Molecular Phylogenetics and Evolution 136 (2019) 87–103

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Assessing the diversity of Western North American *Juga* (Semisulcospiridae, Gastropoda)

Ellen E. Strong[a,*], Nathan V. Whelan[b,c]

[a] Department of Invertebrate Zoology, Smithsonian Institution, National Museum of Natural History, 10th and Constitution Ave NW, Washington DC 20560, USA
[b] United States Fish and Wildlife Service, Southeast Conservation Genetics Lab, Warm Springs Fish Technology Center, Auburn, AL 36849, USA
[c] School of Fisheries, Aquaculture and Aquatic Sciences, Auburn University, Auburn, AL 36849, USA

A B S T R A C T

Juga is a genus of freshwater gastropods distributed in Pacific and Interior drainages of the Pacific Northwest from central California to northern Washington. The current classification has relied heavily on features of the shell, which vary within and across drainages, and often intergrade without sharp distinctions between species. The only previous molecular analysis included limited population sampling, which did not allow robust assessment of intra- versus interspecific levels of genetic diversity, and concluded almost every sampled population to be a distinct OTU. We assembled a multilocus mitochondrial (COI, 16S) and nuclear gene (ITS1) dataset for ~100 populations collected across the range of the genus. We generated primary species hypotheses using ABGD with best-fit model-corrected distances and further explored our data, both individual gene partitions and concatenated datasets, using a diversity of phylogenetic and species delimitation methods (Bayesian inference, maximum likelihood estimation, StarBEAST2, bGMYC, bPTP, BP&P). Our secondary species delimitation hypotheses, based primarily on the criterion of reciprocal monophyly, and informed by a combination of geography and morphology, support the interpretation that *Juga* comprises a mixture of geographically widespread species and narrow range endemics. As might be expected in taxa with low vagility and poor dispersal capacities, analysis of molecular variance (AMOVA) revealed highly structured populations with up to 80% of the observed genetic variance explained by variation between populations. Analyses with bGMYC, bPTP, and BP&P appeared sensitive to this genetic structure and returned highly dissected species hypotheses that are likely oversplit. The species diversity of *Juga* is concluded to be lower than presently recognized, and the systematics to require extensive revision. Features of the teleoconch considered significant in species-level and subgeneric classification were found to be variable within some species, sometimes at a single site. Of a number of potentially new species identified in non-peer reviewed reports and field guides, only one was supported as a distinct OTU.

1. Introduction

As for many invertebrates, freshwater snails are understudied relative to their diversity. Systematic studies of freshwater snails from northwestern North America have focused primarily on groundwater-dependent springsnails and pebblesnails in the families Lithoglyphidae, Amnicolidae, and Hydrobiidae. These studies have revealed that their diversity frequently has been underestimated and that many species have restricted distributions and are highly imperilled (e.g., Hershler and Frest, 1996; Hershler et al., 2003, 2007; Hershler and Liu, 2010). Conversely, some populations hypothesized to represent undescribed species have been found to be broadly disjunct members of widely distributed species already described (e.g., Liu et al., 2015, 2016). A conspicuous member of the freshwater snail fauna from the northwestern United States is the genus *Juga* H. and A. Adams, 1854 (Semisulcospiridae). The systematics of *Juga* has received little attention since the 1940's, and the diversity has not been robustly explored using molecular methods. Thus, the taxonomic status of these species is uncertain.

Juga comprises large-bodied snails distributed in Pacific and Interior drainages from central California to northern Washington, mainly at low to medium elevations, and are rare east of the Cascades. They are found in a diversity of habitats, including perennial seeps, springs and spring runs, groundwater-influenced creeks and streams, as well as large rivers, usually in cold, oligotrophic, well-oxygenated waters with stable bottoms. Some species appear tolerant of silt and low energy conditions and may occur in ponds and lakes (Furnish et al., 1997). With individuals capable of reaching 3.5 cm in adult shell length and population densities as high as 1500 individuals/m^2, they are prominent members of their benthic macroinvertebrate communities and can comprise as much as 90% of the invertebrate biomass in some streams (Hawkins and Furnish, 1987). Formerly placed in the cerithioidean family Pleuroceridae, molecular and morphological data support their affinities among the Asian genera *Hua*, *Koreoleptoxis* and *Semisulcospira* (Strong and Frest, 2007; Strong and Köhler, 2009; Köhler, 2017) in the Semisulcospirinae, which was elevated to family rank by Strong and Köhler (2009). It was hypothesized that *Juga* dispersed to North America across the Thulean land bridge and diverged from their Asian

* Corresponding author.
E-mail address: StrongE@si.edu (E.E. Strong).

https://doi.org/10.1016/j.ympev.2019.04.009
Received 10 August 2018; Received in revised form 6 April 2019; Accepted 7 April 2019
Available online 08 April 2019
1055-7903/ Published by Elsevier Inc.

relatives during the Palaeocene-Eocene thermal maximum ~55 million years ago (Strong and Köhler, 2009).

The number and circumscription of recognized *Juga* species has fluctuated over the years (e.g., Henderson, 1935; Goodrich, 1942; Burch, 1989) and there remains no consensus; eleven species are considered valid in the taxonomic authority list of Johnson et al. (2013), while the conservation watch list NatureServe (2017) recognizes as many as 25 species or subspecies. As for other freshwater gastropods, this confusion has resulted from the reliance on features of the shell (size, color, strength and persistence of spiral and axial ornamentation) for circumscribing and identifying species. These features may vary within species of *Juga* across drainages and from headwater to downstream habitats (Goodrich, 1942), likely reflecting a combination of ecophenotypic and heritable variation (Minton et al., 2008; Whelan et al., 2012). Furthermore, *Juga* species are thought to have diverged primarily through isolation by distance and genetic drift (Campbell et al., 2016), such that speciation may not have been accompanied by morphological differentiation. Consequently, shell characters may intergrade without sharp distinctions between species, further complicating efforts to classify them.

Classification of *Juga* above the species level has also relied on shell characters. Until recently, extant species of *Juga* were distributed among three subgenera based on ornament of the early shell (Taylor, 1966): *Juga s.s.* with plications (i.e. axial folds or ribs) on the early teleoconch, *Calibasis* Taylor, 1966 with lirate (i.e. spiral) early sculpture, and *Oreobasis* Taylor, 1966 with weak to no early sculpture. Variability in ornament development, compounded by erosion of the early whorls in many specimens, has made it challenging to distinguish consistently between species and subgenera and to assign individuals confidently to them. Based on an anatomical study of the type species of each subgenus, Strong and Frest (2007) found that *Juga s.s.* and *Oreobasis* did not differ significantly in features of reproductive or alimentary anatomy, including of the radula, and synonymized the two.

A series of non-peer reviewed reports and field guides has been published in the last 25 years (e.g., Frest and Johannes, 1993, 1995, 1999), suggesting that the number of species currently considered valid is a significant underestimate. Many potentially new species have been identified based on a combination of features of the shell, their distribution, and ecology, and have highly localized distributions. However, these undescribed entities (e.g., Basalt Juga, Blue Mountains Juga, Brown Juga, Cinnamon Juga, Crooked River Juga, Indian Ford Juga, One-Band Juga, Opal Springs Juga, Purple Juga, Three-Band Juga) have been abandoned in open nomenclature and their validity has not been robustly tested with molecular tools. Most of these putative species have never been figured and information about their distributions is obscure and often not easily retrievable. Reference to the same putative species by different names in different reports has generated even more confusion (see e.g., Frest and Johannes, 1999, 2010-2011).

Populations of *Juga* are experiencing increasing threats from habitat degradation and loss, and some are of conservation concern (e.g., Furnish and Monthey, 1998, 1999; Furnish, 2007). Pollution, nutrient runoff, trampling by cattle, stream modification and urbanization, siltation from logging, and diversion or impoundment for agriculture, livestock, drinking water, recreation, and other private and public uses pose the most significant threats (Taylor, 1981; Frest and Johannes, 1995; Furnish, 2007). Like other freshwater Caenogastropoda, they are dioecious and deposit egg capsules on hard substrates that hatch as crawling juveniles (Strong et al., 2008). It is thought that *Juga* individuals may require three years to reach sexual maturity with life spans of five to seven years on average (Furnish, 1990). This combination of slow maturation times, low fecundity and limited dispersal capacity hinders their ability to recover from human-mediated impacts (Strong et al., 2008). Although no species of *Juga* currently are federally listed as Threatened or Endangered, nor considered to be species of Special Concern, some species have been identified by state agencies as meriting protection. For example, the state of California recognizes *J. acutifilosa*, *J. chacei*, *J. occata,* and *J. orickensis* to be "special animals", and they are included on the California list of "species at risk". Eight species or subspecies (*J. acutifilosa*, *J. hemphilli hemphilli*, *J. hemphilli dallesensis*, *J. hemphilli maupinensis*, *J. newberryi*, *J. silicula*, Indian Ford Juga, Basalt Juga) are identified in the Interagency Special Status /Sensitive Species Program (ISSSSP), a program for conservation and management established by the U.S. Forest Service and Oregon/Washington Bureau of Land Management. Of the 25 species or subspecies recognized by NatureServe (2017), all but two are ranked critically imperilled (G1) or imperilled (G2), five of which have not been formally described. Two of these potentially new species (Basalt Juga, Cinnamon Juga) have been afforded federal protection under the "survey and manage" provisions of the Northwest Forest Plan (USDA and USDI, 1994, 2000) and petitioned for federal listing given their restricted distributions, declining population numbers and threats from highway and railway development, pollution, logging, grazing, and water diversions (USFWS, 2011). The 12-month finding on the petition concluded that species not formally described could not be considered listable entities pending the necessary genetic comparisons to establish their taxonomic validity (USFWS, 2012). Indeed, a potentially new species of Hydrobiidae included in the petition, the Columbia Duskysnail (*Colligyrus* n. sp. 1), was subsequently found to be conspecific with another already recognized species, *Colligyrus greggi* (Liu et al., 2015), demonstrating the importance of thorough systematic revision prior to formal listing.

The only molecular analysis of *Juga* conducted to date is that of Campbell et al. (2016). Their dataset comprised partial COI sequences for 103 individuals and partial 16S sequences for 35 individuals. Sequences were analyzed using maximum parsimony and Bayesian inference; operational taxonomic units (OTUs) were defined as "distinctive lineages" based on vaguely articulated molecular and morphological criteria. "Reference lots" (Campbell et al., 2016: 162-163) deposited in museum collections are paravouchers (*sensu* Groenenberg et al., 2011). Consequently, it is not possible to establish an explicit link between genotype and phenotype among the specimens sequenced in their study. This is crucial for understanding variation in populations and assessing the utility of shell characters for species circumscription and identification. The conclusion of their analysis was that the sampled populations of *Juga* represented, *"…numerous undescribed OTUs and high local endemism…"* (Campbell et al., 2016: 168). No fewer than 33 OTUs were recognized, and almost every population sampled was concluded to represent a distinct OTU. This interpretation was supported despite the fact that some populations were represented by only a single sequenced specimen, and that not all OTUs, as defined, were monophyletic (e.g., *Juga Calibasis* OTU 7, fig. 2; *Juga Oreobasis* OTU 3, fig. 3; Campbell et al., 2016). The recognized OTUs frequently were found to be separated from their closest relatives by less than 1% uncorrected pairwise distance in COI, and by as little as 0.15% between some recognized subspecies (Campbell et al., 2016; supplementary file 3). No OTU was found to be geographically widespread. They also concluded that none of the three extant named subgenera were monophyletic but that clades of equivalent rank merited recognition. They did not critically explore their data using any of the methods of species delimitation now in routine use for single-locus datasets [e.g., Automatic Barcode Gap Discovery (ABGD), Poisson Tree Processes (PTP), and Bayesian general mixed Yule-coalescent (bGMYC)] to test the validity of their conclusions. Furthermore, phylogenetic hypotheses were inferred with maximum parsimony or with Bayesian analysis; in the latter, the COI dataset was not partitioned by codon position, which may negatively affect accuracy of the resulting phylogenetic hypotheses (Brown and Lemmon, 2007; Kainer and Lanfear, 2015).

Accurately assessing the diversity of *Juga* is important given that the number of actual species could represent three times the number currently recognized, many meriting conservation attention. Thus, the goal of this study is to assess the diversity of *Juga* using an expanded molecular dataset with dense geographic and population sampling, and to

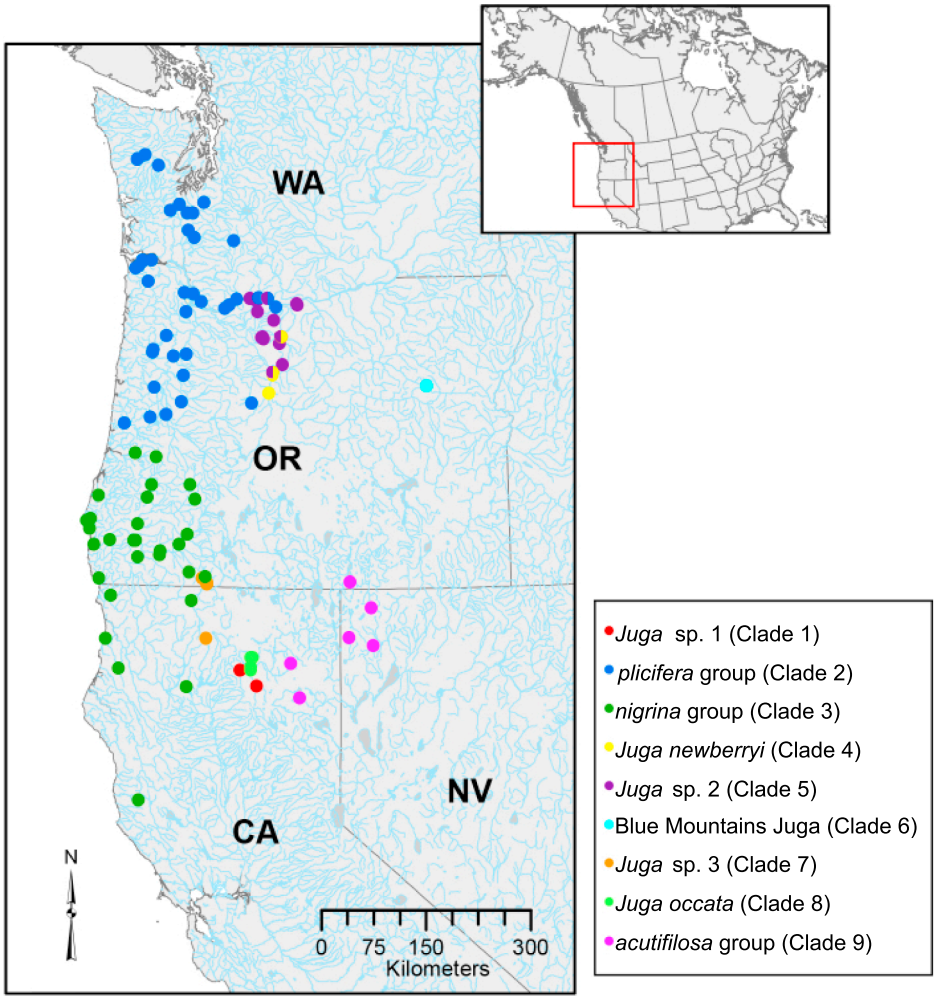

Fig. 1. Map of sampling localities. Colors correspond to clades in Fig. 2. Two-toned circles indicate sites where clades co-occur. Clade number as defined on Fig. 2, indicated in parentheses.

explore our dataset using a variety of species delimitation methods. Our aim is to determine if *Juga* represents few, highly variable, widespread species, numerous narrow-range endemics, or a mixture of the two, and to assess the utility of shell morphology in establishing and recognizing taxa at and above the species level.

2. Materials and methods

2.1. Sample collection

Samples were collected or obtained from colleagues from ~100 sites spanning *Juga*'s known distribution from the coast ranges of central California in the south, to the Olympic Peninsula in western Washington in the north, and as far east as the Blue Mountains in eastern Oregon and the Black Rock Desert in northwestern Nevada (Fig. 1). Topotypic or near topotypic specimens of all 28 available nominal species group names were sampled. In some cases, the original type locality was imprecisely defined, or populations at the type locality have been extirpated. In these cases, near-topotypic populations and morphotypes closely matching the types were sought. The included samples also span the known ecological diversity of the genus, from springs and spring runs, to large rivers, lakes and ponds. Specimens were heat-shocked (Fukuda et al., 2008) in the field prior to preservation in 95% EtOH to prevent the animals from retracting into their shells. Sample site information, USNM voucher registration numbers, NMNH Biorepository tissue identification numbers, and GenBank accession numbers, are provided in Tables 1 and S1.

2.2. Molecular data generation

Two mitochondrial genes, COI and 16S, were sequenced for each individual and a subset of individuals were sequenced for the nuclear ribosomal internal transcribed spacer 1 (ITS1) gene. Whole genomic DNA was extracted from a ~ 1 mm^3 tissue clip of the foot using an Autogenprep965 (Autogen, Holliston, MA) automated phenol:chloroform extraction with a final elution volume of 50 µL. A 691 base pair (bp) fragment of COI was amplified using the jgLCOI primer (Geller et al., 2013) in combination with a newly designed reverse primer for cerithioideans (Cerithioid_COIR: TATWCCAAATCCNGGWARAAT); a 504-512bp fragment of 16S was amplified with the universal 16SAR/BR primers (Palumbi et al., 1991). PCR reactions for mitochondrial genes were performed with 1 µL of undiluted DNA template in 20 µL reactions. Reaction volumes for COI consisted of 10 µL of Promega Go-Taq Hotstart Master Mix, 0.15 µM each primer, 0.25 µg/µL BSA, 1.25% DMSO and an amplification regime of an initial denaturation at 95 °C for 7 min, followed by 45 cycles of denaturation at 95 °C for 45 s, annealing at 42 °C for 45 s, extension at 72 °C for 1 min and a final extension at 72 °C for 3 min. Reaction volumes for 16S were 1x Biolase (Bioline, Taunton, MA) reaction buffer, 500 µM dNTPs, 3 mM MgCl2, 0.15 µM each primer, 0.25 µg/µL BSA, 1 unit Biolase DNA polymerase and an amplification regime of initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation at 95 °C for 30 s, annealing at

Table 1
Sampling sites, arranged by species.

Site	County	State	Latitude	Longitude
Clade 1: Juga sp. 1				
Goose Valley Spring	Shasta	CA	40.92954	−121.726
Hat Creek at Bridge Campground	Shasta	CA	40.73002	−121.437
Clade 2: *plicifera* group				
Ames Creek	Lewis	WA	46.45156	−121.993
Benson State Recreation Area, along Interstate 84, near Multnomah Falls	Multnomah	OR	45.57697	−122.126
Black River	Grays Harbor	WA	46.82987	−123.186
Black River, north of bridge along SE 128 Ave, Littlerock	Thurston	WA	46.90204	−123.024
Burnt Bridge Creek. Bridge at NE 110th Ave, Vancouver	Clark	WA	45.65811	−122.56
Calapooia River at McKercher Park	Linn	OR	44.35853	−122.876
at confluence with Bear Creek, Maki Road	Clatsop	OR	46.15432	−123.667
Cowlitz River at Spencer Rd Trout Hatchery	Lewis	WA	46.48514	−122.726
end of Crossroads Lane (boat ramp)	Lane	OR	44.18786	−123.144
Drift Creek on Cascade Highway NE	Marion	OR	44.96928	−122.809
Green Mountain Road, just upstream of confluence of North Fork and South Fork Klaskanine rivers (approx 300 m upstream of the confluence)	Clatsop	OR	46.08651	−123.741
just downstream of Gnat Creek Fish Hatchery, along US Hwy 30	Clatsop	OR	46.17701	−123.503
just downstream of the mouth of Hood River	Hood River	OR	45.71486	−121.513
just downstream of US Hwy 20 (FS campground)	Deschutes	OR	44.35671	−121.611
Lake Quinault at Falls Creek Campground, from boat launch east to private docks along South Shore Road	Grays Harbor	WA	47.47012	−123.846
Lake Quinault at Gatton Creek Campground, along rocky beach	Grays Harbor	WA	47.47381	−123.839
Le Bong Creek at Stevenson	Skamania	WA	45.69756	−121.913
Long Tom River at Poodle Creek Road	Lane	OR	44.14338	−123.43
Major Creek	Klickitat	WA	45.71542	−121.351
Mary's River, Mary's River Park, Philomath, Oregon	Benton	OR	44.53434	−123.374
Mill Creek, Buell County Park, near Sheridan, Oregon	Polk	OR	45.02294	−123.418
Mill Creek at Mill Creek Park	Polk	OR	44.9866	−123.426
Nisqually River	Thurston	WA	46.93345	−122.561
NW 179th St.	Clark	WA	45.75032	−122.715
Old US Hwy 30	Clatsop	OR	46.16633	−123.673
Quinault River, just N of S. Shore Rd., between Fletcher and Bunch Canyons	Jefferson	WA	47.52975	−123.699
South Fork Newaukum River	Lewis	WA	46.57511	−122.835
Siuslaw River at Brickerville	Lane	OR	44.06064	−123.885
Skamania at Skelton Road	Skamania	WA	45.62187	−122.05
Skookumchuck River	Thurston	WA	46.79536	−122.761
Skookumchuck River, Bucoda Volunteer Park, Bucoda	Thurston	WA	46.79609	−122.867
South Yamhill River at McMinnville City Park	Yamhill	OR	45.20714	−123.18
Spider Lake, off of FR 23	Mason	WA	47.40684	−123.437
Thomas Creek at Chapin Park in Scio	Linn	OR	44.70445	−122.847
Thompson City Park, just upstream of Interstate 84, The Dalles	Wasco	OR	45.60456	−121.189
Tualatin Hills Park, Cedar Mills	Washington	OR	45.52142	−122.838
Tweedle Rd.	Clatsop	OR	45.89631	−123.554
US Hwy 30, Scapoose	Columbia	OR	45.77081	−122.879
Willamette River at Salem Wallace Marine Park	Polk	OR	44.946	−123.044
Youngs River Falls, just upstream of Youngs River Rd.	Clatsop	OR	46.06745	−123.789
Clade 3: *nigrina* group				
Clear Creek at French Gulch	Shasta	CA	40.70483	−122.637
Cow Creek Rd. at FS road 32-7-19	Douglas	OR	42.7724	−123.574
Edson Campground	Curry	OR	42.81515	−124.411
Elk River Road, just upstream of US Hwy 101	Curry	OR	42.78614	−124.481
~0.35 mi south of Hwy 138 (south of junction of NF4714 and NF4720)	Douglas	OR	43.30182	−122.682
Jack Creek	Curry	OR	42.06198	−124.219
Jenny Creek at Pinehurst	Jackson	OR	42.118	−122.366
Montgomery Creek at Montgomery Woods State Reserve	Mendocino	CA	39.23469	−123.396
Neil Creek	Jackson	OR	42.17654	−122.65
North Fork Galice Creek Rd.	Josephine	OR	42.55312	−123.632
North Umpqua River at Amacher Park	Douglas	OR	43.28162	−123.356
near jct of Birdseye Creek Rd. and Birdseye West Rd (37–4-4)	Jackson	OR	42.38448	−123.175
Redwood Creek at Chezem Road	Humboldt	CA	40.91292	−123.814
Redwood Creek at Orick	Humboldt	CA	41.2888	−124.057
Rogue River at Agness	Curry	OR	42.55783	−124.06
Rogue River at Carpenter Island Park	Josephine	OR	42.55954	−123.598
Rogue River at Huntley Park	Curry	OR	42.48021	−124.33
Rogue River at Rogue River City	Jackson	OR	42.43128	−123.171
Shasta River below Hwy 263 Bridge	Siskiyou	CA	41.80701	−122.594
St. Hwy 234, Dodge Bridge County Park	Jackson	OR	42.52609	−122.843
St. Hwy 199, 1.4 mile southwest of Wonder	Josephine	OR	42.34796	−123.559
along St. Hwy 199, south side of Gasquette, former site of Adams Station	Del Norte	CA	41.84275	−123.995
along St. Hwy 42, just southwest of Winston	Douglas	OR	43.11768	−123.427

(*continued on next page*)

Table 1 (*continued*)

Site	County	State	Latitude	Longitude
St. Hwy 425, just south of Riverton	Coos	OR	43.1162	−124.29
St. Hwy 62 just upstream of Casey State Recreation Area. Takelma Drive, near McGregor Park.	Jackson	OR	42.65887	−122.694
Umpqua River at Sawyers Rapids	Douglas	OR	43.68216	−123.668
US Hwy 101, Humbug Mountain State Park	Curry	OR	42.68371	−124.422
upstream of NF2823 bridge (d/s of log jam)	Douglas	OR	43.10709	−122.585
Yoncalla Creek at Boswell Road	Douglas	OR	43.6408	−123.298
Clade 4: *Juga newberryi*				
Deschutes River at Rainbow Landing	Jefferson	OR	44.75784	−121.227
Oak Springs	Wasco	OR	45.22164	−121.083
Opal Springs	Jefferson	OR	44.49055	−121.298
US Hwy 26	Jefferson	OR	44.75838	−121.227
Clade 5: *Juga* sp. 2				
Confluence of Muddy Hollow and China Hollow	Sherman	OR	45.64714	−120.809
Dog Creek	Skamania	WA	45.71059	−121.671
Fifteenmile Creek	Wasco	OR	45.43089	−121.225
Gate Creek at the 4820 Road	Wasco	OR	45.21917	−121.431
Harpham Flat	Wasco	OR	45.13754	−121.122
Major Creek	Klickitat	WA	45.71542	−121.351
Muddy Hollow 2	Sherman	OR	45.63111	−120.793
Oak Springs	Wasco	OR	45.22164	−121.083
Post Canyon Dr.	Hood River	OR	45.69697	−121.575
South Fork Gate Creek at the 4830 Road	Wasco	OR	45.19596	−121.413
South Junction	Wasco	OR	44.86017	−121.06
Souva Creek off the 4,830,120 Road	Wasco	OR	45.2081	−121.448
St. Hwy 281, Hood River	Hood River	OR	45.69708	−121.523
US Hwy 26	Jefferson	OR	44.75838	−121.227
West Fork Neal Creek at 1700 Road	Hood River	OR	45.54211	−121.522
Clade 6: Blue Mountains Juga				
Phipps Meadow along US Hwy 26	Grant	OR	44.58158	−118.442
Clade 7: *Juga* sp. 3				
Shoat Springs, Copco Rd.	Jackson	OR	42.046	−122.336
Headwaters of Sacramento River at Mount Shasta City Park	Siskiyou	CA	41.32871	−122.327
Keene Creek at Lincoln	Jackson	OR	42.1047	−122.413
Unnamed Creek at Skookum Creek Road	Jackson	OR	42.02901	−122.338
Clade 8: *Juga occata*				
Baum Lake	Shasta	CA	40.93447	−121.549
Fall River at Spinner Fall Lodge	Shasta	CA	41.09723	−121.549
Pitt River at US Hwy 299	Shasta	CA	40.98066	−121.547
Spring Creek at Spring Creek Road	Shasta	CA	41.10176	−121.519
Clade 9: *acutifilosa* group				
Bitner Ranch	Washoe	NV	41.73641	−119.469
Adjacent to Coleman Lake (dry)	Lake	OR	42.06981	−119.84
Divine Spring run near Home Camp	Washoe	NV	41.35402	−119.854
North Little High Rock Canyon Wilderness, Black Rock Desert	Washoe	NV	41.2595	−119.4343
Spring in Murrer's Lower Meadow	Lassen	CA	40.58752	−120.697
Willow Creek at Hayden Hill Road	Lassen	CA	41.02017	−120.853

48 °C for 30 s and extension at 72 °C for 45 s, followed by a final extension at 72 °C for 5 min. A 431–507 bp fragment of the ITS1 gene was amplified using the ITS1 and ITS2 primers from White et al. (1990). Reaction volumes for ITS1 were 1x Biolase reaction buffer, 500 μM dNTPs, 1.5 mM MgCl2, 0.3 μM each primer, 0.25 μg/μL BSA, 1 unit Biolase DNA polymerase and amplification regime of initial denaturation at 94 °C for 7 min, followed by 35 cycles of denaturation at 94 °C for 40 s, annealing at 59 °C for 40 s and extension at 72 °C for 60 s, followed by a final extension at 72 °C for 10 min. PCR products were purified using the Exo-SAP-IT protocol (GE healthcare, Piscataway, NJ). BigDye 3.1 (ABI, Foster City, CA) sequencing reactions and sequencing on an ABI 3730XL DNA analyzer capillary array were done following manufacturer's instructions. Chromatograms were visually inspected and corrected as necessary in Geneious Pro 11.1.2 (Biomatters). COI sequences were translated into amino acids to check for stop codons and frameshift mutations.

Not every individual was successfully amplified and sequenced for both mitochondrial genes. The final COI dataset comprises 571 sequences, and the 16S dataset comprises 566 individuals; 560 individuals were successfully sequenced for both. A subset of 272 individuals was sequenced for ITS1. 267 individuals were sequenced for all three markers (Tables 2 and S1). Two outgroups were selected from the Pleuroceridae, *Elimia comma* and *Elimia carinifera* (see Table S1). Individual genes were aligned with MUSCLE (Edgar, 2004) using default parameters as implemented in Geneious Pro. COI alignments were trimmed to 658 base pairs representing the standard invertebrate

Table 2
Datasets analyzed herein. "Gblocks", nonconserved regions of the alignment removed with Gblocks (Castresana, 2000). "Square", data matrix of newly obtained sequences, comprising all individuals for which ITS1 and at least one mitochondrial gene were successfully sequenced. *alignment gaps are considered missing data.

Dataset	N	Genes	Alignment length	% missing data*	Analyses
COI	573	COI	658	0	ABGD
COI_GenBank	641	COI	658	$1.4x10^{-4}$	ABGD
mtDNA	579	COI, 16S	1166	1.7	
ITS	274	ITS1	676	31.8	
ITS_Gblocks	274	ITS1	534	17.1	
COI_16S_ITS	579	COI, 16S, ITS1	1842	25.9	
COI_16S_ITS_Gblocks	579	COI, 16S, ITS1	1700	20.3	bGMYC, bPTP
COI_16S_ITS_Gblocks_square	274	COI, 16S, ITS1	1700	6.2	bGMYC, bPTP, BP&P, StarBEAST2

barcoding region (Folmer et al., 1994). Nonconserved regions of the ITS1 alignment were removed with Gblocks version 0.91b (Castresana, 2000) using the Gblocks web server (http://molevol.cmima.csic.es/castresana/Gblocks_server.html).

We also compiled a COI dataset for 641 individuals that included all newly generated COI sequences and sequences downloaded from GenBank (Tables 2 and S2). We did not use every *Juga* COI sequence in GenBank as some had gaps and ambiguous base calls that we interpreted as low-quality sequences. Few 16S and no ITS1 sequences for *Juga* have been deposited in GenBank, so we included only COI sequences in our GenBank-supplemented dataset to limit the quantity of missing data.

In total, we compiled eight datasets that varied with regards to taxon sampling, the inclusion or exclusion of nonconserved regions, and the quantity of missing data: (1) two COI datasets, one comprising only newly generated sequences and one augmented with sequences downloaded from GenBank; (2) a combined mitochondrial gene dataset (COI, 16S) with only newly generated sequences; (3) two ITS1 datasets, that differed in the inclusion or exclusion of nonconserved regions; (4) and three concatenated mitochondrial and nuclear gene datasets (COI, 16S, ITS1), including and excluding nonconserved regions, and that differed in the quantity of missing data. See Table 2 for details.

2.3. Primary species hypotheses

We formulated candidate or primary species hypotheses (PSHs) (*sensu* Puillandre et al., 2012a,b) using Automatic Barcode Gap Delimitation (ABGD) as implemented on the ABGD web server (http://wwwabi.snv.jussieu.fr/public/abgd/; Puillandre et al., 2012a). As COI is the standard barcoding gene for invertebrates, we only used the COI gene with ABGD. Analyses were performed omitting the *Elimia* outgroups. Following Barley and Thomson (2016), the best-fit substitution model for each COI dataset was determined with jmodeltest 2.1.10 (Guindon and Gascuel, 2003; Darriba et al., 2012), and model-corrected distances were calculated in PAUP*. ABGD parameters were as follows: Pmin = 0.01, Pmax = 0.08, steps = 20, relative gap width (X) = 0.5, and Nb bins = 40. Pmin and Pmax values were selected based on preliminary analyses such that the values spanned the barcode gap, but did not include values far outside the gap as to minimize the number of steps needed for analyses. The value of X relates to the sensitivity of the method to gap width. Given the overlap in the distribution of pairwise distances within and between groups (see Results, below) we reduced this value from the default 1.5. As ABGD is often run using K80 model-corrected distances (transition to transversion ratio of 2.0; see Barley and Thomson, 2016), we also generated species delimitation hypotheses using K80 distances with all other values as above.

2.4. Phylogenetic inference

All eight datasets were analyzed using Bayesian inference. The best-fit partitioning scheme and substitution models were inferred with PartitionFinder 2 (Lanfear et al., 2016), using the greedy search algorithm and Bayesian information criterion, which favored the following scheme: COI: TrNef + Γ, HKY + I, GTR + Γ for the first, second and third codon positions, respectively; 16S: HKY + Γ; ITS1: TrNef + Γ.

Trees were inferred in MrBayes 3.2.6 (Ronquist et al., 2012) for all eight datasets using the best-fit partitioning scheme and substitution models indicated by PartitionFinder, except model averaging of nucleotide exchangeabilities was employed with reversible jump Markov chain Monte Carlo ("nst = mixed" in MrBayes). Inference in MrBayes used four independent Metropolis coupled Markov chain Monte Carlo runs of 20,000,000 generations, sampling every 500 generations. Convergence was assessed with MrBayes and Tracer 1.4 Rambaut and Drummond (2007). Effective sample size of each parameter was greater than 200 and the potential scale reduction factor of each parameter was 1.00, suggesting that independent runs had converged. A burn-in of 25% of the posterior distribution was discarded and a majority rule consensus tree was constructed with MrBayes.

To assess if overparameterization was a potential source of error in our analyses, we also performed a nonpartitioned Bayesian analysis of the COI dataset for 573 individuals, and of the concatenated dataset for 579 individuals with nonconserved regions removed. Trees were inferred in MrBayes with identical parameters as in the other analyses, except no partitions were defined and the substitution model was set to nst = mixed.

The concatenated dataset for 274 individuals with nonconserved regions removed was also analyzed using maximum likelihood estimation in IQ-TREE (Nguyen et al., 2015). Best-fit partitions and models, as indicated by PartitionFinder 2, were used (COI: TrNef + Γ, HKY + I, GTR + Γ for the first, second and third codon positions, respectively; 16S: HKY + Γ; ITS1: TrNef + Γ). Nodal support was assessed using 1000 ultrafast bootstrap replicates (Hoang et al., 2017). Nodes with posterior probabilities (PP) lower than 0.90 and bootstrap values (BS) lower than 70% were considered unsupported.

2.5. Species delimitation

We further evaluated the primary species hypotheses produced by ABGD to formulate secondary species hypotheses (SSHs). SSHs were based primarily on the criterion of reciprocal monophyly in analyses of the expanded mitochondrial and nuclear gene dataset, and informed by a combination of geographic distribution and morphological data in an integrative framework (Puillandre et al., 2012b). These hypotheses were compared to the current classification (Johnson et al., 2013), to the interpretation of Campbell et al. (2016), and to hypotheses generated by a variety of tree-based algorithmic species delimitation methods in routine use, including bGMYC, bPTP and BP&P. Two of these methods (bGMYC and BP&P) use the multispecies coalescent model, while the Bayesian implementation of the Poisson tree processes model is based on coalescent theory. Given the results of recent simulation studies that indicate programs implementing the multispecies coalescent model delimit genetic structure rather than species (Sukumaran and Knowles, 2017; Leaché et al., 2018), we wished to explore the performance of these methods in analyses of freshwater snail species

with low vagility and poor dispersal capacity (e.g., Kappes, 2012).

Bayesian generalized mixed Yule-Coalescent (bGMYC; Reid and Carstens, 2012) species delimitation, a Bayesian extension of GMYC species delimitation (Fujisawa and Barraclough, 2013), was run in the R 3.5.3 package (R Development Core Team, 2019) bGMYC (Reid and Carstens, 2012). bGMYC requires ultrametric trees so we used mitochondrial gene trees from StarBEAST2 posterior distributions; only 25 trees were used to decrease computational times associated with running bGMYC on over 500 individuals. We used the R package APE (Paradis et al., 2004) to collapse tree tips represented by individuals with identical mitochondrial haplotypes into a single tip. This was done so input trees better conformed to model assumptions of bGMYC (Reid and Carstens, 2012; Fujisawa and Barraclough, 2013). bGMYC was run using 500,000 MCMC generations, burn-in of 350,000, sampling every 5 generations, and starting values of 25 for the threshold parameter and 1 for the coalescent rate change parameters. Default values were used for all other parameters.

Bayesian PTP analyses (Zhang et al., 2013) were performed using 20 trees from the posterior distribution of MrBayes analyses on both the COI_16S_ITS_Gblocks and COI_16S_ITS_Gblocks_square datasets; only 20 trees were used for computational efficiency. Trees inferred with MrBayes were used because PTP does not require ultrametric trees, which has been argued as a strength of this method (Zhang et al., 2013). Identical tips were collapsed as done with bGMYC. Non-*Juga* outgroups were trimmed from analyses following the recommendation of Zhang et al. (2013). We ran Bayesian PTP for 300,000 generations, sampling every 100 generations following a 25% burn-in. Default values were used for all other parameters.

In addition to the species delimitation methods described above, we used the multi-locus method BP&P version 3.3, which uses the multispecies coalescent model and a reversible jump MCMC to test species boundaries. Unlike other methods used here, BP&P requires *a priori* allocation of sequenced individuals into putative species. We performed three BP&P species delimitation analyses using the same *a priori* assignments of individuals to putative species as used in the three StarBEAST analyses (see below). This included the maximum inferred split produced by the bGMYC analyses (34 putative species) to establish an upper bound in the number of species that could be delimited by BP&P (see Fig. 2, and Results, below). All BP&P analyses were done with the following parameters: maximum clade credibility trees inferred with StarBEAST2 as guide trees (see below); the species delimitation reversible-jump MCMC algorithm 1 with $\alpha = 2$ and $m = 1$; a gamma prior with values $\alpha = 2$ and $\beta = 1000$ for the population size parameter; a gamma prior with values $\alpha = 2$ and $\beta = 2000$ for divergence time at the root of the species tree; a Dirichlet process prior for all other divergence time parameters. For each analysis, three independent BP&P runs were done for 100,000 MCMC generations, sampling every two generations, with a burn-in of 8000 generations.

2.6. Species tree estimation

As concatenation methods can result in well-supported but incorrect trees (Degnan and Rosenberg, 2006; Kubatko, 2007), we generated a species tree hypothesis with StarBEAST2 (Heled and Drummond, 2010), a Bayesian species tree reconstruction method that uses the multispecies coalescent model (Rannala and Yang, 2003) to resolve gene tree conflict. StarBEAST2 analyses were run in BEAST 2.4.8 (Bouckaert et al., 2014) with dataset Juga_COI_16S_ITS_Gblocks_square; analyses with dataset Juga_COI_16S_ITS_Gblocks, which included more individuals, and also considerably more missing data, failed to converge and are not reported. StarBEAST2 requires *a priori* assignment of individuals to different species, which was first done using our secondary species delimitation hypotheses (SSHs), corresponding to the nine main clades resulting from phylogenetic analyses of the concatenated dataset (Fig. 2). Given the results of ABGD with model-corrected distances (see Fig. 2, and Results, below), a second analysis that further subdivided Clade 1 into two putative species was also run. A third analysis that allocated individuals to species indicated by bGMYC was also done. See Tables S3–S5 for the *a priori* allocations of each sequenced individual into putative species. For all analyses, sites and clocks were unlinked for each partition, and we used best-fit partitions and substitutions models indicated by PartitionFinder. As mitochondrial genes are inherited as a single locus, the tree parameter was linked for both mitochondrial genes. Genes were partitioned as with MrBayes analyses. Each partition was assigned a GTR substitution model and rate heterogeneity was modeled with a four category, discrete gamma distribution. We used a random local clock model (Drummond and Suchard, 2010) and a Yule tree prior for all analyses. We performed two independent MCMC runs for 100,000,000 generations, sampling every 1000 generations. Convergence was assessed with Tracer 1.4 Rambaut and Drummond (2007) and assumed to have occurred if effective sample sizes were greater than 200 and trace plots appeared stationary for each parameter.

2.7. Genetic diversity and population structure

Genetic diversity of each species resulting from our secondary species delimitation (=SSH; Table S1, Fig. 2) and the distribution of haplotypes among sites was visualized with haplotype networks for the COI dataset. Median-joining networks (Bandelt et al., 1999) were inferred with epsilon set to 0 using Popart 1.7 (http://popart.otago.ac.nz). Individuals were grouped and color-coded by sampling locality (Table 1). Networks were visualized with Popart and further edited in Adobe Illustrator.

To further quantify genetic structuring within species, we performed an analysis of molecular variance (AMOVA; Excoffier et al., 1992) to calculate the amount of variance among samples that can be attributed to differences among populations (φ_{ST}). The Blue Mountains Juga (Clade 6) was excluded from the analyses as it was found at only a single site. AMOVA was performed with the R 3.5.3 package (R Development Core Team, 2019) Poppr (Kamvar et al., 2014), using the ade4 method (Dray and Dufour, 2007); significance was tested with the Poppr function "randtest" using 1000 permutations. Given the low-dispersal capacity of freshwater snails (e.g., Kappes, 2012), we considered each sampling site as a distinct population. In one instance for *Juga newberryi* (Clade 4), we ran a second analysis with two samples merged (Deschutes River at Rainbow Landing, US Hwy 26) (see Table 3). These samples resulted from collections in subsequent years at two sites that were ~50 m apart in an attempt to recollect *Juga* sp. 2 along this section of the Deschutes River.

2.8. Data sharing

All datasets and the resulting trees in newick or nexus format, and all images of the sequenced vouchers, have been uploaded to a FigShare repository: https://doi.org/10.6084/m9.figshare.6843842. All newly generated sequences have been deposited in GenBank (Table S1).

3. Results

3.1. Primary species hypotheses

ABGD analyses with best-fit model-corrected distances (TrN + Γ) delimited either eight or nine primary species hypotheses, with or without GenBank sequences, respectively (Fig. 2). K80 model-corrected distances (transition to transversion ratio of 2.0), which are standard for most ABGD analyses (Barley and Thomson, 2016), produced 24 primary species hypotheses. The frequency histograms of pairwise distances (Fig. 3) indicate that there is no distinct "barcode gap", but that the distribution of model-corrected pairwise distances within and between clusters overlaps in the range of ~4–8% divergence, depending on the dataset and model correction used.

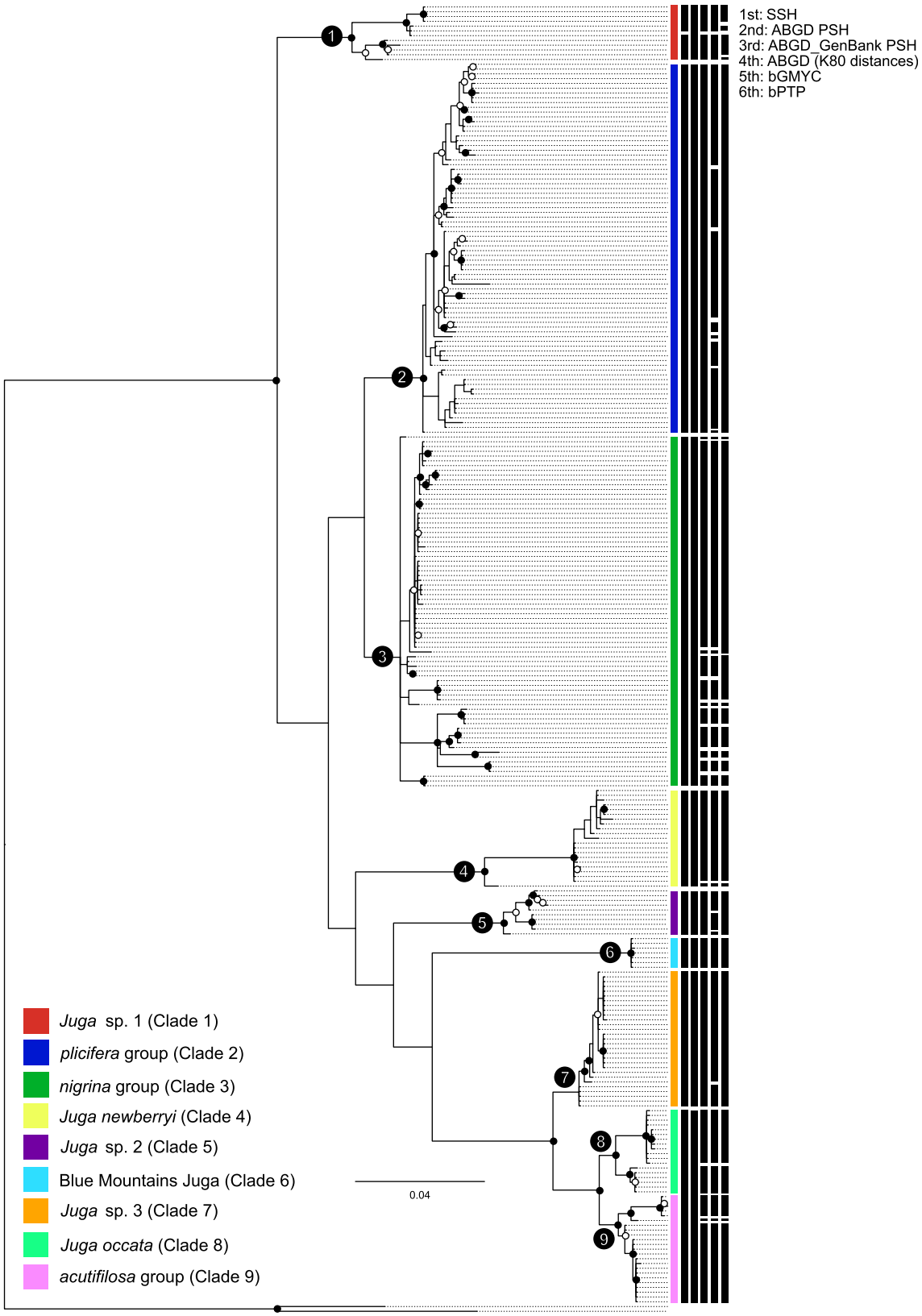

Fig. 2. Bayesian phylogram for 274 individuals based on a concatenated COI, 16S and ITS1 dataset, with nonconserved regions of the ITS1 dataset removed. Primary species hypotheses (=PSHs; from ABGD analyses using best-fit model-corrected distances), are compared to the results from ABGD using K80 model-corrected distances, and from two species delimitation methods (bGMYC, bPTP). Clade-level, secondary species hypotheses (SSHs) are indicated by numbers 1–9 and by colored bars at right; colors correspond to sampling localities in Fig. 1. Tree is rooted with *Elimia comma* and *Elimia carinifera*. Dots correspond to Bayesian posterior probabilities; white PP = 0.90–0.98, black PP = 0.99–1. Scale bar indicates number of nucleotide substitutions per site.

Table 3
Results of analysis of molecular variance (AMOVA) among populations of the nine main clades (=SSHs) using the COI dataset. The Blue Mountains Juga was not included as it was sampled from only a single site. "*" Indicates the results when two samples for *J. newberryi* were merged. See text for details.

Species	Variance	% total	φ_{ST}	*p*-value
Clade 1: *Juga* sp. 1	0.3872	62.2	0.6222	< 0.001
Clade 2: *plicifera* group	0.3088	58.0	0.5800	< 0.001
Clade 3: *nigrina* group	0.2314	56.7	0.5666	< 0.001
Clade 4: *Juga newberryi*	0.0323	20.5	0.2054	0.004
	0.0659*	27.4*	0.2737*	< 0.001*
Clade 5: *Juga* sp. 2	0.3650	66.6	0.6663	< 0.001
Clade 6: Blue Mountains Juga	N/A	N/A	N/A	N/A
Clade 7: *Juga* sp. 3	0.2133	55.7	0.5570	< 0.001
Clade 8: *Juga occata*	0.1681	61.3	0.6130	< 0.001
Clade 9: *acutifilosa* group	0.3373	79.9	0.7992	< 0.001

3.2. Phylogenetic inference

Bayesian analyses of both mitochondrial genes, separately or in combination, and of the concatenated mitochondrial and nuclear gene datasets (Table 2), produced nine main clades within *Juga*, usually with robust support (Figs. 2, S1–S3, S6 and S7). All nine clades received PP = 1 on the COI gene trees (Figs. S1, S2), and all except two received PP = 1 on the mtDNA tree (Fig. S3), with one receiving high support (Clade 8; PP = 0.99) and the other unsupported (Clade 5; PP = 0.75). The ITS1 dataset was more conserved than the mitochondrial gene datasets. Analysis of the ITS1 dataset, with and without nonconserved regions removed, supported monophyly of only four of the nine main clades produced in analyses of the mitochondrial and concatenated datasets (Clades 1, 4, 5, 6), but all four were supported with PP = 1; the other five clades collapsed but were not contradicted (Figs. S4, S5). In the analyses of the concatenated datasets (Figs. 2, S6 and S7), all nine clades were robustly supported (PP = 1) except one (Clade 3) with variable support (PP = 0.67, 0.87, 1) which was sensitive to missing data and the inclusion or exclusion of nonconserved regions of the ITS1 alignment.

The nonpartitioned Bayesian analysis of the COI dataset for 573 individuals and of the concatenated dataset for 579 individuals with nonconserved regions removed (results not shown), were congruent with the partitioned analyses in supporting the nine main clades and the relationships between them. All main clades were robustly supported (PP = 1) in both nonpartitioned analyses, except Clades 5 and 7, which received PP = 0.99 in the nonpartitioned analysis of the COI dataset only. Support for internal nodes did not respond in a predictable way, and were lower in the nonpartitioned analysis of the COI dataset, but much higher in the nonpartitioned analysis of the concatenated dataset (PP > 0.98).

Maximum likelihood analysis of the concatenated dataset for 579 individuals with nonconserved regions removed also produced the nine main clades with high support (BS ≥ 94) and was congruent with the Bayesian analysis in the relationships between clades. Bootstrap values for internal nodes were moderate to high (BS = 78–100) (Fig. S8).

3.3. Secondary species hypotheses

Our secondary species hypotheses (SSHs; Fig. 2) were based primarily on the criterion of reciprocal monophyly in analyses of the expanded mitochondrial (COI, 16S) and nuclear gene (ITS1) dataset, and were informed by geographic distribution and morphological data in an integrative framework. The resulting SSHs are equivalent to the nine main clades produced in the phylogenetic analyses as delimited above. Six of the nine main clades correspond to primary species hypotheses produced by the two ABGD analyses using best-fit model-corrected distances. Our SSHs differ from the primary species hypotheses in recognizing Clades 8 and 9, which were united in a single primary species hypothesis in both of the ABGD analyses. Despite displaying variation in spiral shell ornament that intergrades between them, the two clades were reciprocally monophyletic in all phylogenetic analyses except of the ITS1 dataset (Figs. S4 and S5), and are allopatric; Clade 8 occurs in the central Pit River system, while Clade 9 occurs in the upper Pit River and in nearby parts of Great Basin.

In the case of Clade 1, the primary species hypotheses produced by ABGD were sensitive to sampling, as evidenced by the splitting of this clade in the analysis with fewer individuals. However, members of the two subclades are morphologically indistinguishable, including where they occur in sympatry (Hat Creek; Fig. 4G–L). In all other cases where two putative species co-occur [Major Creek (Clades 2, 5), and Oak Springs and US Hwy 26 (Clades 4, 5)], they are readily distinguished based on shell morphology. Consequently, we conclude that splitting of this clade by ABGD is the result of sampling error, and we recognize a single species.

3.4. Phylogenetic inference and conflict among datasets

Very different hypotheses of relationship between the nine main clades were inferred in the mitochondrial gene trees (Figs. S1–S3) compared to those in the nuclear gene trees (Figs. S4, S5). For example, mitochondrial gene trees supported Clade 6 as sister to all other *Juga*, while the nuclear gene trees supported Clade 1 in this position. With one exception, interior nodes were strongly supported with PP ≥ 0.96 in the nuclear gene trees, but were often unsupported (PP = 0.75–0.89) in the mitochondrial gene trees. In analyses of the concatenated mitochondrial and nuclear gene datasets (Figs. 2, S6 and S7), the overall

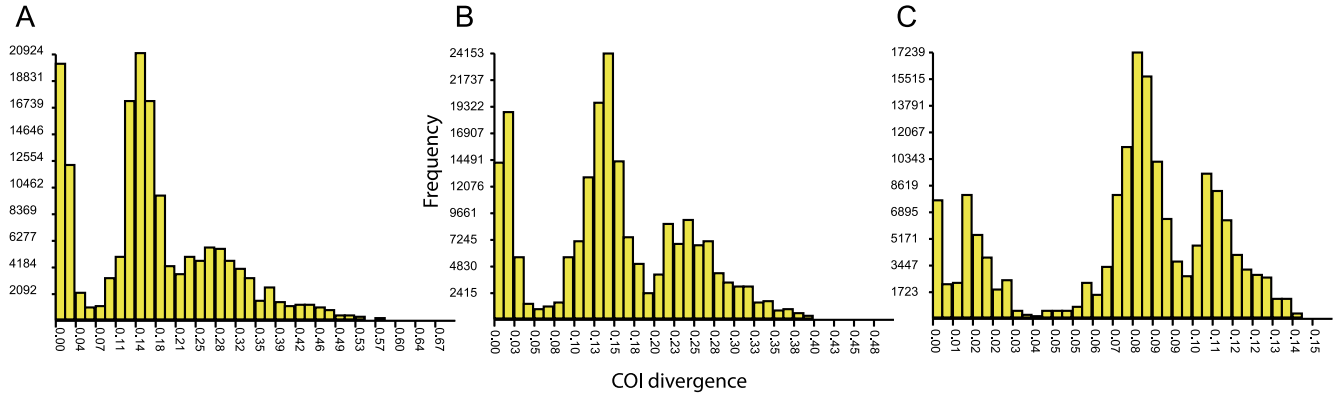

Fig. 3. Frequency histograms of model-corrected pairwise distances produced by ABGD analysis of three COI datasets. A. COI dataset for 573 individuals using best-fit model-corrected distances. B. COI dataset supplemented with GenBank sequences for 641 individuals using best-fit model-corrected distances. C. COI dataset for 573 individuals using K80 model-corrected distances.

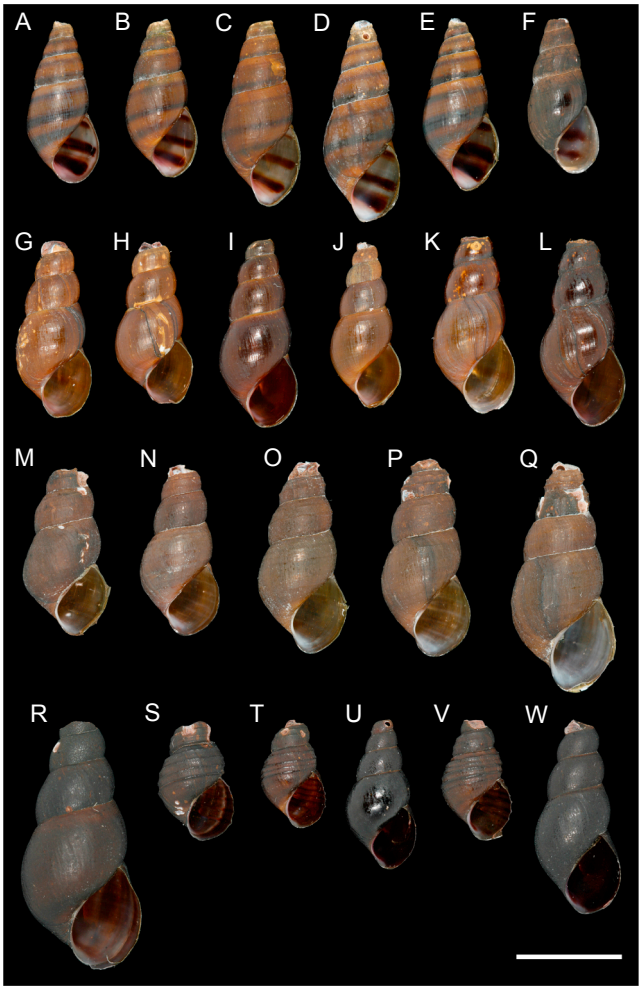

Fig. 4. Shell morphology and genetic disparity. A–E. *Juga* sp. 2, Dog Creek, WA. Specimen E differs from the others, all of which have the same COI haplotype, by 6% model-corrected pairwise divergence in COI. **A.** USNM 1413227; **B.** USNM 1413228; **C.** USNM 1413229; **D.** USNM 1413230; **E.** USNM 1413231; **F.** *Juga* sp. 2, Major Creek, WA. USNM 1413185. Identical COI haplotype with specimen E. **G-J.** *Juga* sp. 1, Hat Creek at Bridge Campground, CA. Specimen J differs from the others, which all have the same COI haplotype, by 7.6% model-corrected pairwise divergence in COI. **G.** USNM 1413092; **H.** UNSM 1413094; **I.** USNM 1413096; **J.** USNM 1413097. **K–L.** *Juga* sp. 1, Goose Valley Spring, CA. Both K and L differ from specimen J by 2.2% model-corrected pairwise divergence. **K.** USNM 1413132; **L.** USNM 1413136. **M–P.** *Juga occata*, Fall River at Spinner Fall Lodge, CA. Specimen O differs from the others, all of which have the same COI haplotype, by 3.6% model-corrected pairwise divergence. **M.** USNM 1413080; **N.** USNM 1413081; **O.** USNM 1413082; **P.** USNM 1413083. **Q.** *Juga occata*, Baum Lake, CA. USNM 1413149. Differs from specimen O by 0.81% model-corrected pairwise divergence. **R–W.** *Juga* sp. 2, Unnamed creek at Skookum Creek Rd., OR. Specimens with variable sculpture patterns, and no corresponding differences in COI; specimens R, S, T, V, W, all have the same COI haplotype. Specimen U differs from the others by 0.82% model-corrected pairwise divergence in COI. **R.** USNM 1413276; **S.** USNM 1413277; **T.** USNM 1413278; **U.** USNM 1413279; **V.** USNM 1413280; **W.** USNM 1413281. Scale bar, 1 cm.

structure of the tree strongly mirrored that of the ITS1 gene trees given the support for deeper nodes in the nuclear gene dataset, but with the resolution at shallower levels of the mitochondrial gene trees supporting monophyly of the nine main clades.

Species tree analyses with StarBEAST2 produced different deep relationships depending on how we defined *a priori* species designations. In the analysis with Clade 1 considered a single species (Fig. 5A), Clade 6 was sister to all other *Juga*, in agreement with the mitochondrial gene trees. However, support for this, and other deep relationships, were low (PP = 0.37–0.80). Species tree inference with Clade 1 subdivided *a priori* into two putative species resulted in Clades 6–9 united as sister to the rest of *Juga*. However, as in the first analysis, most relationships were unsupported (Fig. 5B). Analysis with individuals assigned to putative species based on bGMYC delimitation (34 putative species; see below) united Clades 5 + 6 as sister to all other *Juga*, but again support for most relationships was low (Fig. 5C).

3.5. Genetic diversity and population structure

Analyses of the COI dataset revealed an often highly complex and heterogeneous landscape of genetic diversity within populations and structuring between them. Excluding three COI sequences that were not full length, the COI dataset comprised 125 unique haplotypes, 94 (75%) of which were confined to single sites; almost 40% of the sites sampled produced only a single COI haplotype.

The haplotype networks for each of the nine main clades (=SSHs; Figs. 6, 7) generally displayed a pattern of many haplotypes separated by only a few mutations, but haplotypes separated from each other by greater than 10 mutations were common. Haplotypes often clustered by sampling location, but some gene flow was evident as haplotypes were often shared among sites (e.g., Fig. 7C). At sites where multiple haplotypes were found, they exceeded 1% uncorrected pairwise divergence at over 20% of the sites sampled. This is at the high end of the level assumed indicative of species-level distinction by Campbell et al. (2016).

At several sites, co-occurring haplotypes could differ by up to 3.6% model-corrected pairwise divergence (e.g., Clade 8, Fall River at Spinner Fall Lodge; Clade 2, Benson State Recreation Area) and in extreme cases, by as much as 6.0% (Clade 5, Dog Creek) or 7.6% (Clade 1, Hat Creek) divergence among members of the same clade. These high levels of sequence divergence occurred among individuals that were essentially identical in shell morphology (Fig. 4). In these cases, genetically divergent individuals were found to be closely related to neighboring populations from the same clade. For example, in Clade 1 (Fig. 6A), the most genetically divergent individual from Hat Creek (Fig. 4J) was more closely related to individuals from Goose Valley Spring (e.g., Fig. 4K–L), from which it differed by 0.16%–2.2% model-corrected pairwise divergence. In Clade 5 (Fig. 7A), the most genetically divergent individual sequenced from Dog Creek (Fig. 4E) had the same COI haplotype as individuals from nearby Major Creek (e.g., Fig. 4F). In Clade 8 (Fig. 7D), the most genetically divergent individual from the Fall River at Spinner Fall Lodge (Fig. 4O) was more closely related to individuals from Baum Lake (e.g., Fig. 4Q) from which it differed by 0.16–1.0% divergence. Conversely, marked variation in shell ornamentation could be found among individuals with identical COI haplotypes collected at a single site. For example, individuals collected syntopically from a spring-fed creek in southern Oregon displayed smooth to spirally sculptured forms among individuals with the same COI haplotype (Fig. 4R–W); individuals from this population had two different ITS1 alleles, but ITS1 allele did not correspond to presence or absence of shell sculpture.

The genetic structure evident in the haplotype networks was also apparent on the inferred phylogenetic trees. For example, within Clade 2, there are 22 internal nodes with posterior probabilities of 0.90 or higher and 30 different COI haplotypes (Fig. 2). Given often substantial differences in distribution of haplotypes among samples, within-clade phylogenetic structure often corresponded to a site or to subsets of individuals at a site (Figs. 6, 7). Maximum model-corrected genetic divergence in COI within the nine main clades was highly heterogeneous, and ranged from 1.0% (Clade 7) at the low end, to 4.0% (Clade 8), 4.2% (Clade 9), 4.9% (Clade 2), 5.5% (Clade 3), 6.5% (Clade 4), 6.9% (Clade 5), and 8.0% (Clade 1). The one clade lacking any phylogenetic structure was represented by a single population collected in the Blue Mountains of eastern Oregon (Clade 6; 0.16%). This clade was

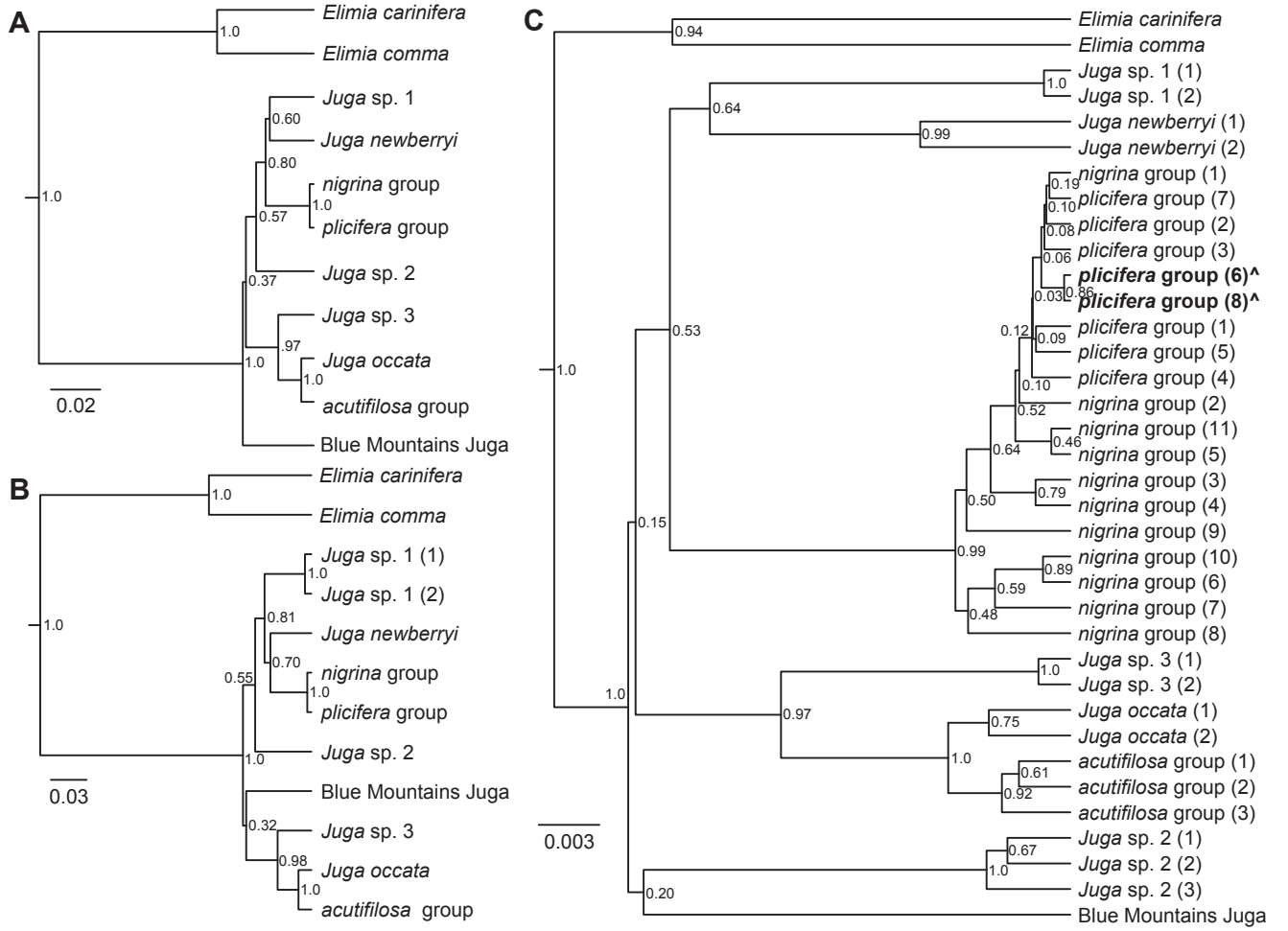

Fig. 5. Maximum clade credibility species trees produced by StarBEAST2 for the concatenated COI, 16S and ITS1 dataset for 274 individuals, with nonconserved regions of the ITS1 dataset removed. **A.** *a priori* species designations matching primary species allocations into nine main clades. **B.** *a priori* species designations matching primary species allocations, and with Clade 1 subdivided into two subclades. **C.** *a priori* species designations matching bGMYC species delimitation. Tips labelled with "^" were combined into one species by BP&P, but all other tips were considered separate species by BP&P. See Tables S3–S5 for *a priori* allocations. Posterior probabilities are shown at the nodes. Scale bar indicates number of nucleotide substitutions per site.

genetically homogeneous with only two COI haplotypes separated by a single nucleotide polymorphism.

Results of analysis of molecular variance (AMOVA) demonstrated that all of the species sampled from more than a single site (excluding only Clade 6) have a marked and highly significant population genetic structure (Table 3). At the low end of the spectrum, variation among populations within Clade 4 (*Juga newberryi*) explained 20.5% of the variance (φ_{ST} = 0.2054). This value increased to 27.4% (φ_{ST} = 0.2737) when the two samples from Deschutes River at Rainbow Landing and US Hwy 26 were merged (see Materials and Methods, above). In seven of the eight species, over 55% ($p < 0.001$) of the observed genetic variance was explained by the variation among populations ($\varphi_{ST} > 0.5570$), and in one case (Clade 9: *acutifilosa* group) this value reached almost 80% (φ_{ST} = 0.7992; $p < 0.001$).

3.6. Tree-based algorithmic species delimitation

Species delimitation with bPTP, bGMYC and BP&P produced highly dissected species hypotheses that frequently disagreed in the number of putative species recognized and how individuals were allocated among them (Figs. 2 and S7). Species delimitation with bPTP produced 24 or 37 putative species with the COI_16S_ITS_Gblocks_square and COI_16-S_ITS_Gblocks datasets, respectively, demonstrating sensitivity of the method to taxon or character sampling, or to both (Figs. 2 and S7). For example, using the smaller dataset (COI_16S_ITS_Gblocks_square) with less missing data, bPTP subdivided Clade 3 into nine species, but subdivided the same clade into 17 putative species using the larger dataset with more missing data. As noted above, ABGD with K80 model-corrected distances also produced 24 putative species, agreeing in number, and often in composition (Clades 2, 4–9), with the estimate produced by bPTP.

Analysis with bGMYC of the COI_16S_ITS_Gblocks_square dataset identified 34 putative species, the most of any method. Three clades that were inferred to be single species by all other methods (Clades 2, 5 and 7) were split into at least two putative species by bGMYC, and in the case of Clade 2, into as many as eight. The allocations produced by bGMYC agreed in composition with those produced by bPTP for Clades 4, 6, 8 and 9, but did not appear as sensitive to sampling and produced similar numbers of putative species (34 vs. 33) for both datasets (Figs. 2 and S7).

Unlike the preceding analyses, BP&P requires *a priori* species designations and can only collapse, not subdivide, *a priori* groupings. We explored a range of *a priori* allocations, including the more conservative scheme matching our secondary delimitation hypotheses (SSHs), to the scheme favored by bGMYC with 34 putative species. With one exception, BP&P split species to as low a level as allowed by the initial input with high support (PP = 1); in the analysis with putative species allocated based on bGMYC, two terminals in the *plicifera* group (Clade 2)

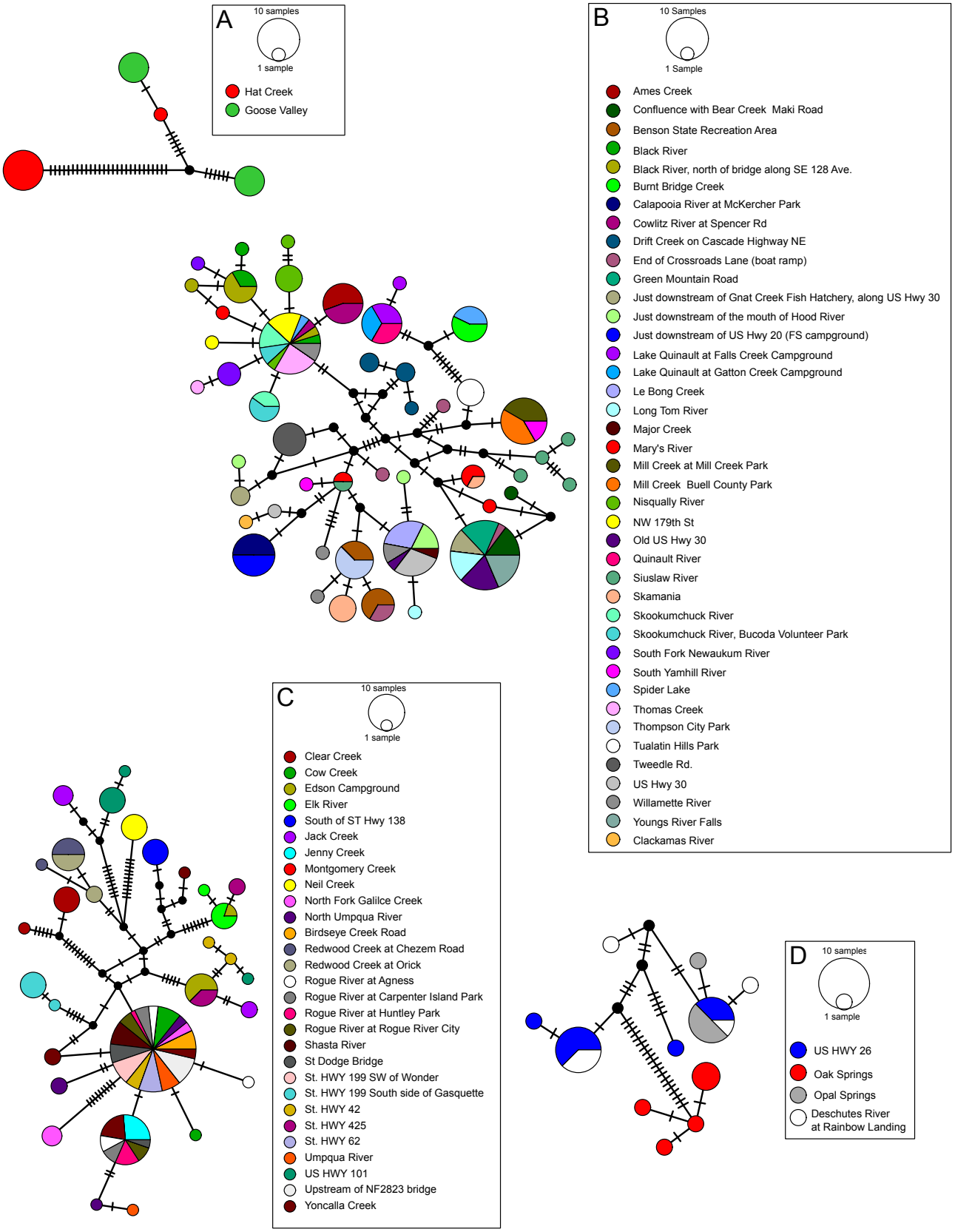

Fig. 6. Median-joining haplotype networks for main clades (=secondary species hypotheses). A. Clade 1 (*Juga* sp. 1). B. Clade 2 (*plicifera* group). C. Clade 3 (*nigrina* group). D. Clade 4 (*Juga newberryi*). Circles are proportional to the number of haplotypes and are color coded by population. Black dots are inferred, but unsampled, haplotypes. Each tick mark represents one mutation.

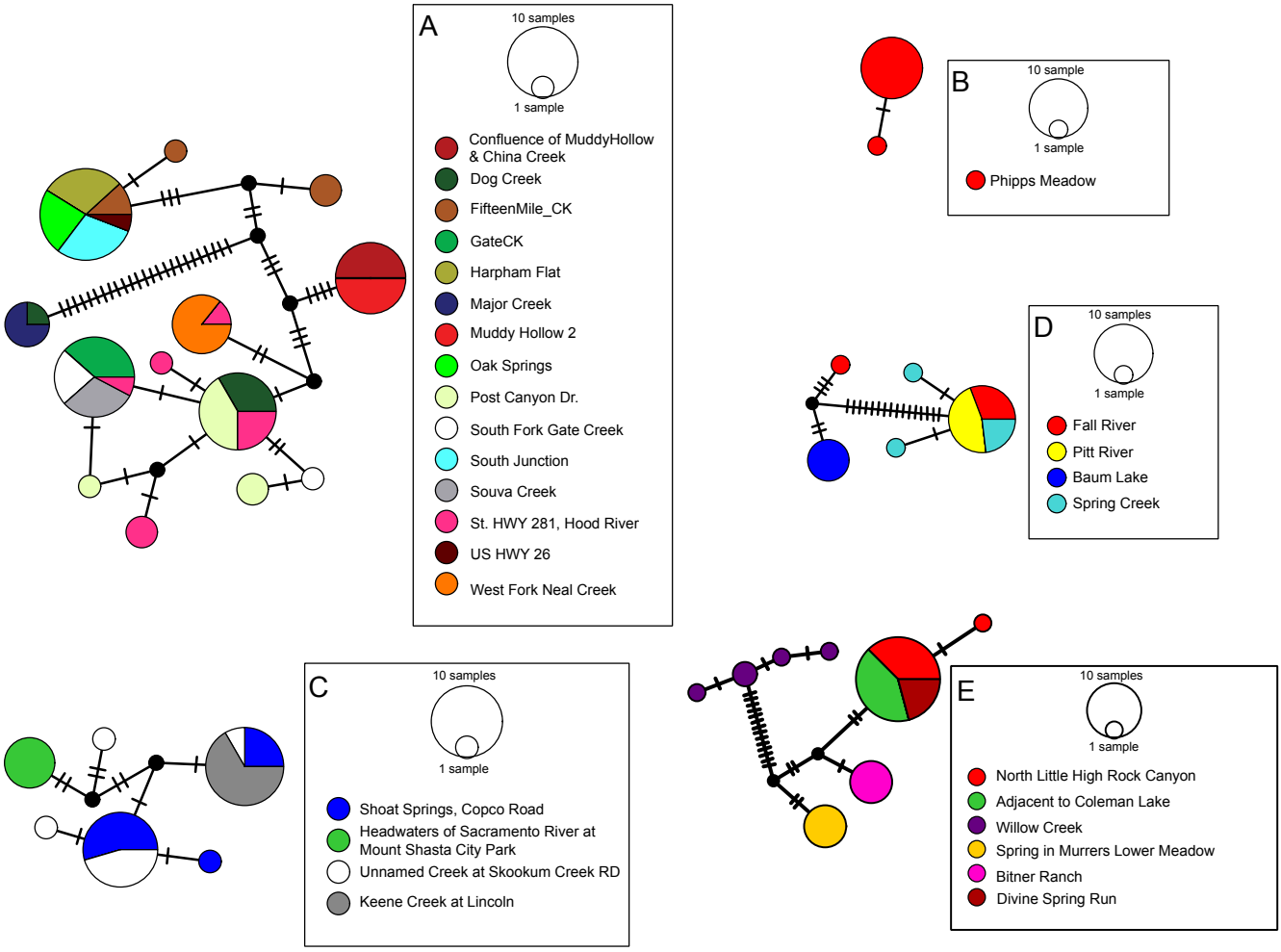

Fig. 7. Median-joining haplotype networks for main clades (=secondary species hypotheses). A. Clade 5 (*Juga* sp. 2). B. Clade 6 (Blue Mountains Juga). C. Clade 7 (*Juga* sp. 3). D. Clade 8 (*Juga occata*). **E.** Clade 9 (*acutifilosa* group). Circles are proportional to the number of haplotypes and are color coded by population. Black dots are inferred, but unsampled, haplotypes. Each tick mark represents one mutation.

were collapsed into a single species (PP = 0.73, no other species delimitation scheme had PP > 0.20).

4. Discussion

4.1. Species diversity of Juga

We generated primary species hypotheses using ABGD with best-fit model-corrected distances, which we refined using other independent lines of evidence, including an expanded multilocus molecular dataset, morphology, and geographic distribution data to produce secondary species hypotheses. In one case where ABGD produced conflicting primary species hypotheses that were sensitive to sampling (Clade 1; Fig. 2), morphologically identical individuals were inferred to belong to the same putative species. This conservative approach will underestimate species diversity in cases of morphologically cryptic species. However, it is not infrequent that so-called "cryptic" species revealed by molecular data, upon closer scrutiny yield shell characters that support the molecular results (e.g., Delicado and Ramos, 2012; Templado et al., 2016; Fedosov et al., 2017). Here we found no such evidence (Fig. 4G–L).

These nine species correspond to the nine clades recovered by Campbell et al. (2016), the difference being in the interpretation of their composition and taxonomic rank. Campbell et al. (2016) interpreted each of the clades as comprised of multiple, geographically restricted OTUs. We dispute this interpretation given that the OTUs recognized by Campbell et al. (2016) were not always reciprocally monophyletic; furthermore, the level of genetic distinction they considered sufficient for species recognition frequently was seen in our study at a single site. With the expanded sampling of this analysis, we interpret their OTUs as lying along a poorly sampled continuum of morphological and genetic variation and are clearly over split. The shift in the distribution of intra- versus interspecific pairwise distances evident when the COI dataset is expanded or analyzed with different models (Fig. 3), demonstrates the importance of model choice and dense geographic and population sampling for accurately inferring the distinction between intra- and interspecific divergence (Puillandre et al., 2012a). Our results also underscore that future studies should not rely on a universal threshold, or barcode gap, for species recognition (e.g., Astrin et al., 2016; Rosel et al., 2017), and that this may be true even among closely related species with highly heterogeneous phylogeographic histories.

While representing significant progress over the previous shell-based classification, of course we consider that our secondary species hypotheses should be subject to further testing through expanded population and molecular sampling. In particular, poorly represented clades with high levels of genetic diversity and phylogenetic structure should be examined more closely (Clades 1 and 4). In these cases, we inferred the pronounced phylogenetic structuring to be an artifact of sampling and have favored a conservative approach. In our experience

with other freshwater cerithioideans, expanded geographic sampling often recovers haplotypes that break up these long branches (unpubl. data), as occurred here for Clade 1 in the analysis of the COI dataset that included GenBank specimens (Figs. S1, S2). By the same token, additional sampling may support the primary species hypothesis that Clades 8 and 9 represent a single species. Despite the fact that the divergence is shallower between these two clades than for the others, we feel that recognition of the two clades as distinct is justified given that they are robustly supported as reciprocally monphyletic in all phylogenetic analyses except of the ITS1 dataset and are allopatric.

4.2. Phylogenetic inference and conflict among datasets

Mitochondrial and nuclear gene trees produced very different hypotheses of relationship between the nine clades. Given the conflict in placement of the root and in sister group relationships of the main clades, support for interior nodes in analyses of the concatenated mitochondrial and nuclear gene datasets is inconsistent and often low. Deep relationships on the concatenated tree largely reflect those inferred with the ITS1 dataset, and appear mostly influenced by signal of the ITS1 gene. This conflict was not resolved in any of the three StarBEAST analyses, which produced topologies with mostly unsupported interior nodes (Fig. 5). Expanding the molecular dataset to additional mitochondrial and nuclear loci will be required to resolve relationships between the nine main clades.

4.3. Genetic diversity and population structure

Freshwater snails like *Juga*, with no larval dispersal stage and low vagility, face dispersal barriers such as habitat isolation in the case of spring-associated species, and network-limited dispersal in wider ranging riverine species. As to be expected in organisms with such low dispersal capacity, analyses revealed a heterogeneous genetic landscape of diverse and highly structured populations with many private alleles. While many populations produced a single, unique COI haplotype, others produced mixtures of haplotypes that frequently exceeded the levels of pairwise divergence considered indicative of species level discrimination by Campbell et al. (2016), and in several instances were separated by as much as 7.6% model-corrected pairwise divergence (Clade 1; Fig. 6A). However, genetic divergence frequently did not covary with variation in shell morphology. In some cases, individuals with highly divergent COI haplotypes collected at a single site were essentially identical in shell morphology (e.g., Clade 1, *Juga* sp. 1; Fig. 4G–J), while conversely, smooth to spirally ornamented individuals could present the same COI haplotype (e.g., Clade 5, *Juga* sp. 2; Fig. 4R–W). ITS1 allele also did not correspond to shell morphology.

We infer this complex genetic landscape to be the result of dynamic phylogeographic processes including population bottlenecks from founder effects, population fragmentation and loss, and secondary contact following sometimes prolonged periods of isolation. Dispersal by birds, headwater capture or tectonic-driven changes in drainage patterns, are likely avenues mediating sporadic gene flow and secondary contact among highly fragmented populations. This island-like population structure has been documented in other freshwater and land snails (see e.g., Thomaz et al., 1996; Becker et al., 2016), and is known to preserve genetic variation and delay lineage sorting as the effective population size remains large.

4.4. Algorithmic methods of species delimitation

One of the main objectives of this study was to assess how our secondary species delimitation hypotheses (SSHs) compared to several tree-based algorithmic species delimitation methods in routine use (i.e. bGMYC, bPTP, BP&P). These methods are widely seen as providing testable and repeatable means of generating hypotheses of species diversity and have been applied in consensus, pluralistic, or validation approaches that seek agreement among methods and have proven useful in several studies on non-adaptive radiations or cryptic species complexes (e.g., Barley et al., 2013; Bagley et al., 2015; Obertegger et al., 2018). However, the accuracy of these methods relies on the correctness of the speciation model. Recently, simulation studies have shown that at least some species delimitation methods, notably those implementing the multispecies coalescent model, perform poorly under conditions that violate assumptions of the model, particularly with regard to gene flow and random mating, or panmixia (e.g., Sukumaran and Knowles, 2017; Leaché et al., 2018). Under these conditions, the units delimited by these methods reflect structure, not species boundaries. Consequently, in taxa with high levels of genetic structuring, species diversity may be significantly overestimated.

Analyses with bGMYC and bPTP returned highly dissected hypotheses of species boundaries within *Juga*. For instance, bPTP inferred 24 putative species and bGMYC inferred 34 in analyses of the COI_16S_ITS_Gblocks_square dataset. BP&P requires *a priori* designation of potential species, but the results split species to as low a level as allowed by the initial input except in one case when two entities in the *plicifera* group (Clade 2) were collapsed into a single species. The resulting schemes frequently disagreed in the composition of the identified units and did not support the interpretation of Campbell et al. (2016) despite approaching similar numbers of OTUs. When in conflict, the recognized groupings usually reflected structure within or between populations, were not always reciprocally monophyletic, and were not corroborated by other lines of evidence including morphology or geography. Clade 6, which represents a long, isolated branch, both phylogenetically and geographically, and has little genetic structure, is the only clade for which all methods of species delimitation agreed.

It seems likely that the complex and highly heterogeneous genetic landscape in *Juga* presents significant violations of the assumptions of these methods and that the estimates they produced are oversplit. Similarly, ABGD with K80 model-corrected distances produced comparable estimates of putative species diversity as those supported by bGMYC, bPTP and BP&P. The K80 model has been shown to produce inaccurate estimates in simulation studies for complex datasets, and to consistently underperform compared to more complex models (Barley and Thomson, 2016). Consequently, it is likely that the K80 model is not sufficiently complex to adequately describe the observed data. We conclude that these methods are of limited utility for inferring species limits in groups like *Juga* with significant genetic structure and complex phylogeographic histories.

4.5. Implications for the systematics of Juga

With nine species-level clades, diversity of *Juga* is concluded to be lower than presently thought (Johnson et al., 2013; NatureServe, 2017). Only in two instances do the resulting clades correspond to traditional taxonomic groupings: *J. newberryi* (Clade 4) from the Deschutes drainage, and *J. occata* (Clade 8) from the northern California Pit River system. Even in these two cases, some of the forms revealed here to belong to these two species have been relegated to other taxonomical species or to represent putatively undescribed species. For example, *J. newberryi* includes both the Opal Springs (or Crooked River) Juga and Purple Juga, both previously thought to represent potentially new species (Frest and Johannes, 1995). *Juga occata*, traditionally conceived as a robust, heavily ornamented large river form surviving in the main course of the Pit River (Frest and Johannes, 1993; Furnish, 2005), is here revealed to include populations that occur in lakes and spring fed creeks that are less heavily sculptured and would be referred to *J. acutifilosa* as previously conceived (e.g., Fig. 4M–Q). Several other clades combine multiple taxa that are currently recognized as valid (Johnson et al., 2013). Pending a comprehensive systematic revision, these are distinguished here as the *plicifera* group (Clade 2; including *J. plicifera, J. hemphilli, J. silicula*) from south-central Oregon to northern Washington; the *nigrina* group (Clade 3; with *J. nigrina, J. chacei*) from

central California and south-central Oregon; and the *acutifilosa* group (Clade 9; combining *J. acutifilosa, J. interioris, J. laurae*) from northern California, southeastern Oregon and northwestern Nevada. The clade from the central Columbia and Deschutes drainages in north-central Oregon (Clade 5) has been referred to as the Basalt Juga, the Three-Band Juga and the One-Band Juga among other names (Frest and Johannes, 1995; Furnish, 2005). Campbell et al. (2016) referred members of this clade to *J. bairdiana* (Lea, 1862) and *J. hemphilli maupinensis* (Henderson, 1935), of which the former would have priority. However, pending a comprehensive revision and review of all relevant type material, it is not clear if this is the oldest available name for members of this clade. Three clades seem to be without an available species group name and will require formal description: *Juga* sp. 1 (Clade 1) from southern Pit and eastern Sacramento tributaries, the highly distinct Blue Mountains Juga (Clade 6) from eastern Oregon (Frest and Johannes, 1995), and *Juga* sp. 3 (Clade 7) from the northern California upper Sacramento and south-central Oregon upper Klamath drainages. The latter includes what has been referred to as the Cinnamon Juga from springs along the upper Sacramento River in Siskiyou County; the spirally ornamented south-central Oregon forms in this clade have been referred to *J. acutifilosa*. Thus, the circumscription and synonymies of most species require extensive revision.

All other putative new species abandoned in open nomenclature are not supported as distinct OTUs in any of our analyses. Consequently, the majority seem to represent localized morphological variants of more widespread species. Of the two petitioned for listing, the Basalt Juga and Cinnamon Juga, both were returned as members of more widespread but highly fragmented species, *Juga* sp. 2 and *Juga* sp. 3, respectively. Additional surveys are needed to determine if they merit conservation action. This result further highlights the importance of rigorous systematic analysis prior to listing (Liu et al., 2015).

The nine monophyletic clades supported here present a mixture of widely distributed species and narrow range endemics, with one recovered at only a single site. Interestingly, the two most widespread species are distributed in primarily large rivers from coastal drainages. These are also the most variable species with shell morphologies that broadly overlap in size, color, ornamentation and banding patterns. These two clades were robustly supported (PP = 1.0) as sister taxa in the StarBEAST (Fig. 5) and concatenated analyses (Figs. 2, S6 and S7), and ITS1 (Figs. S4, S5) showed evidence of incomplete lineage sorting and could not resolve the monophyly of the two; however, they were not supported as sister taxa on the mitochondrial gene trees (Figs. S1–S3). Additional independent loci are needed to resolve this conflict.

Features of teleoconch sculpture traditionally considered significant in species circumscription, including shell size, color and banding patterns, were confirmed to be variable within species and to intergrade between species, rendering them of limited utility for species circumscription. The strength and persistence of spiral and axial ornamentation, the primary characters for distinguishing between the named subgenera, also were confirmed to be variable within and between species, and sometimes between individuals at a single site. Thus, we caution against basing hypotheses of species status in Semisulcospiridae on shell structure alone, and our findings are likely applicable to other freshwater Cerithioidea (e.g., Pleuroceridae, Melanopsidae). Furthermore, subgeneric designations based primarily on shell characters do not appear biologically meaningful.

In contrast to modern studies of groundwater-dependent springsnail and pebblesnail diversity in the Pacific Northwest (e.g., Hershler and Frest, 1996; Hershler et al., 2003, 2007; Hershler and Liu, 2010), the diversity of *Juga* is found to be lower than presently appreciated. Two of the species as circumscribed herein, have broad distributions, while the others are geographically restricted and highly fragmented, making them susceptible to human-mediated impacts. The Blue Mountains Juga was recovered from only a single site, but more surveys are necessary to establish how many populations of this fragile, highly distinct species still survive.

5. Conclusions

We used ABGD analyses to formulate primary species hypotheses that we further refined using phylogenetic analysis of an expanded mitochondrial and nuclear gene dataset within an integrative framework. Our secondary species hypotheses, based on a combination of reciprocal monophyly and informed by a combination of geography and morphology, recognize a mixture of widespread, highly variable species and narrow range endemics. Features of teleoconch sculpture considered significant in subgeneric classification are confirmed to be variable within some putative species (e.g., Fig. 4R–W). The diversity of *Juga* is concluded to be lower than presently recognized; only two species-level clades supported in the present study correspond to species currently recognized (*J. newberryi*, *J. occata*). Three other clades do not seem to have any available species-group names among the list of synonyms, only one of which corresponds to one of the undescribed putative species identified in the grey literature (Blue Mountains Juga). The majority of populations previously identified in the grey literature or by Campbell et al. (2016) as representing potentially new species do not comprise distinct OTUs in our analyses. This includes two undescribed putative species that have been petitioned for listing under the U.S. Endangered Species Act (Basalt Juga, Cinnamon Juga) (USFWS, 2012).

This research presents a valuable empirical case study exploring the performance of several tree-based algorithmic species delimitation methods in analyses of species with low vagility, poor dispersal capacity, and complex phylogeographic histories. Our results are consistent with those of recent simulation studies demonstrating that methods implementing the multispecies coalescent model are sensitive to genetic structure and will overestimate species diversity. Thus, we view this as a cautionary tale about the uncritical use of tree-based algorithmic methods of species delimitation in species groups with high levels of genetic diversity and population structuring.

Acknowledgements

This project was supported by an award from the Bureau of Land Management, Oregon State Office (L11AC20325, modification no. 2). This work was made possible in part by a grant of high-performance computing resources and technical support from the Alabama Supercomputer Authority. We gratefully acknowledge Jeffrey Garner (Alabama Division of Wildlife and Freshwater Fisheries) and Philippe Bouchet (Paris Museum of Natural History) for assistance in the field; Amanda Windsor, Herman Wirshing and Jordan Casey (NMNH Laboratories of Analytical Biology) for generating the sequences; Elizabeth Diamond (USNM) for photographing the sequenced vouchers; and Dan Cole (USNM) for generating the distribution map. We thank Kelli Van Norman (BLM), Candace Fallon and Emilie Blevins (The Xerces Society for Invertebrate Conservation), and Darci Rivers-Pankratz (U.S. Forest Service, Interagency Special Status/Sensitive Species Program) for collecting specimens and/or for arranging specimens to be collected on our behalf. Special thanks to Kelli Van Norman for making this work possible. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service. We are grateful to three anonymous reviewers for their careful reading of the manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2019.04.009.

References

Astrin, J.J., Höfer, H., Spelda, J., Holstein, J., Bayer, S., Hendrich, L., Huber, B.A., Kielhorn, K.-H., Krammer, H.-J., Lemke, M., Monje, J.C., Morinière, J., Rulik, B.,

Petersen, M., Janssen, H., Muster, C., 2016. Towards a DNA barcode reference database for spiders and harvestmen of Germany. PLoS ONE 11 (9), e0162624. https://doi.org/10.1371/journal.pone.0162624.
Bagley, J.C., Alda, F., Breitman, M.F., Bermingham, E., van den Berghe, E.P., Johnson, J.B., 2015. Assessing species boundaries using multilocus species delimitation in a morphologically conserved group of Neotropical freshwater fishes, the *Poecilia sphenops* species complex (Poeciliidae). PLoS ONE 10 (4), e0121139. https://doi.org/10.1371/journal.pone.0121139.
Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.
Barley, A.J., Thomson, R.C., 2016. Assessing the performance of DNA barcoding using posterior predictive simulations. Mol. Ecol. 25, 1944–1957.
Barley, A.J., White, J., Diesmos, A.C., Brown, R.M., 2013. The challenge of species delimitation at the extremes: diversification without morphological change in Philippine sun skinks. Evolution 67, 3556–3572. https://doi.org/10.1111/evo.12219.
Becker, M., Zielske, S., Haase, M., 2016. Conflict of mitochondrial phylogeny and morphology based classification in a pair of freshwater gastropods (Caenogastropoda, Truncatelloidea, Tateidae) from New Caledonia. ZooKeys 603, 17–32.
Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537. https://doi.org/10.1371/journal.pcbi.1003537.
Brown, J.M., Lemmon, A.R., 2007. The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. Syst. Biol. 56, 643–655.
Burch, J.B., 1989. North American Freshwater Snails. Malacological Publications, Hamburg, MI, pp. 365.
Campbell, D.C., Clark, S.A., Johannes, E.J., Lydeard, C., Frest, T.J., 2016. Molecular phylogenetics of the freshwater gastropod genus *Juga* (Cerithioidea: Semisulcospiridae). Biochem. Syst. Ecol. 65, 158–170.
Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.
Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2, e68.
Delicado, D., Ramos, M.A., 2012. Morphological and molecular evidence for cryptic species of springsnails [genus *Pseudamnicola (Corrosella*) (Mollusca, Caenogastropoda, Hydrobiidae)]. ZooKeys 190, 55–79.
Dray, S., Dufour, A.B., 2007. The ade4 package: implementing the duality diagram for ecologists. J. Stat. Softw. 22, 1–20.
Drummond, A.J., Suchard, M.A., 2010. Bayesian random local clocks, or one rate to rule them all. BMC Biol. 8, 114. https://doi.org/10.1186/1741-7007-8-114.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of molecular variance inferred from matrix distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479–491.
Fedosov, A.E., Stahlschmidt, P., Puillandre, N., Aznar-Cormano, L., Bouchet, P., 2017. Not all spotted cats are leopards: evidence for a *Hemilienardia ocellata* species complex (Gastropoda: Conoidea: Raphitomidae). Euro. J. Taxonomy 268, 1–20. https://doi.org/10.5852/ejt.2017.268.
Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech, 3, 294–299.
Frest, T.J., Johannes, E.J., 1995. Interior Columbia Basin mollusk species of Special Concern. In: Final Report prepared for Interior Columbia Basin Ecosystem Management Project, Walla Walla, Washington. Deixis Consultants, Seattle, Washington, pp. 274.
Frest, T.J., Johannes, E.J., 1993. Mollusc species of Special Concern within the range of the Northern Spotted Owl. Department of Agriculture, Forest Service, Portland, Oregon. Deixis Consultants, Seattle, Washington, Final Report prepared for Forest Ecosystem Management Working Group, U. S, pp. 98.
Frest, T.J., Johannes, E.J., 1999. Field Guide to Survey and Manage Freshwater Mollusk Species from the Northwest Forest Plan, BLM/OR/WA/PL-99/045 + 1792. USDI Bureau of Land Management, Portland, OR, pp. 117.
Frest, T.J., Johannes, E.J., 2010–2011... Review of the species of *Juga* (western U.S. Cerithioidea, Pleuroceridae, Semisulcospirinae). Malacol. Rev. 43–44, 1–61.
Fujisawa, T., Barraclough, T.G., 2013. Delimiting species using single-locus data and the generalized mixed yule coalescent approach: a revised method and evaluation on simulated datasets. Syst. Biol. 62, 707–724.
Fukuda, H., Haga, T., Tatara, Y., 2008. Niku-nuki: a useful method for anatomical and DNA studies on shell-bearing molluscs. Zoosymposia 1, 15–38.
Furnish, J.L., 2005. Sensitive aquatic mollusks of the US Forest Service Pacific Southwest Region. USDA Forest Service, Pacific Southwest Region, Vallejo, CA, pp. 23.
Furnish, J.L., Monthey, R., 1998. Conservation Assessments for Mollusk Species associated with springs and spring runs: *Fluminicola* new species 2, 3, 11; *Vorticifex klamathensis sinitsini*; *Juga* (*Oreobasis*) new species 2; and *Lyogyrus* new spp. 1. In: Report submitted to USDI Bureau of Land Management, Salem, OR, December 1998, pp. 17.
Furnish, J.L., Monthey, R., 1999. Management Recommendations for Aquatic Mollusks. Version 2.0. Report submitted to USDI Bureau of Land Management, Salem, OR, December 1998.
Furnish, J., Monthey, R., Applegarth, J., 1997. Survey protocol for terrestrial mollusk species from the Northwest Forest Plan. Version 3.1. In: Report to the USDI Bureau of Land Management, Salem, Oregon, October 29, 1997, pp. 52.
Furnish, J.L., 1990. Factors affecting the growth, production and distribution of the stream snail *Juga silicula* (Gould). In: Ph.D. Dissertation, Oregon State University, pp. 173.
Furnish, J.L., 2007. Guide to Sensitive Aquatic Mollusks of the U.S. Forest Service Pacific Southwest Region. In: Report submitted to the USDA Forest Service Pacific Southwest Region, May 2007, pp. 23.
Geller, J., Meyer, C., Parker, M., Hawk, H., 2013. Redesign of PCR primers for mitochondrial cytochdrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surverys. Mol. Ecol. Resour. 13, 851–861.
Goodrich, C., 1942. The pleuroceridae of the Pacific coastal drainage, including the western interior basin. Occas. Papers . Museum Zool. 469, 1–4.
Groenenberg, D.S.J., Neubert, E., Gittenberger, E., 2011. Reappraisal of the "Molecular phylogeny of western Palaearctic Helicidae s.l. (Gastropoda: Stylommatophora)": when poor science meets GenBank. Mol. Phylogenet. Evol. 61, 914–923.
Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.
Hawkins, C.P., Furnish, J.K., 1987. Are snails important competitors in stream ecosystems? Oikos 49, 209–220.
Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27, 570–580. https://doi.org/10.1093/molbev/msp274.
Henderson, J., 1935. West American species of *Goniobasis*, with descriptions of new forms. Nautilus 48, 94–99.
Hershler, R., Frest, T.J., 1996. A review of the North American freshwater snail genus *Fluminicola* (Hydrobiidae). Smithsonian Contrib. Zool. 583, 1–41.
Hershler, R., Liu, S.-P., Frest, T.J., Johannes, E.J., 2007. Extensive diversification of pebblesnails (Lithoglyphidae: *Fluminicola*) in the upper Sacramento river basin, northwestern United States. Zool. J. Linn. Soc. 149, 371–422.
Hershler, R., Frest, T.J., Liu, S.-P., Johannes, E.J., 2003. Rissooidean snails from the Pit river basin, California. Veliger 46, 275–304.
Hershler, R., Liu, S.-P., 2010. Two new, possibly threatened species of *Pyrgulopsis* (Gastropoda: Hydrobiidae) from southwestern California. Zootaxa 2343, 1–17.
Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S., 2017. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.
Johnson, P.D., Bogan, A.E., Brown, K.M., Burkhead, N.M., Cordeiro, J.R., Garner, J.T., Hartfield, P.D., Lepitzki, D.A.W., Mackie, G.L., Pip, E., Tarpley, T.A., Tiemann, J.R., Whelan, N.V., Strong, E.E., 2013. Conservation status of freshwater gastropods of Canada and the United States. Fisheries 38, 247–282.
Kainer, D., Lanfear, R., 2015. The effects of partitioning on phylogenetic inference. Mol. Biol. Evol. 32, 1611–1627.
Kamvar, Z.N., Tabima, J.F., Grünwald, N.J., 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2, e281.
Kappes, H., 2012. Slow, but steady: dispersal of freshwater molluscs. Aquat. Sci. 74, 1–14.
Köhler, F., 2017. Against the odds of unusual mtDNA inheritance, introgressive hybridisation and phenotypic plasticity: systematic revision of Korean freshwater gastropods (Semisulcospiridae, Cerithioidea). Invertebr. System. 31, 249–268.
Kubatko, L.S., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56, 17–24.
Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772–773. https://doi.org/10.1093/molbev/msw260.
Leaché, A.D., Zhu, T., Rannala, B., Yang, Z., 2018. The spectre of too many species. Syst. Biol. 00, 113.
Liu, H.-P., Hershler, R., Rossel, C.S., 2015. Taxonomic status of the Columbia duskysnail (Truncatelloidea, Amnicolidae, Colligyrus). ZooKeys 514, 1–13.
Liu, H.-P., Marceau, D., Hershler, R., 2016. Taxonomic identity of two amnicolid gastropods of conservation concern in lakes of the Pacific Northwest of the USA. J. Molluscan Stud. 82, 464–471.
Minton, R.L., Norwood, A.P., Hayes, D.M., 2008. Quantifying phenotypic gradients in freshwater snails: a case study in *Lithasia* (Gastropoda: Pleuroceridae). Hydrobiologia 605, 173–182.
NatureServe., 2017. NatureServe Explorer: An online encyclopedia of life [web application]. Version 7.1. NatureServe, Arlington, Virginia. Available http://explorer.natureserve.org. (Accessed: April 8, 2018).
Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.
Obertegger, U., Cieplinski, A., Fontaneto, D., Papakostas, S., 2018. Mitonuclear discordance as a confounding factor in the DNA taxonomy of monogonont rotifers. Zoolog. Scr. 47, 122–132.
S.R. Palumbi, A. Martin, S. Romano, W.O. McMillan, L. Stice, G. Grabowski. The simple fool's guide to PCR, version 2.0, privately published document compiled by S. Palumbi, Dept. Zoology, Univ. Hawaii, Honolulu, HI; 1991.
Paradis, E., Claude, J., Strimmer, K., 2004. APE: analysis of phylogenetics and evolution in R language. Bioinformatics 20, 289–290.
Puillandre, N., Modica, M.V., Zhang, Y., Sirovich, L., Boisselier, M.C., Cruaud, C., Holford, M., Samadi, S., 2012b. Large-scale species delimitation method for hyperdiverse groups. Mol. Ecol. 21, 2671–2691. https://doi.org/10.1111/j.1365-294X.2012.05559.x.
Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012a. ABGD, automatic barcode gap discovery for primary species delimitation. Mol. Ecol. 21, 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x.
R Development Core Team, 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
A. Rambaut. A.J. Drummond. Tracer v1.4. Available from http://beast.bio.ed.ac.uk/Tracer; 2007.

Rannala, B., Yang, Z., 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 164, 1645–1656.

Reid, N.M., Carstens, B.C., 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evol. Biol. 12, 196. https://doi.org/10.1186/1471-2148-12-196.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.

Rosel, P.E., Hancock-hanser, B.L., Archer, F.I., Robertson, K.M., Martien, K.K., Leslie, M.S., Berta, A., Cipriano, F., Viricel, A., Viaud-Martinez, K.A., Taylor, B.L., 2017. Examining metrics and magnitudes of molecular genetic differentiation used to delimit cetacean subspecies based on mitochondrial DNA control region sequences. Mar. Mammal Sci. 33, 76–100.

Strong, E.E., Frest, T., 2007. On the anatomy and systematics of *Juga* from western North America (Gastropoda: Cerithioidea: Pleuroceridae). Nautilus 121, 43–65.

Strong, E.E., Gargominy, O., Ponder, W.F., Bouchet, P., 2008. Global diversity of gastropods (Gastropoda: Mollusca) in freshwater. Hydrobiologia 595, 149–166.

Strong, E.E., Köhler, F., 2009. A morphological and molecular analysis of "*Melania*" *jacqueti* Dautzenberg & Fischer, 1906: from anonymous orphan to critical basal offshoot of the Semisulcospiridae (Gastropoda: Cerithioidea). Zoolog. Scr. 38, 483–502. https://doi.org/10.1111/j. 1463-6409. 2008. 00385.x.

Sukumaran, J., Knowles, L., 2017. Multispecies coalescent delimits structure, not species. Proc. Nat. Acad. Sci., USA 114, 1607–1612.

Taylor, D.W., 1966. Summary of North American Blancan nonmarine mollusks. Malacologia 4, 1–172.

Taylor, D.W., 1981. Freshwater mollusks of California: a distributional checklist. California Fish Game 67, 140–163.

Templado, J., Richter, A., Calvo, M., 2016. Reef building Mediterranean vermetid gastropods: disentangling the *Dendropoma petraeum* species complex. Medit. Marine Sci. 17, 13–31.

Thomaz, D., Guiller, A., Clarke, B., 1996. Extreme divergence of mitochondrial DNA within species of pulmonate land snails. In: Proceedings of the Royal Society London B, pp. 363–368.

USDA [United States Department of Agriculture], USDI [United States Department of the Interior]. Standards and guidelines for management of habitat for late-successional and old-growth forest related species within the range of the northern spotted owl. Attachment A to the record of decision for amendments to Forest Service and Bureau of Land Management planning documents within the range of the northern spotted owl. Available from http://www.reo.gov/documents/reports/newsandga.pdf; 1994.

USDA [United States Department of Agriculture], USDI [United States Department of the Interior]. Final Supplemental Environmental Impact Statement for Amendment to the Survey & Manage, Protection Buffer, and other Mitigation Measures Standards and Guidelines VOLUME II – APPENDICES. BLM/OR/WA/PT-00/065 + 1792. Available from https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprd3841012.pdf; 2000.

USFWS [United States Fish and Wildlife Service], 2011. Endangered and threatened wildlife and plants; 90-day finding on a petition to list 29 mollusk species as threatened or endangered with critical habitat. Fed. Reg. 76, 61826–61853.

USFWS [United States Fish and Wildlife Service], 2012. Endangered and threatened wildlife and plants; 12-month finding on a petition to list 14 aquatic mollusks as endangered or threatened. Fed. Reg. 77, 57922–57948.

Whelan, N.V., Johnson, P.D., Harris, P.M., 2012. Presence or absence of carinae in closely related populations of *Leptoxis ampla* (Anthony, 1855) (Gastropoda: Cerithioidea: Pleuroceridae) is not the result of ecophenotypic plasticity. J. Molluscan Stud. 78, 231–233.

White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR Protocols: A Guide to Methods and Applications. Academic Press Inc, San Diego, pp. 315–322.

Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A., 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29, 2869–2876. https://doi.org/10.1093/bioinformatics/btt499.