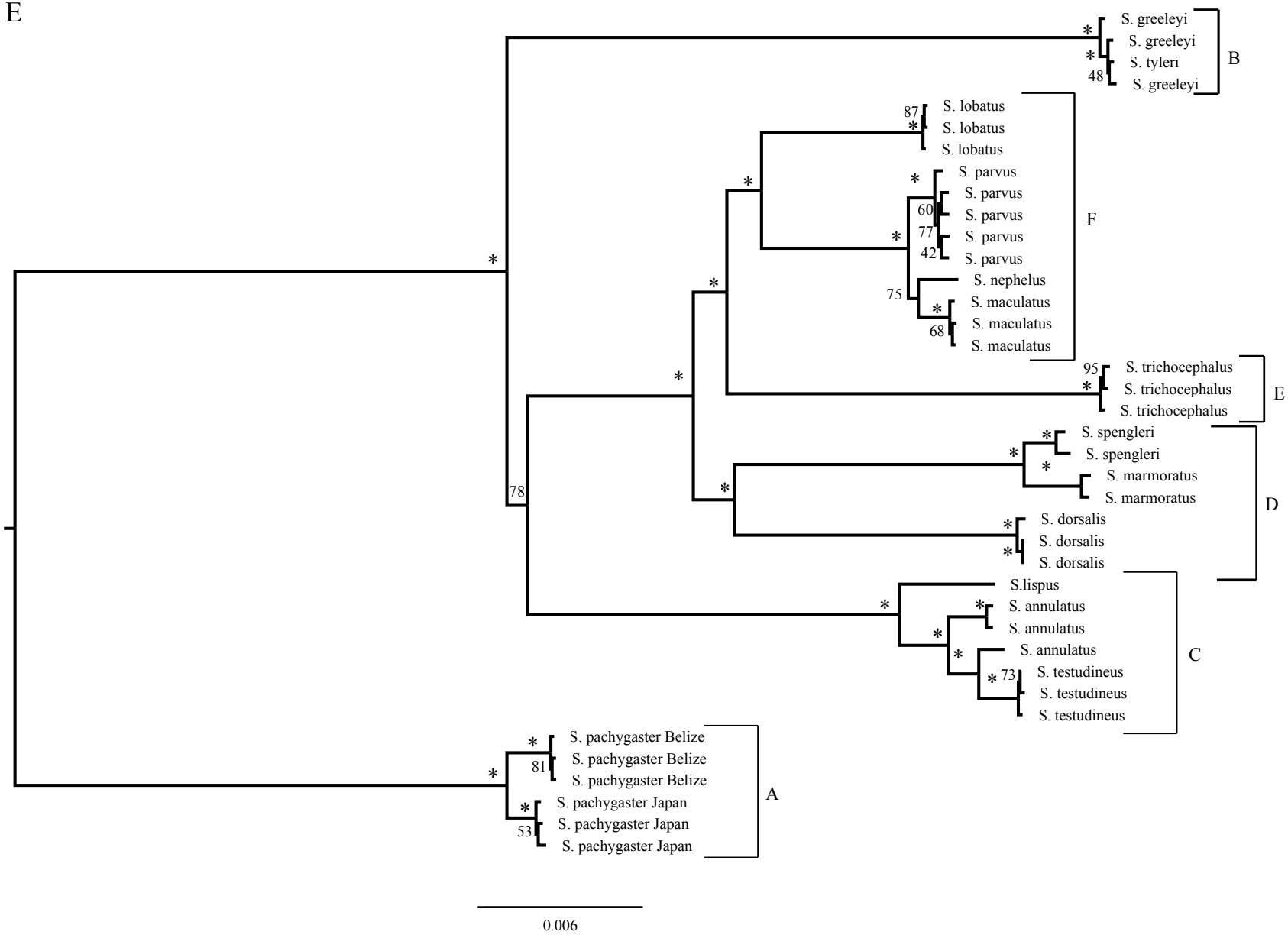

S. greeleyi
S. tyleri
S. greeleyi
S. greeleyi
52
B
S. parvus
S. parvus
S. parvus
S. parvus
S. parvus
73
57
51
S. maculatus
S. maculatus
S. maculatus
S. nephelus
73
65
S. lobatus
S. lobatus
S. lobatus
92
F
S. trichocephalus
S. trichocephalus
S. trichocephalus
93
E
S. dorsalis
S. dorsalis
S. dorsalis
S. marmoratus
S. marmoratus
S. spengleri
S. spengleri
78
D
S. lispus
S. annulatus
S. testudineus
S. testudineus
S. testudineus
S. annulatus
S. annulatus
81
C
S. pachygaster Belize
S. pachygaster Belize
S. pachygaster Belize
S. pachygaster Japan
S. pachygaster Japan
S. pachygaster Japan
75
48
A
0.006

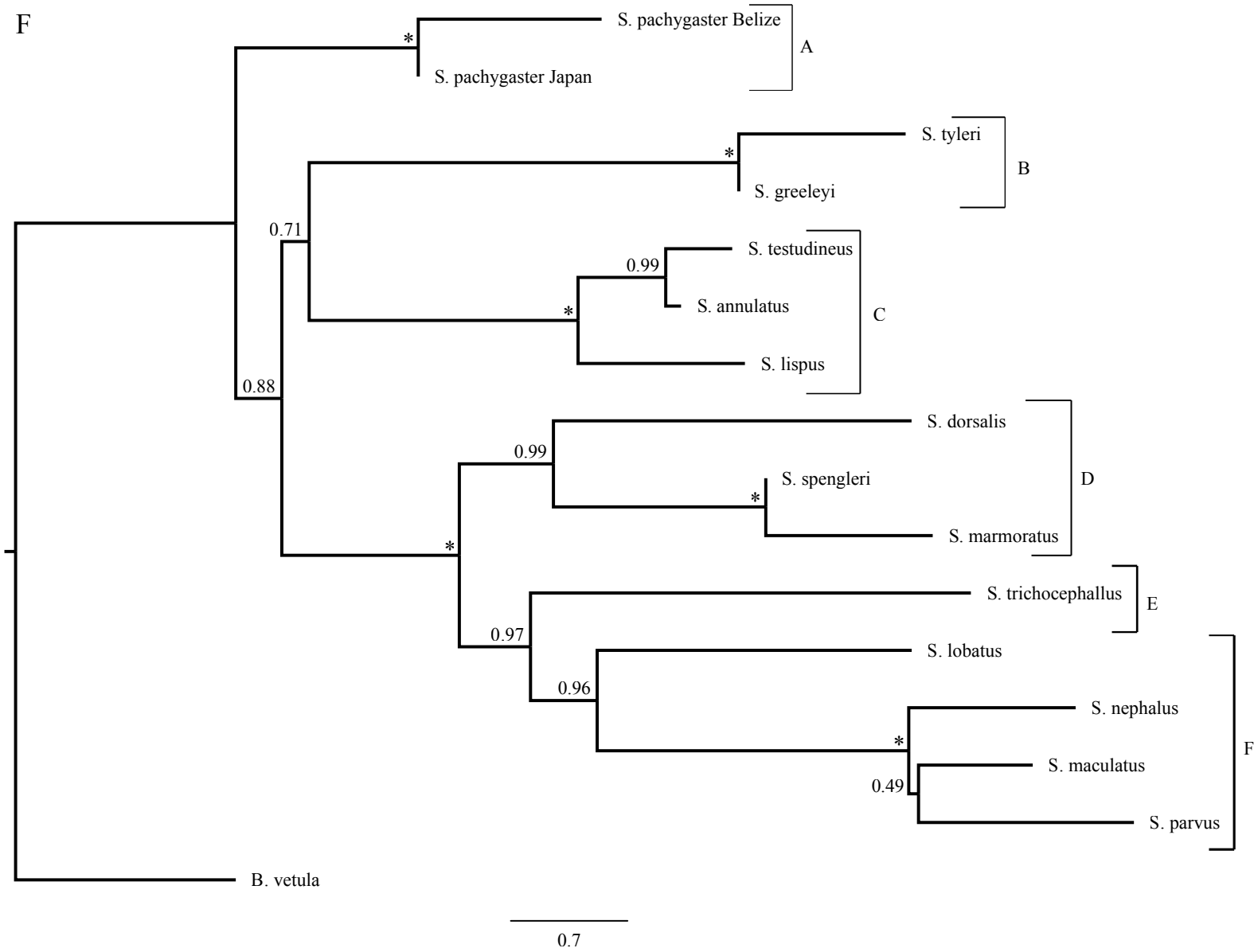
C

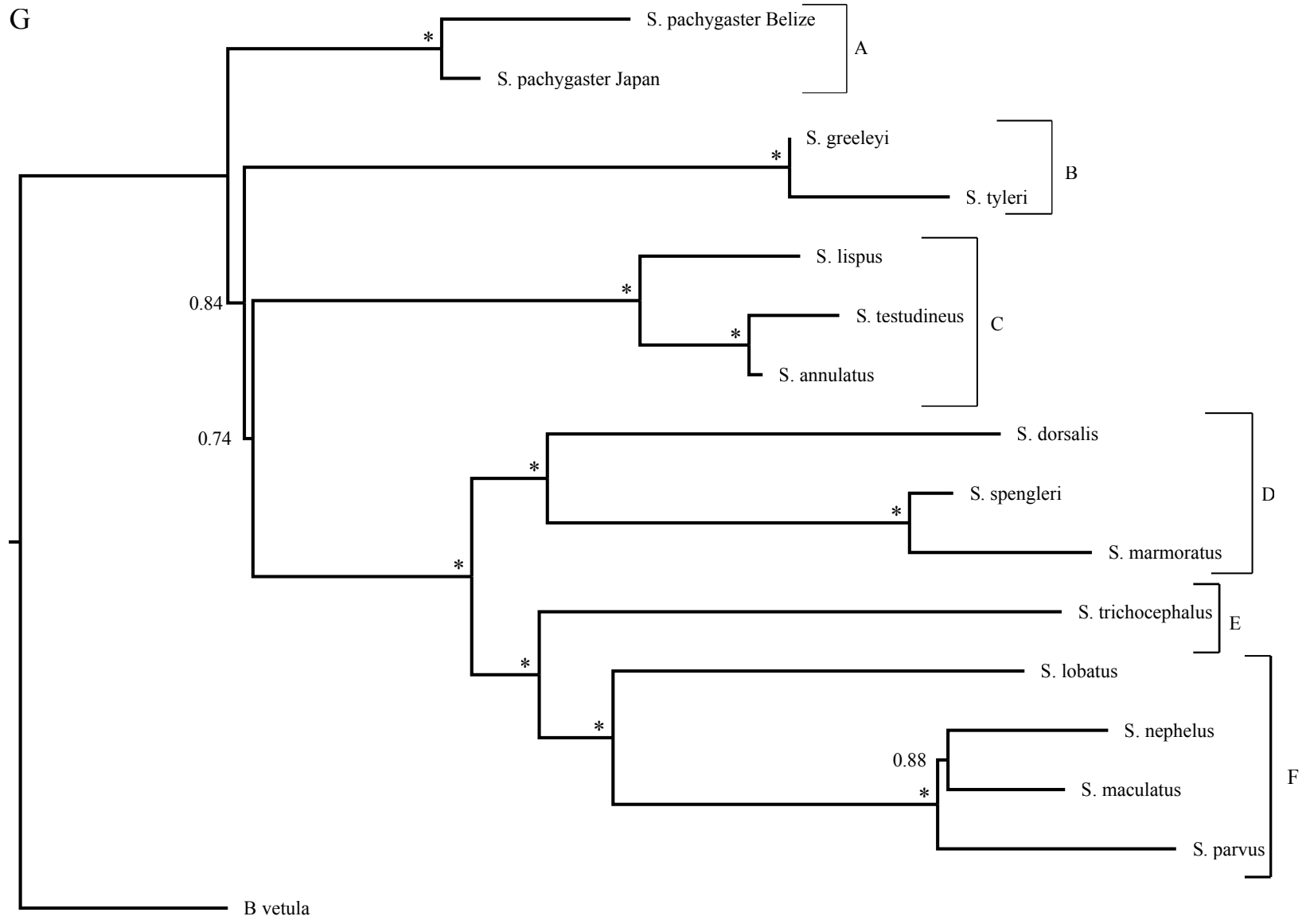

S. greeleyi
S. greeleyi
S. tyleri
S. greeleyi
44
B
S. lobatus
S. lobatus
S. lobatus
89
S. parvus
S. parvus
S. parvus
S. parvus
S. parvus
61
86
35
S. nephelus
S. maculatus
S. maculatus
S. maculatus
72
70
F
S. trichocephalus
S. trichocephalus
S. trichocephalus
95
E
S. dorsalis
S. dorsalis
S. dorsalis
S. spengleri
S. spegleri
S. marmoratus
S. marmoratus
74
D
S. lispus
S. annulatus
S. testidinues
S. testidinues
S. testidinues
S. annulatus
S. annulatus
76
C
S. pachygaster Belize
S. pachygaster Belize
S. pachygaster Belize
S. pachygaster Japan
S. pachygaster Japan
S. pachygaster Japan
75
52
A
0.006

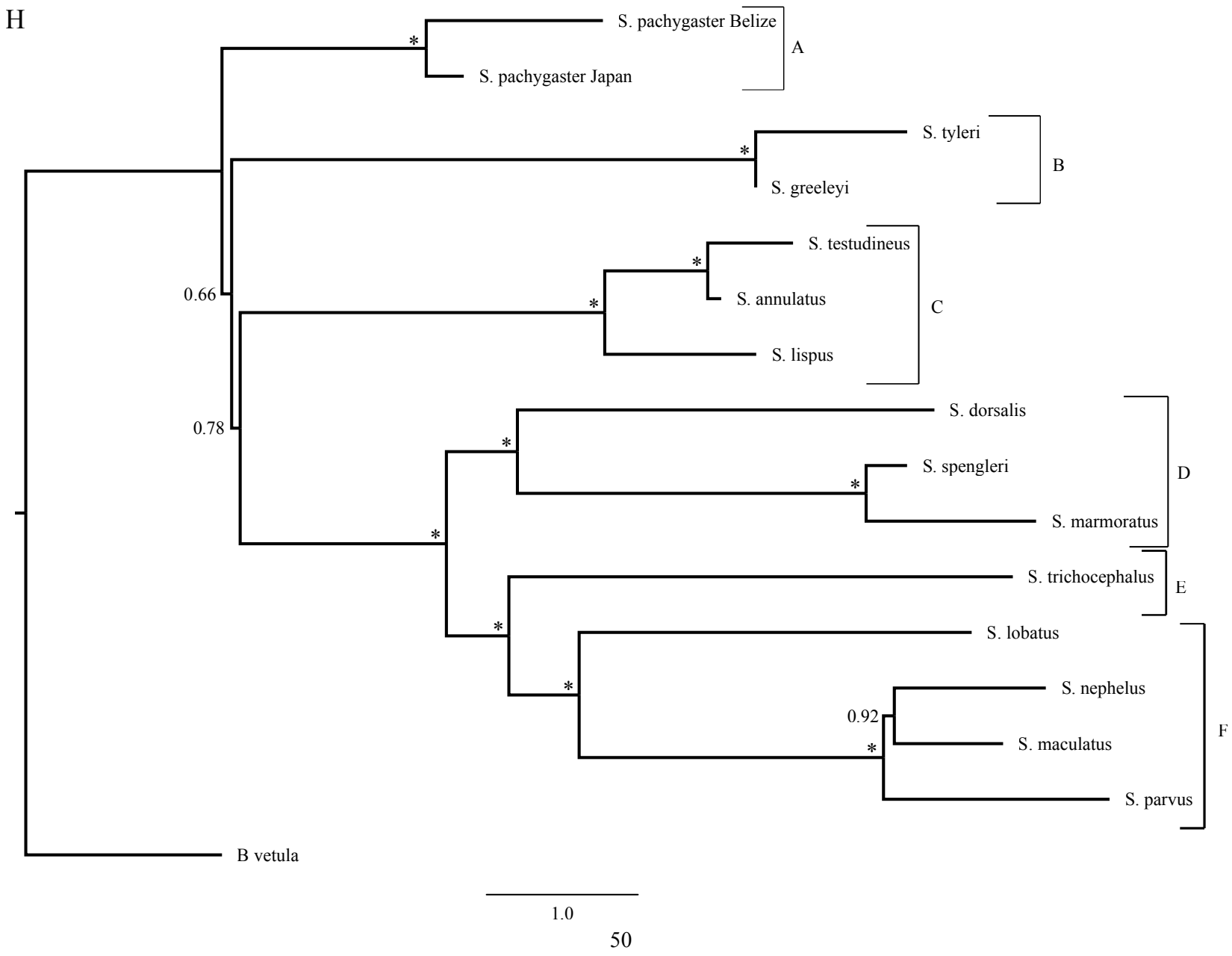

D
S. greeleyi
S. greeleyi
S. tyleri
S. greeleyi
B
88
S. lobatus
S. lobatus
S. lobatus
42
S. parvus
82
S. parvus
S. parvus
56
S. parvus
S. parvus
F
S. nephelus
72
S. maculatus
65
S. maculatus
S. maculatus
96
S. trichocephalus
S. trichocephalus
S. trichocephalus
E
S. spengleri
S. spengleri
S. marmoratus
S. marmoratus
78
D
S. dorsalis
S. dorsalis
S. dorsalis
S. lispus
S. annulatus
S. annulatus
S. annulatus
S. testudineus
72
S. testudineus
S. testudineus
C
S. pachygaster Belize
76
S. pachygaster Belize
S. pachygaster Belize
S. pachygaster Japan
S. pachygaster Japan
53
S. pachygaster Japan
A
0.006

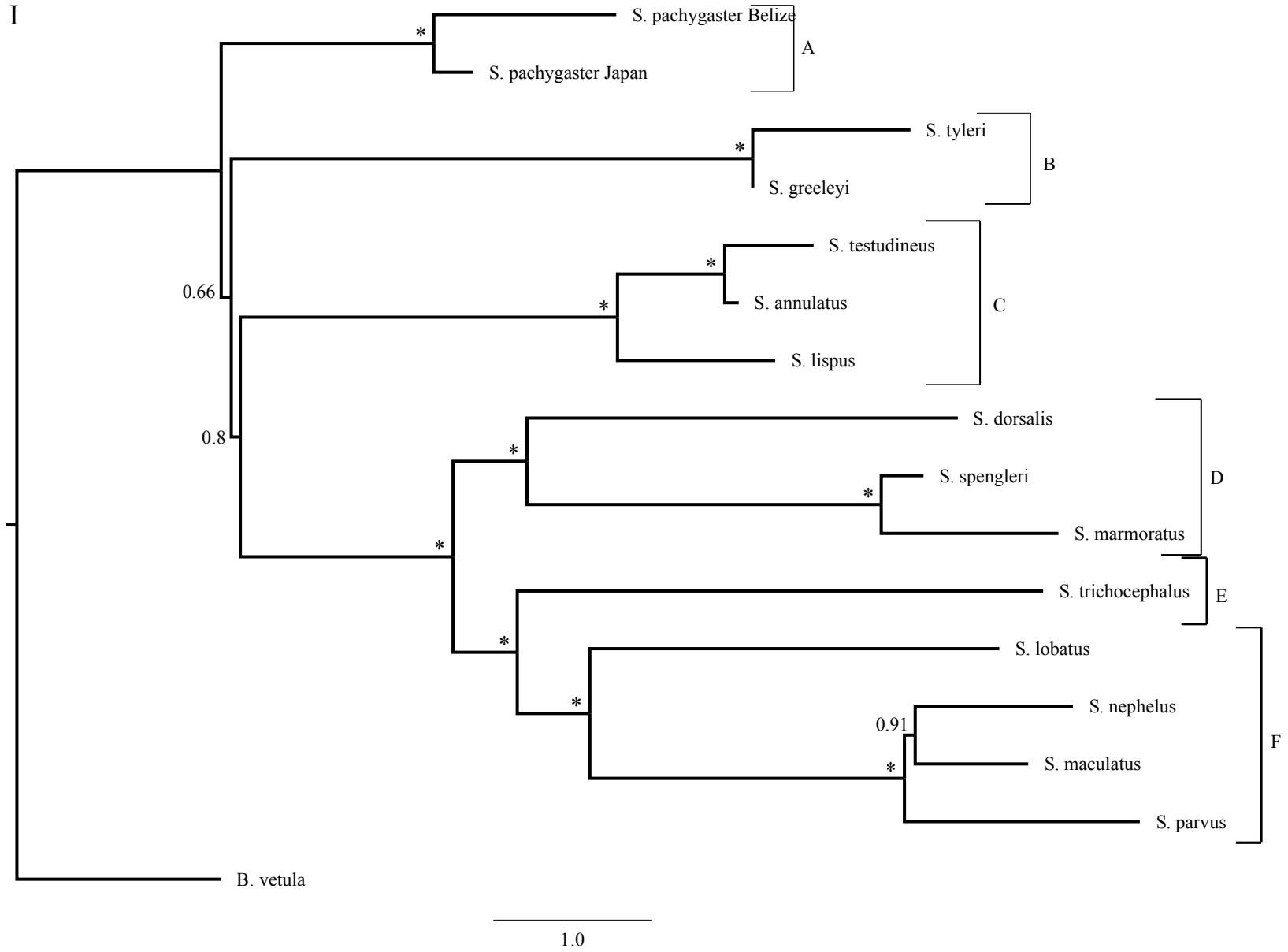
E


S. greeleyi
S. greeleyi
S. tyleri
S. greeleyi
48
B
87
S. lobatus
S. lobatus
S. lobatus
S. parvus
S. parvus
60
S. parvus
77
S. parvus
42
S. parvus
S. nephelus
75
S. maculatus
68
S. maculatus
S. maculatus
F
95
S. trichocephalus
S. trichocephalus
S. trichocephalus
E
S. spengleri
S. spengleri
S. marmoratus
S. marmoratus
78
S. dorsalis
S. dorsalis
S. dorsalis
D
S.lispus
S. annulatus
S. annulatus
S. annulatus
73
S. testudineus
S. testudineus
S. testudineus
C
S. pachygaster Belize
81
S. pachygaster Belize
S. pachygaster Belize
S. pachygaster Japan
53
S. pachygaster Japan
S. pachygaster Japan
A
0.006


F
S. pachygaster Belize
S. pachygaster Japan
A
S. tyleri
S. greeleyi
B
0.71
S. testudineus
0.99
S. annulatus
S. lispus
C
0.88
S. dorsalis
0.99
S. spengleri
S. marmoratus
D
S. trichocephallus
E
0.97
S. lobatus
0.96
S. nephalus
S. maculatus
0.49
S. parvus
F
B. vetula
0.7


G
S. pachygaster Belize
S. pachygaster Japan
A
S. greeleyi
S. tyleri
B
S. lispus
S. testudineus
S. annulatus
C
0.84
0.74
S. dorsalis
S. spengleri
S. marmoratus
D
S. trichocephalus
E
S. lobatus
S. nephelus
0.88
S. maculatus
S. parvus
F
B vetula
1.0


H
S. pachygaster Belize
S. pachygaster Japan
A
S. tyleri
S. greeleyi
B
S. testudineus
S. annulatus
S. lispus
C
0.66
0.78
S. dorsalis
S. spengleri
S. marmoratus
D
S. trichocephalus
E
S. lobatus
S. nephelus
0.92
S. maculatus
S. parvus
F
B vetula
1.0
50


I
S. pachygaster Belize
S. pachygaster Japan
A
S. tyleri
S. greeleyi
B
S. testudineus
S. annulatus
S. lispus
C
0.66
0.8
S. dorsalis
S. spengleri
S. marmoratus
D
S. trichocephalus
E
S. lobatus
S. nephelus
0.91
S. maculatus
S. parvus
F
B. vetula
1.0

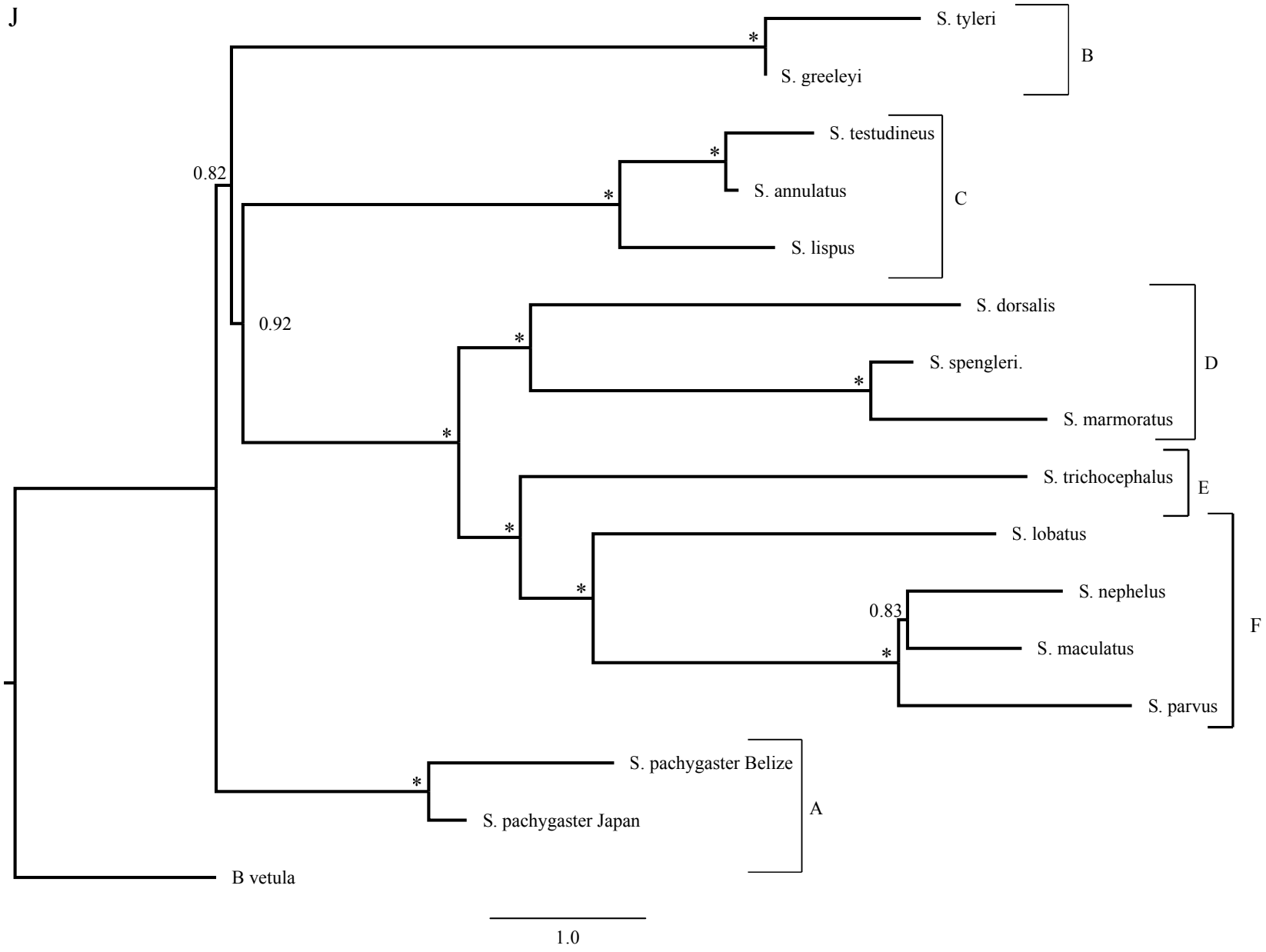
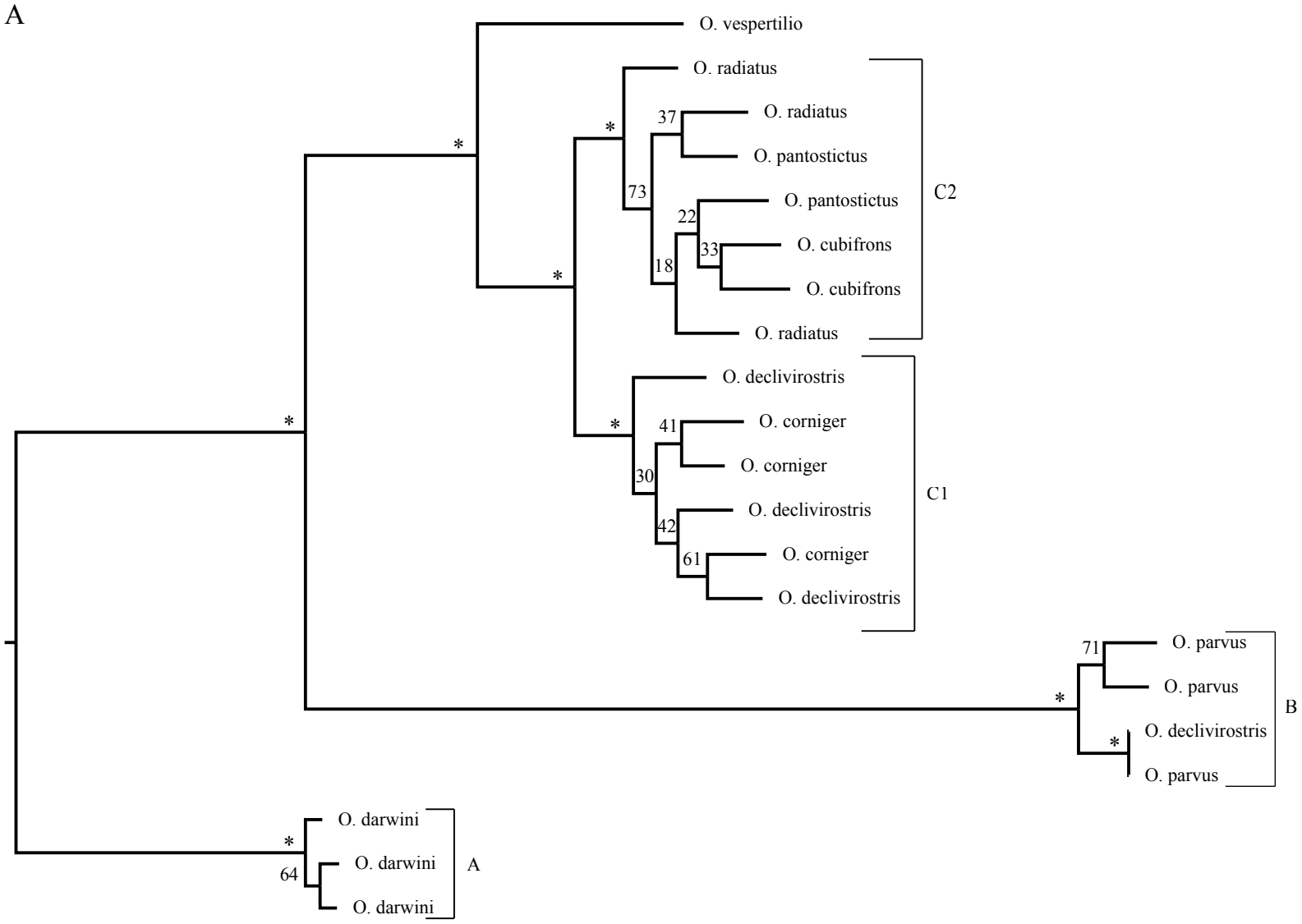
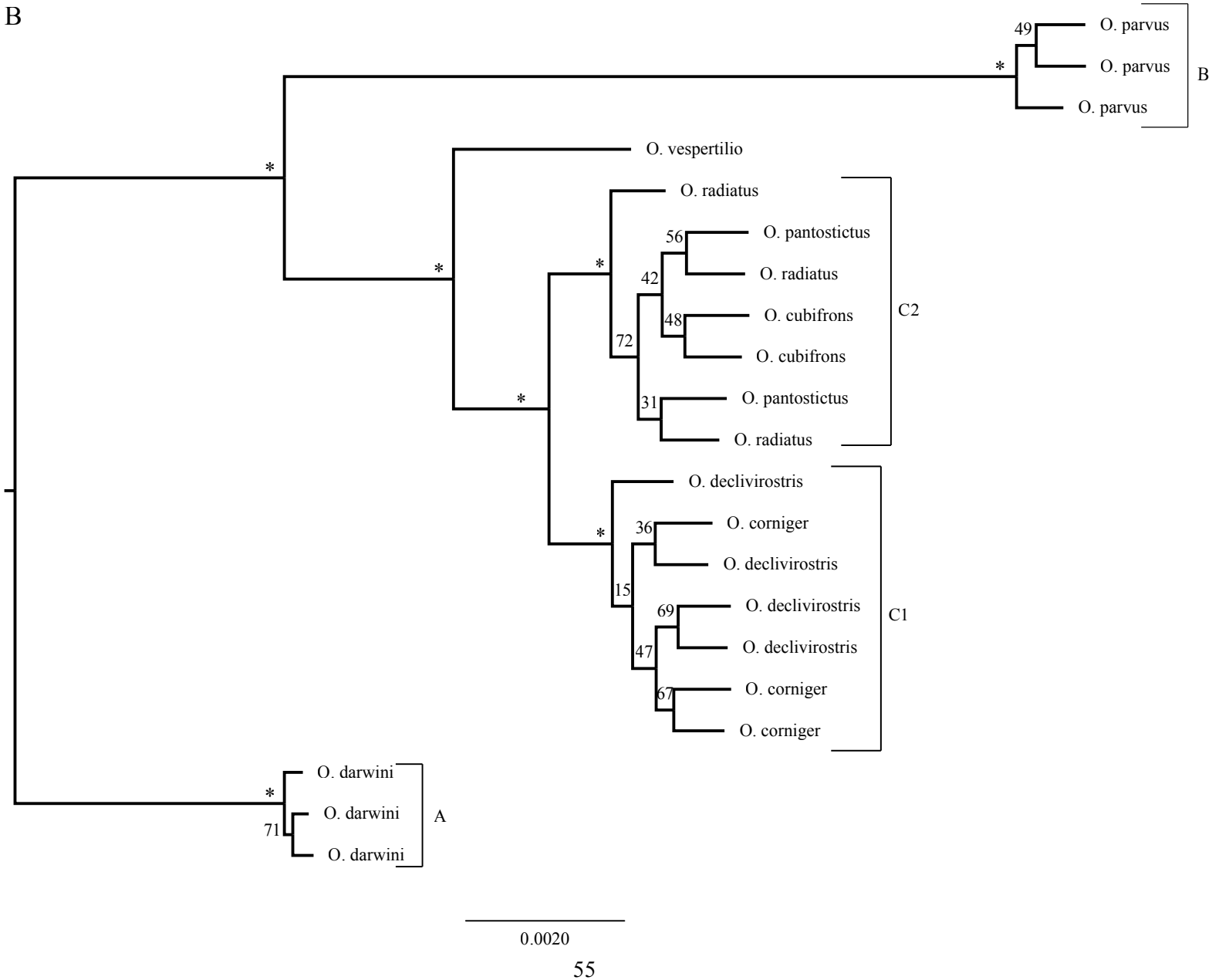
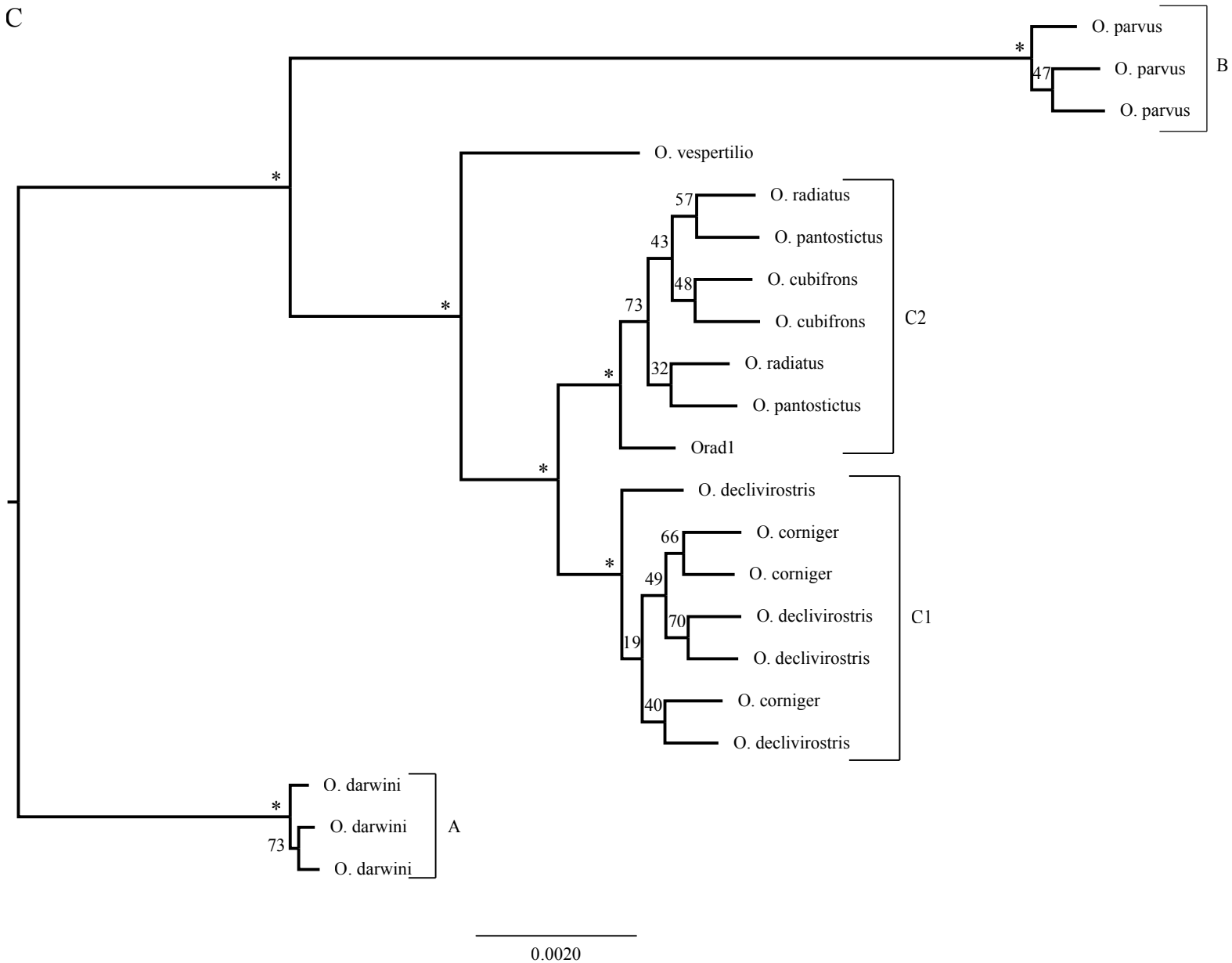

J
S. tyleri
S. greeleyi
B
S. testudineus
S. annulatus
S. lispus
C
0.82
0.92
S. dorsalis
S. spengleri.
S. marmoratus
D
S. trichocephalus
E
S. lobatus
S. nephelus
0.83
S. maculatus
S. parvus
F
S. pachygaster Belize
S. pachygaster Japan
A
B vetula
1.0

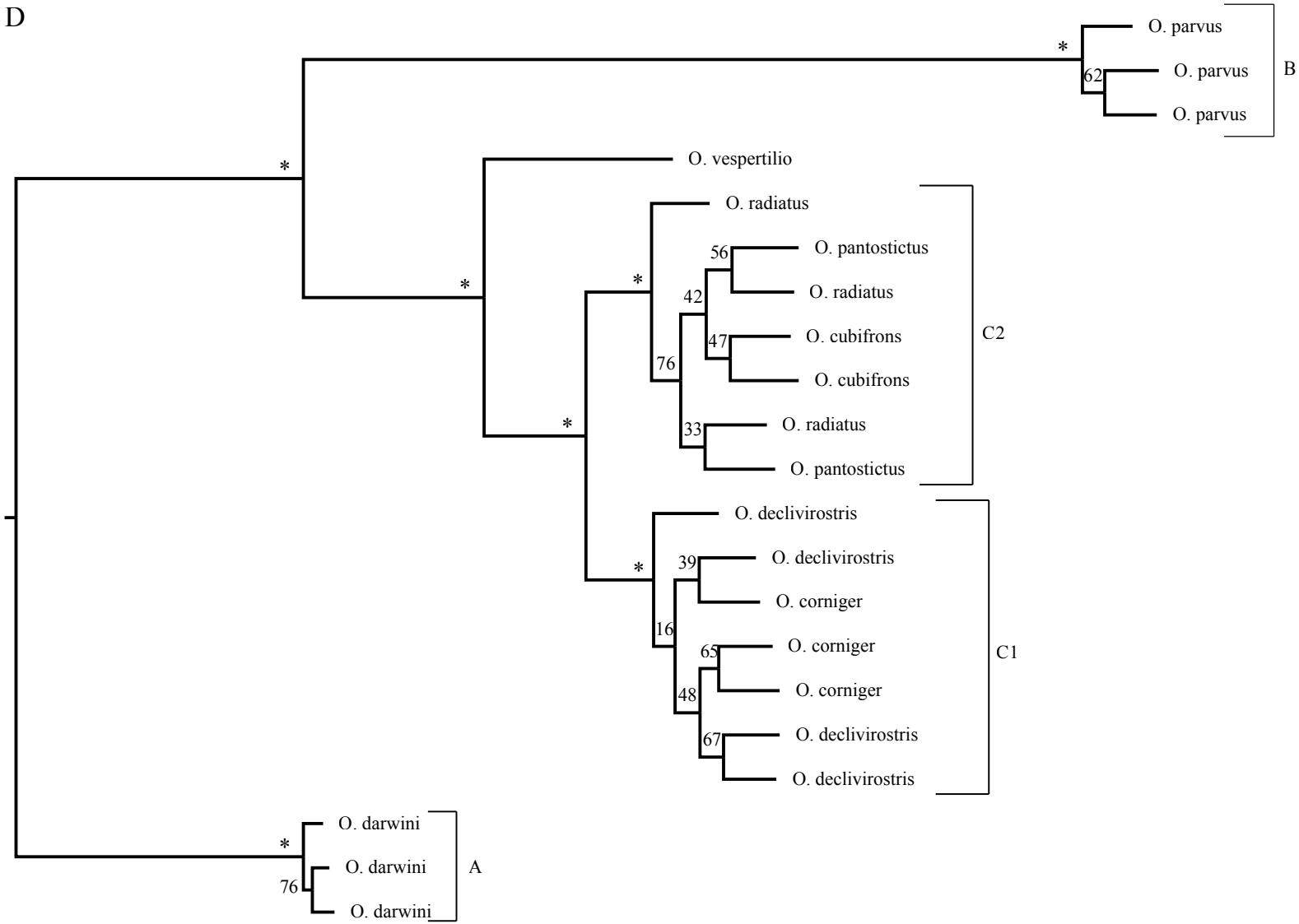
Figure 8: *Ogcocephalus* phylogenetic hypotheses generated using ML with the full concatenated datasets and the species trees generated using Bayesian analysis. Full support ML bootstrap % = 100 and LPP = 1.0 denoted by the *. A) Concatenated ML analysis of the complete dataset B) Concatenated ML analysis of the dataset with 10% missing data C) Concatenated ML analysis of the dataset with 25% missing data D) Concatenated ML analysis of the dataset with 50% missing data E) Concatenated ML analysis of the incomplete dataset. ML trees have had the outgroup pruned for ease of viewing. F) Species tree analysis of the complete dataset G) Species tree analysis of the dataset with 10% missing data H) Species tree analysis of the dataset with 25% missing data I) Species tree analysis of the dataset with 50% missing data J) Species tree analysis of the incomplete dataset. Branch length for ML analysis is scaled at 0.02 mutations per site and branch lengths in the coalescent analysis are measure in coalescent units.


A
O. vespertilio
O. radiatus
O. radiatus
O. pantostictus
O. pantostictus
O. cubifrons
O. cubifrons
O. radiatus
C2
O. declivirostris
O. corniger
O. corniger
O. declivirostris
O. corniger
O. declivirostris
C1
O. parvus
O. parvus
O. declivirostris
O. parvus
B
O. darwini
O. darwini
O. darwini
A
37
73
22
33
18
41
30
42
61
71
64
*
0.0020

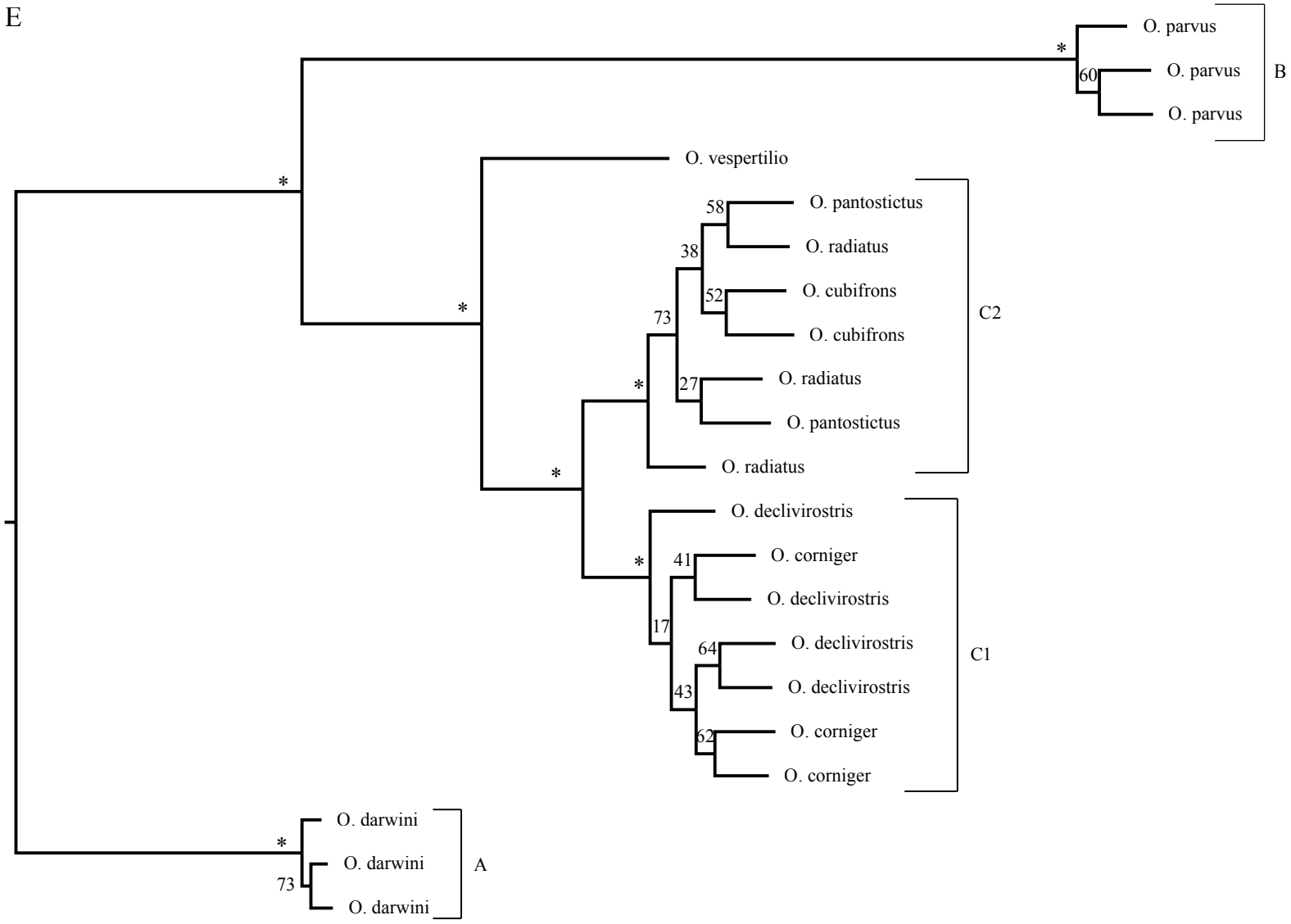

B
O. parvus
O. parvus
O. parvus
B
O. vespertilio
O. radiatus
O. pantostictus
O. radiatus
O. cubifrons
O. cubifrons
O. pantostictus
O. radiatus
C2
O. declivirostris
O. corniger
O. declivirostris
O. declivirostris
O. declivirostris
O. corniger
O. corniger
C1
O. darwini
O. darwini
O. darwini
A
49
56
42
48
72
31
36
15
69
47
67
71
0.0020

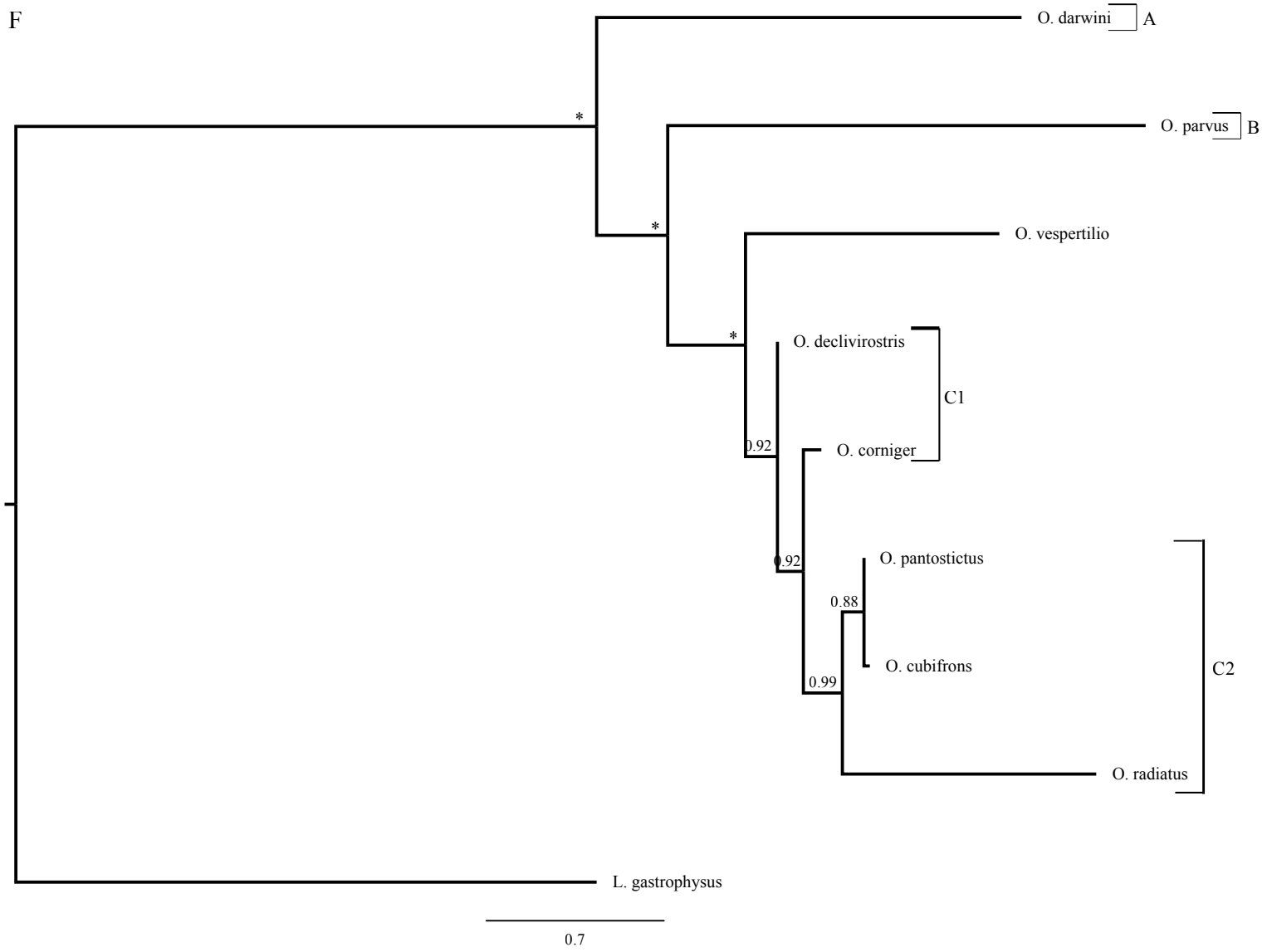

C
O. parvus
O. parvus
O. parvus
B
47
O. vespertilio
O. radiatus
O. pantostictus
O. cubifrons
O. cubifrons
O. radiatus
O. pantostictus
Orad1
C2
57
43
48
73
32
O. declivirostris
O. corniger
O. corniger
O. declivirostris
O. declivirostris
O. corniger
O. declivirostris
C1
66
49
70
19
40
O. darwini
O. darwini
O. darwini
A
73
*
0.0020

D

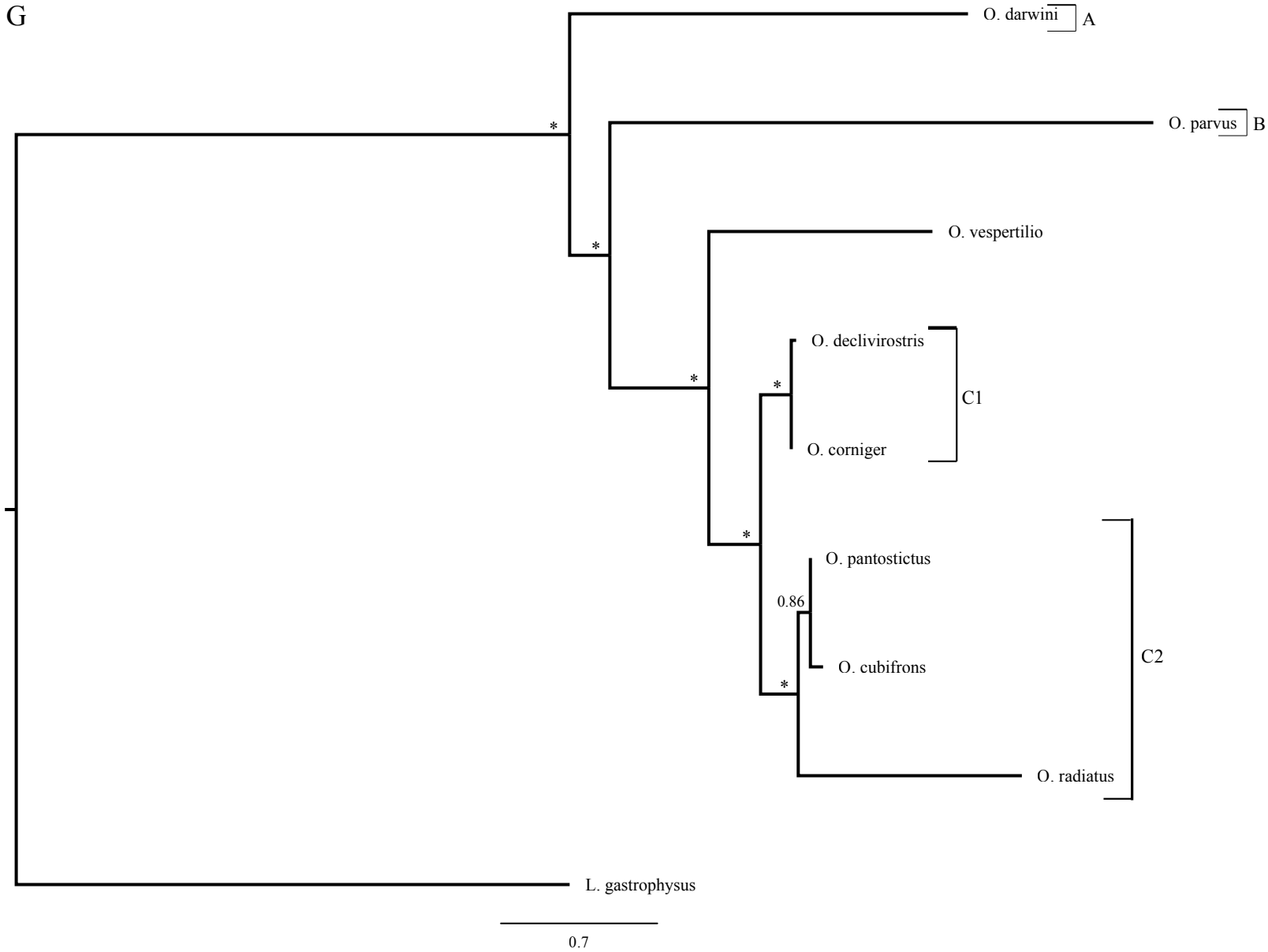

O. parvus
O. parvus
O. parvus
B
62
O. vespertilio
O. radiatus
O. pantostictus
O. radiatus
O. cubifrons
O. cubifrons
O. radiatus
O. pantostictus
C2
56
42
47
76
33
O. declivirostris
O. declivirostris
O. corniger
O. corniger
O. corniger
O. declivirostris
O. declivirostris
C1
39
16
65
48
67
O. darwini
O. darwini
O. darwini
A
76
0.0020

E

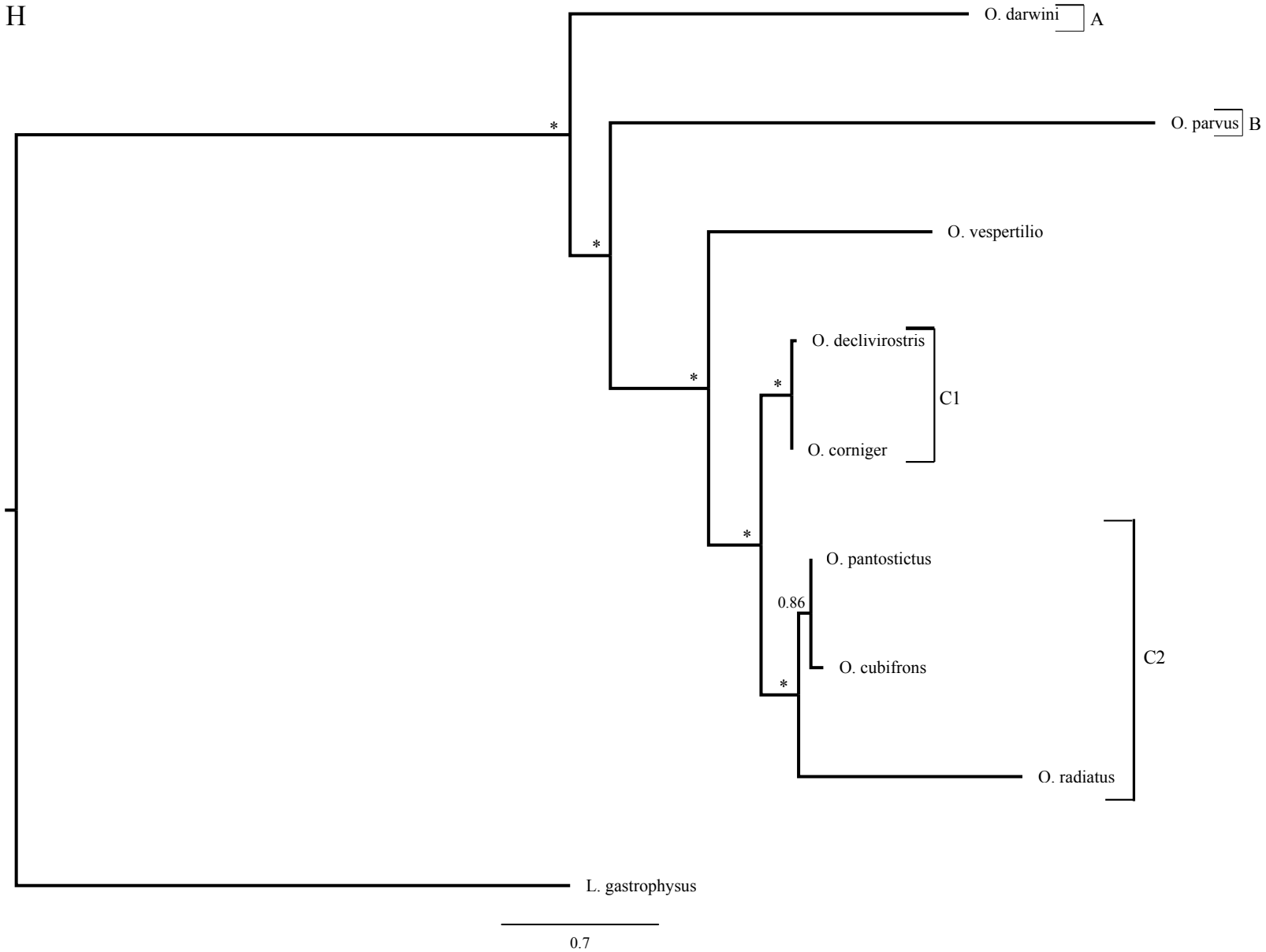

O. parvus
O. parvus
O. parvus
B
60
*
O. vespertilio
*
*
58
O. pantostictus
38
O. radiatus
52
O. cubifrons
73
O. cubifrons
C2
*
27
O. radiatus
O. pantostictus
*
O. radiatus
O. declivirostris
*
41
O. corniger
O. declivirostris
17
64
O. declivirostris
43
O. declivirostris
62
O. corniger
O. corniger
C1
O. darwini
*
O. darwini
73
O. darwini
A
0.0020


F
O. darwini
A
O. parvus
B
O. vespertilio
O. declivirostris
C1
O. corniger
O. pantostictus
O. cubifrons
C2
O. radiatus
L. gastrophysus
*
*
*
0.92
0.92
0.88
0.99
0.7

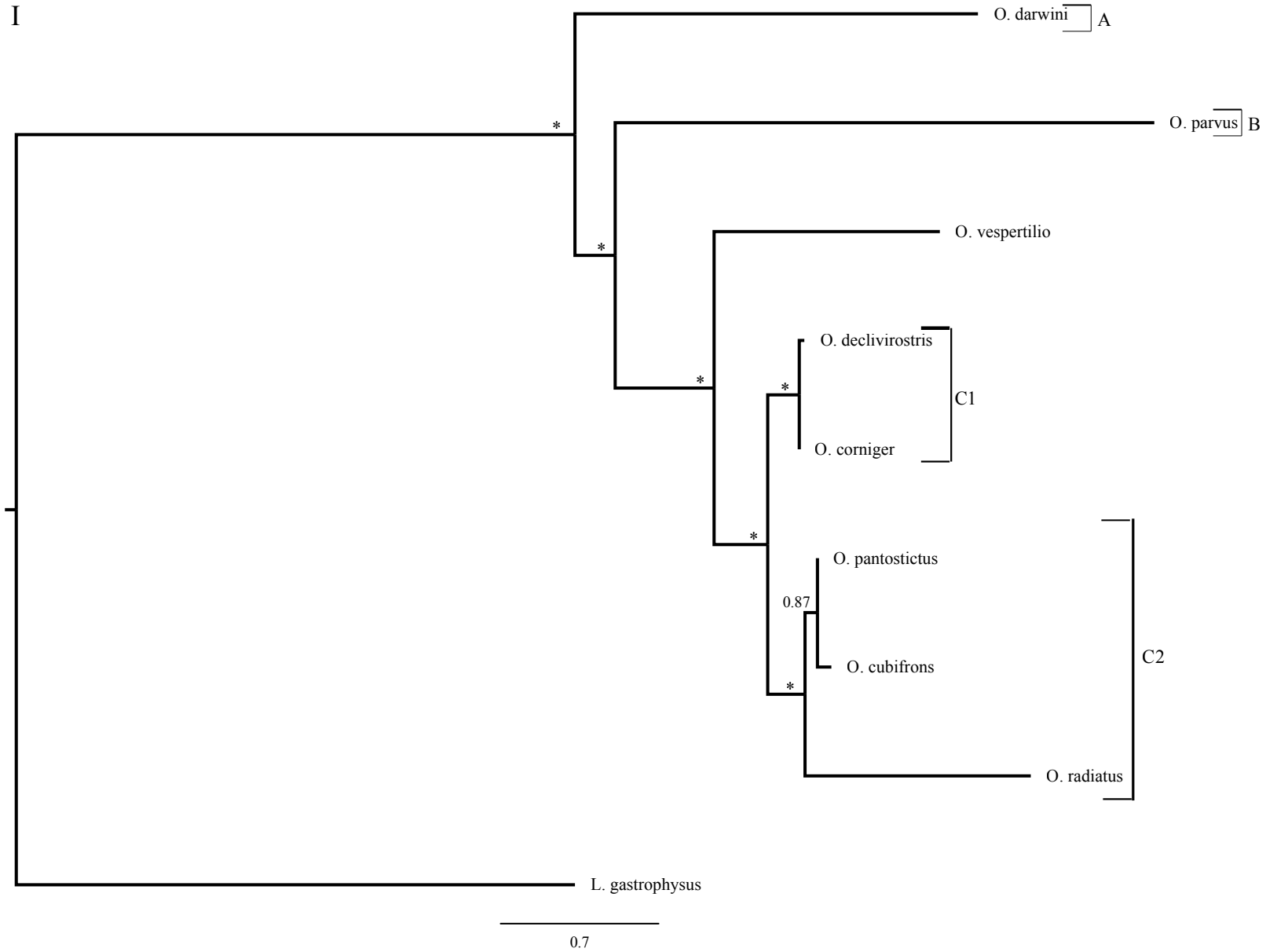
G

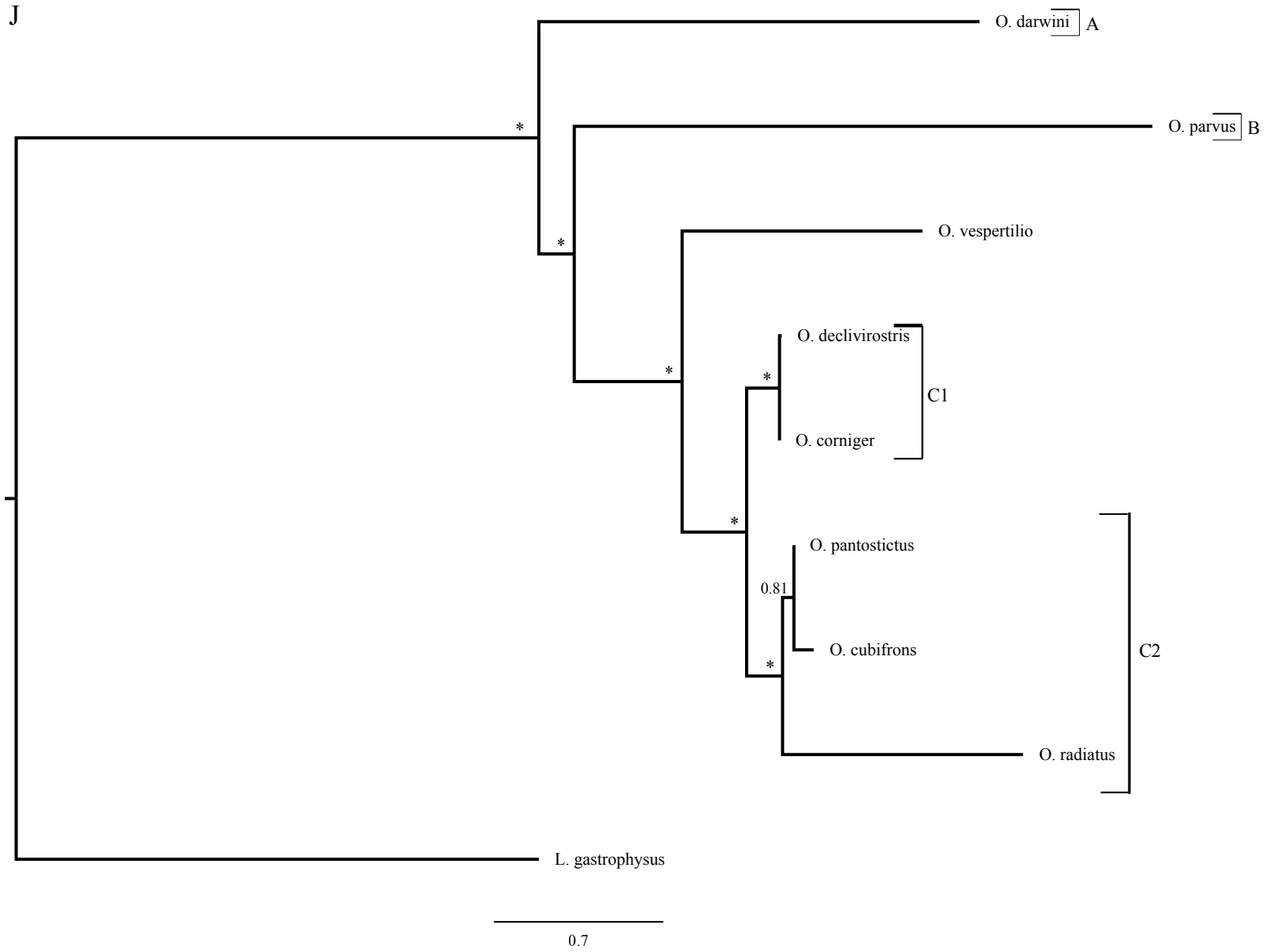

O. darwini
A
O. parvus
B
O. vespertilio
O. declivirostris
C1
O. corniger
O. pantostictus
0.86
O. cubifrons
C2
O. radiatus
L. gastrophysus
0.7

H


O. darwini
A
O. parvus
B
O. vespertilio
O. declivirostris
C1
O. corniger
O. pantostictus
0.86
O. cubifrons
C2
O. radiatus
L. gastrophysus
0.7

I


O. darwini
A
O. parvus
B
O. vespertilio
O. declivirostris
C1
O. corniger
O. pantostictus
0.87
O. cubifrons
C2
O. radiatus
L. gastrophysus
0.7


J
O. darwini
A
O. parvus
B
O. vespertilio
O. declivirostris
O. corniger
C1
O. pantostictus
0.81
O. cubifrons
C2
O. radiatus
L. gastrophysus
0.7