

A molecular taxonomy of *Cottus* in western North America

MICHAEL K. YOUNG^{1,*}, REBECCA SMITH^{1,4}, KRISTINE L. PILGRIM¹, DANIEL J. ISAAK², KEVIN S. MCKELVEY¹, SHARON PARKES², JACOB EGGE³, AND MICHAEL K. SCHWARTZ¹

¹USDA Forest Service, National Genomics Center for Wildlife and Fish Conservation, Rocky Mountain Research Station, 800 E. Beckwith Avenue, Missoula, MT 59802

²USDA Forest Service, Rocky Mountain Research Station, 322 East Front Street Suite 401, Boise, ID 83702

³Department of Biology, Pacific Lutheran University, Tacoma, WA 98447

⁴Present address: Department of Ecology & Evolutionary Biology, University of Tennessee, 569 Dabney Hall, Knoxville, TN 37996

ABSTRACT.—The taxonomy of sculpins (*Cottus*, Cottidae) remains one of the last major unresolved puzzles in the systematics of North American freshwater fishes. We used molecular approaches to identify candidate taxa and their distribution across western North America. We crowd-sourced the collection of specimens ($n = 8272$) via outreach to biologists in the western United States and Canada. From that collection, we sequenced—at up to 2 mitochondrial and 2 nuclear genes—a subset ($n = 4009$) of specimens from most basins in the western United States, added sequences from public sequence databases, and applied an array of species delimitation and specimen identification methods to assess phylogenetic and spatial patterns of diversity. Species delimitation methods, primarily relying on a conservative interpretation of the phylogenetic species concept, were broadly concordant and indicated that 43 candidate species were present. Some named taxa were unsupported, whereas others, if recognized, would violate the phylogenetic species concept. Specimen assignment was largely unambiguous and geographic distributions were consistent with phylogeographic patterns in other taxa. Our work establishes a benchmark for understanding the diversity of sculpin in western North America and suggests new species hypotheses both there and in eastern North America.

RESUMEN.—La taxonomía de los sculpins (*Cottus*, Cottidae) sigue siendo uno de los últimos grandes rompecabezas sin resolver en la sistemática de los peces de agua dulce de Norteamérica. Utilizamos métodos moleculares para identificar los taxones candidatos y su distribución en el oeste de Norteamérica. Recolectamos especímenes ($n = 8272$) a través de una campaña de divulgación entre biólogos del oeste de Estados Unidos y Canadá. A partir de esa colección, secuenciamos hasta 2 genes mitocondriales y 2 nucleares de un subconjunto ($n = 4009$) de especímenes, pertenecientes a la mayoría de las cuencas del oeste de Estados Unidos. Adicionalmente, añadimos secuencias de bases de datos de secuencias públicas y aplicamos una serie de métodos de delimitación de especies e identificación de especímenes para evaluar los patrones filogenéticos y espaciales de diversidad. Los métodos de delimitación de especies basados principalmente en una interpretación conservadora del concepto filogenético de especie, fueron ampliamente concordantes e indicaron la presencia de 43 especies candidatas. Algunos de los taxones nombrados no estaban respaldados, mientras que otros, de ser reconocidos, violarían el concepto filogenético de especie. La asignación de especímenes fue en gran medida inequívoca y las distribuciones geográficas fueron coherentes con los patrones filogeográficos de otros taxones. Nuestro trabajo establece un punto de referencia para entender la diversidad de peces *Cottus* en el oeste de Norteamérica y sugiere nuevas hipótesis de especies tanto allí como en el este de Norteamérica.

For centuries, morphological differences have been the basis for recognizing species and developing taxonomies of complex organisms. This approach has been successful for most taxa, yet there remain groups for which taxonomies based on the morphospecies concept are unsatisfying (Honeycutt et al. 2010). Among these are the circumboreal freshwater


fishes of the genus *Cottus* (Actinopterygii: Perciformes: Cottidae) (Betancur-R et al. 2017), commonly known as sculpins. Sculpins are among the most difficult freshwater fishes to identify based on appearance (Jenkins and Burkhead 1994, Moyle 2002, McPhail 2007) due to phenotypic variation in putatively diagnostic characters related to age, sex, habitat,

*Corresponding author: michael.k.young@usda.gov


MKY  orcid.org/0000-0002-0191-6112

RS  orcid.org/0000-0002-4085-5922

KLP  orcid.org/0000-0003-3726-6980

KSM  orcid.org/0000-0002-1399-3166

SP  orcid.org/0000-0003-1138-4321

MKS  orcid.org/0000-0003-3521-3367

geography, sympatry with congeners, or unknown factors (Robins and Miller 1957, Lyons 1990, Dennenmoser et al. 2015, Lucek et al. 2018, McLeish et al. 2020). Problems with species recognition are further compounded by the presence of recent hybrids between species (Zimmerman and Wooten 1981, Strauss 1986, Rudolfson et al. 2019) and taxa apparently of hybrid origin (Nolte et al. 2009). This has resulted in a high degree of taxonomic instability, with species formerly recognized as local endemics often synonymized with more broadly distributed species (Bailey and Bond 1963) or once-widespread species divided into a host of taxa with restricted distributions (Freyhof et al. 2005). Although endemism is consistent with the limited dispersal thought to be typical of sculpins (Hudy and Shiflet 2009, Gray et al. 2018), *C. bairdii* and *C. cognatus* are among the most widely distributed small-bodied fishes in North America (Page and Burr 2011) and can show little genetic population structure at broad spatial scales (Euclide et al. 2018).

Molecular approaches have resolved a number of issues associated with morphologically cryptic or confusing taxa (Bickford et al. 2007, Fišer et al. 2018). Over the last 2 decades, DNA barcoding (Hebert et al. 2003)—primarily directed at the identification of specimens based on distance measures associated with sequences of a portion of a mitochondrial gene—has proven successful because genetic variation within fish species is small relative to the divergence between species (i.e., the barcode gap) and nearly all fish species in North America are represented in public databases (April et al. 2011), making taxonomic assignment of most specimens straightforward (Ratnasingham and Hebert 2013). In addition, this method has been augmented with a host of single- and multiple-locus approaches for species delineation and discovery, in which one or more genes, divergence metrics, and genetic algorithms are used to identify species boundaries and evaluate whether existing taxonomies have phylogenetic support (Carstens et al. 2013, Fontaneto et al. 2015). Reliable molecular species delimitation, however, requires that all lineages of the taxa of interest be represented from across their geographic distributions (Bergsten et al. 2012, Ahrens et al. 2016), and the method is complicated by individuals or clades exhibit-

ing recent or ancient introgression (Solís-Lemus et al. 2016).

Application of these molecular tools to resolving some elements of the taxonomy of sculpins suggests the need for a comprehensive review of this group. April et al. (2011) observed that Cottidae in North America were disproportionately represented both by species that appeared insufficiently diverged to merit taxonomic recognition and by generic taxa composed of deeply divergent yet morphologically cryptic lineages, each deserving its own name. Examples of the latter appear to be particularly prominent among sculpins in western North America, for which species boundaries are unsettled (McPhail 2007, Baumsteiger et al. 2012, Young et al. 2013) and some taxa are of uncertain affiliation (Neely 2003, Kinziger et al. 2005).

Resolving species boundaries and higher-level relationships is important because sculpins are among the most diverse freshwater fishes in temperate Northern Hemisphere ecosystems (Goto et al. 2015). In western North America, the genus *Cottus* comprises 3 major groups (sometimes considered subgenera or proposed as new genera; Smith and Busby 2014, Goto et al. 2015): (1) *Cottopsis* (12 species and subspecies), restricted to Pacific coastal basins; (2) *Uranidea* (9 western species and subspecies), found throughout the U.S. and Canada, with one species, *C. cognatus*, present in Siberia; and (3) *Cottus*, primarily in Europe and northern Asia but represented by one species in North America, *C. ricei*. A fourth group consists of 5 species of uncertain affiliation (Supplementary Material 1). Based primarily on analyses of mitochondrial data, *Cottopsis* is thought to be sister to *Uranidea*, *Cottus*, and the unaffiliated taxa (Kinziger et al. 2005, Goto et al. 2015).

Adding to their complexity is the diversity of habitats these fishes occupy and their variable life histories. Sculpins are found in small streams, large rivers, lakes, and estuaries, often in great abundance (Jenkins and Burkhead 1994, McPhail 2007). Different species of sculpins may be sympatric, but closely related taxa almost never co-occur (Jenkins and Burkhead 1994, Quist et al. 2004). Sculpins are generally considered coolwater or coldwater fishes, of which *C. cognatus* and *C. confusus* occupy the coldest habitats (Bailey and Bond 1963, Reimers and Bond 1967, Lyons 1990).

Members of *Cottopsis* are salinity tolerant and can presumably use marine or estuarine environments as migration corridors, whereas the other groups are restricted to freshwater habitats (Goto et al. 2015). Some members of *Cottopsis* are also facultatively amphidromous, with females in fluvial habitats laying small eggs that give rise to planktonic larvae that drift to the ocean or large mainstem habitats and subsequently migrate upstream after a period of growth (Goto et al. 2015). Similarly, lake-based populations of some species of each subgenus commonly have pelagic young that can disperse widely, whereas fluvial populations are thought to be dispersal limited because females are assumed to lay large eggs that hatch into fry that immediately become benthic (Tabor et al. 2017, but see Sheldon 1968). Additionally, because sculpins have long been used as baitfish by anglers (e.g., Hubbs and Schultz 1932), translocations outside of their historical range are suspected to be common (Young et al. 2013).

Our goal was to use molecular techniques to delineate candidate species and identify specimens (*sensu* Collins and Cruickshank 2013) of *Cottus* across western North America. This work included representatives of 2 species, *C. bairdii* and *C. cognatus*, found primarily in eastern North America but believed to be present farther west. We engaged a host of stakeholders to build a continental-scale data set (Isaak et al. 2018) of thousands of specimens. Using sequences of many of these specimens, we applied an array of species delineation methods to develop a genetically based taxonomy of candidate species. Where possible, candidate species were elevated to named provisional species based on the assignment of representatives from the location of current or former type specimens. Finally, we described the distribution of candidate and provisional species based on molecular assignment of specimens from across this region, evaluated morphological identifications in light of the molecular taxonomy, and highlighted examples of ancient and recent hybridization.

METHODS

Geographic Setting

We focused on the basins associated with the Western Cordillera of North America that

host native populations of *Cottus* for which the shifting patterns of fluvial connectivity could explain the modern distribution and biodiversity in this group. Physiographically, western North America is composed of several master river basins—the Colorado, Sacramento–San Joaquin, Columbia, Fraser, Yukon, McKenzie, Saskatchewan, and Missouri—and a host of smaller coastal river basins, as well as the endorheic Great Basin. Over the last 20 Ma, the paleohydrology of the region has shifted in response to broad tectonic and climatic forces. Crustal extension collapsed the Nevadaplano and led to the formation of the Great Basin, whereas the subsequent exhumation of the Rocky Mountains from crustal uplift generated the modern southward-flowing Colorado River (Spencer et al. 2008). Movement of the North American Plate across the mantle plume represented by the Yellowstone hotspot, combined with Basin and Range faulting, caused the Continental Divide to migrate eastward hundreds of kilometers (Pierce and Morgan 2009). Among the most prominent examples of stream capture were the enlargement of the paleo–Snake River headwaters and the shifting of its downstream course from the Sacramento or Klamath River in the late Miocene and early Pliocene to the Columbia River, roughly since the beginning of the Pleistocene (Smith et al. 2002). Topographic shifts related to the mantle plume, influenced by Pleistocene glaciation, also led to repeated connection and isolation of the upper Snake River, Great Basin, and Colorado River (Broughton and Smith 2016). That isolation was likely heightened by eruptions from this volcanic center. For example, the 0.62 My Lava Creek B eruption deposited tens of meters of ash downwind across thousands of square kilometers (Izett and Wilcox 1982). This also contributed to the high sedimentation rates—400 m/1 Ma—that would have rerouted channels in and across the Great Basin (Hintze 1988). Ephemeral fluvial corridors were also opened when increasing precipitation during the Pleistocene led to the formation and occasional spillover of the giant pluvial Lakes Bonneville and Lahontan and a host of smaller satellite lakes (Reheis et al. 2002). Farther north, repeated continental glaciation at roughly 41- to 100-ky intervals (Hidy et al. 2013) erased aquatic habitats and isolated fauna in refugia south and north of the glacial margins

(Shafer et al. 2010), and rerouted the Missouri River from Hudson Bay to the Mississippi River basin (Howard 1958). During glacial maxima, lowered sea levels extended freshwater habitats seaward by tens of kilometers and facilitated the development of freshwater corridors linking North America and Asia across the Bering Land Bridge (Lindsey and McPhail 1986). At the southern margin of the ice sheets, periglacial lakes provided abundant aquatic habitat, but the failure of the ice dams that formed them exposed the aquatic fauna to dozens of monumental floods (Booth et al. 2003, O'Connor et al. 2020), some of which may have facilitated dispersal. Finally, following the Last Glacial Maximum, the successive formation of lakes at the receding northern and southern ice margins permitted colonization of many now-disjunct river basins (McPhail and Lindsey 1986).

The fossil record offers little resolution of the evolutionary responses of sculpins to these events. The majority of fossil sculpins in western North America are from lacustrine deposits (Cavender 1986), whereas most modern species are primarily fluvial. Cottids are absent from the western North American fossil record until the late Miocene (ca. 8.4 Ma; McClellan and Smith 2020), and although fossil sculpin of late Miocene and Pliocene deposits in Idaho and Oregon are abundant, these specimens may not be represented by modern descendants (Kimmel 1975, Smith et al. 1982, Van Tassel and Smith 2019). Instead, the progenitors of the current fauna may not have diverged from their sister taxa in Eurasia until ~6.2–2.5 Ma (Yokoyama and Goto 2005). Even relations of the modern fauna to more recent fossils are uncertain or the records are incomplete. Mid-Pleistocene fossil sculpin from the Mopung Hills Formation in the Lahontan Basin were hypothesized to be ancestor to either *C. pitensis* or *C. beldingii* (Taylor and Smith 1981, Reheis et al. 2002). A fossil of *C. bairdii* at least 1.4 million years old was reported from the Grand View local fauna in southern Idaho (Smith et al. 1982), whereas lacustrine deposits of pluvial Lake Bonneville no older than 15 ka yielded specimens of *C. extensus* and *C. bairdii*, but not of 2 other species likely present at that time, *C. echinatus* and *C. beldingii* (Broughton and Smith 2016). There are fossils of sculpin from early in the last glacial epoch from the McKenzie

River basin, but whether they represented *C. cognatus* or *C. ricei* was uncertain (Cumbaa et al. 1981). Thus, absent a more comprehensive and definitive fossil record, phylogenetic evidence represents the primary basis for understanding the evolutionary history of sculpins in this region.

Specimens and Genotyping

Samples for molecular analysis ($n = 4009$ from 1463 unique sites) were chosen from among 8272 specimens (Fig. 1, Supplementary Material 2) solicited via personal contacts, presentations, and a *Sculpins of the West* website describing this project (https://www.fs.fed.us/rm/boise/AWAE/projects/fish_tissue_collection.html). We attempted to obtain individuals from every 8-digit hydrologic unit (<https://water.usgs.gov/GIS/huc.html>) within most of the historical range of sculpins throughout the western U.S. and opportunistically from basins in western Canada, in habitats ranging from headwater streams to coastal lakes. We also tried to include 2 or more representatives from every currently or formerly named taxon from at or near their type locations (Supplementary Material 1). Because intensive phylogenetic work has been recently conducted on sculpins in coastal California basins (Baumsteiger et al. 2012, 2014, 2016, Kinzinger et al. 2016), we sought fewer samples from this area. Also, because *C. bairdii* and *C. cognatus* are thought to be present in western North America, we included members of these species from eastern North America. For comparison, we also analyzed specimens of *C. gobio* from Sweden, the first species of this genus to be described by Linnaeus in 1758.

We used the QIAGEN DNeasy Blood and Tissue kit to extract genomic DNA from tissues, following the manufacturer's instructions for tissue, and sequenced 2 mitochondrial regions, cytochrome *c* oxidase subunit 1 (COI) and cytochrome *b* (cytb), and 2 nuclear genes, rhodopsin and *S7*. These 4 genes were chosen because they exhibited a range of evolutionary rates, exposure to selection, and available reference sequences. The COI gene tends to be relatively conserved with respect to amino acid substitutions yet relatively variable for synonymous changes (Hebert et al. 2003), and its use permits comparison to the many sequences in public sequence repositories (e.g., the National Center for Biotechnology

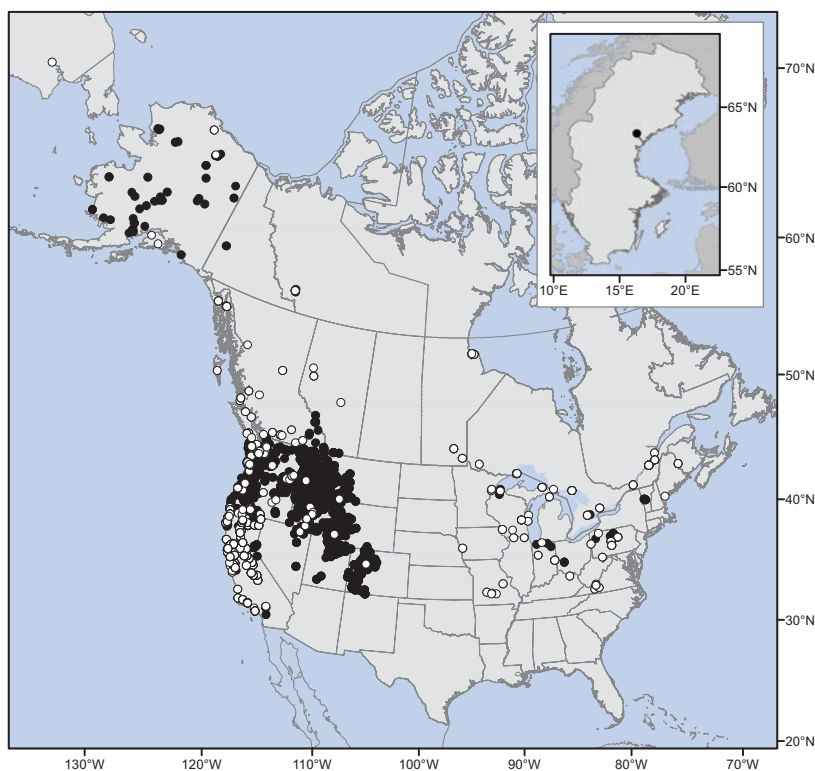


Fig. 1. Locations of new specimens (black dots; $n = 4009$) and specimens in public databases (white dots; $n = 847$) used in species delimitation and specimen assignment in *Cottus*. Inset, location of specimens of *C. gobio*.

Institute and Barcode of Life databases). The *cytb* region tends to be more variable than COI (at least in *Cottus*; Young et al. 2013) and is also among the most commonly sequenced genes among fishes (e.g., Ward 2009). The nuclear genes were chosen to ensure that species delineation was not wholly reliant on the more rapidly evolving mitochondrial regions and to permit identification of the hybrid origins of individuals and clades. Whereas S7 is a selectively neutral nuclear intron likely to exhibit high substitution rates (Haponski and Stepien 2013), rhodopsin was expected to be a slowly evolving gene under strong selection given its role in vision and coloration (Behrens-Chapuis et al. 2015, Hill et al. 2019). Not all individuals were sequenced at all genes. We began by sequencing all individuals at COI. Then, from each COI clade exhibiting a marked level of divergence and represented by more than one individual, we selected one or more individuals to sequence at the 3 additional genes. We used an array of

published and custom-designed primers and a standardized analytical regimen (Supplementary Material 3).

Sequences of COI, *cytb*, and rhodopsin were aligned by eye in MEGA 7.0 (Kumar et al. 2016). All lacked indels and were translated into amino acids to verify that stop codons were absent. Sequences of S7 had multiple indels that were coded using FastGap (Borchsenius 2009)—which implements the gap-coding algorithm of Simmons and Ochoterena (2000)—with the gap codes appended to the S7 sequences. These sequences were aligned with the online version of PRANK (Löytynoja and Goldman 2010; <https://www.ebi.ac.uk/goldman-srv/webprank/>). Sequences of nuclear genes with multiple heterozygous positions were phased into statistically inferred pairs (labeled *a* or *b*) using PHASE and seqPHASE (Stephens et al. 2001, Flot 2010). Those with length-variant forms (evident only in S7) were analyzed in CHAMPURU (Flot 2007) prior to phasing. Each nuclear phase was combined

with the mitochondrial sequences from that individual in concatenated analyses.

Summary statistics of variable sites, parsimony-informative sites, haplotype number and diversity, and nucleotide diversity and differences were calculated in DnaSP v6.12.03 (Rozas et al. 2017). Tests of substitution saturation for the mitochondrial genes were performed using DAMBE 7 (Xia 2018).

Species Delimitation and Specimen Assignment

We sought to identify potential candidate species from the molecular data, and where possible, to assign currently or previously valid names to those candidate species to elevate them to provisional species (see below). For defining species hypotheses, we adopted the phylogenetic species concept—which prioritizes reciprocal monophyly, which we identified using phylogenetic trees—modified to reflect species-level rather than population-level divergence, which we evaluated using a host of additional analyses. We further required the allopatric distribution of closely related candidate species (cf. April et al. 2011) and used field identification of specimens to indirectly consider morphological data, thus also satisfying the requirements for recognizing species under the unified species concept (de Queiroz 2007). Our null hypothesis was that the present taxonomy, for which we followed Fricke et al. (2021), was correct and reflected evolutionary relationships.

Separate data sets were developed for species delimitation and specimen assignment. For the former, the data consisted of (1) concatenated COI and cytb sequences ($n = 564$ sequences of 1657 nucleotides) and (2) concatenated mitochondrial and nuclear sequences ($n = 530$ sequences of 2710 nucleotides from 433 specimens, with 96 specimens having heterozygous nuclear phases). One specimen of *Leptocottus armatus* was used as an out-group in all analyses except evolutionary model evaluations.

We followed a series of steps for species delineation and specimen assignment (Fig. 2). The first step was to examine cytonuclear concordance to identify species complexes and lineages and specimens of hybrid origin. We built maximum-likelihood (ML) phylogenies using the four-gene concatenated data set but analyzed the mitochondrial and nuclear portions separately in IQ-TREE (Nguyen et al.

2015) implemented via the CIPRES gateway (<https://www.phylo.org/>). We assigned 6 preliminary partitions based on gene and codon position for COI and cytb, and then selected edge-linked partitions and the TESTNEWMERGE setting to determine the best-fitting substitution models, which were TIM2e + I + G (position 1), TN + F + I (position 2), and TN + F + R3 (position 3). For the nuclear sequences, there were 5 partitions based on codon position for rhodopsin (positions 1 and 2, K3P + R2; position 3, TVM + F + G) and on sequences (K3P + R2) and gapcodes for S7 (F81 + F + G). We assigned support values using 1000 ultrafast bootstraps to each consensus maximum-likelihood tree. Because ultrafast bootstrap values are relatively unbiased (Minh et al. 2013), we assumed that values >80 reflected acceptable support. We compared well-supported species complexes—blocks of candidate species composed of one or more recognized species and associated lineages—from the nuclear maximum-likelihood tree to well-supported but divergent individual lineages (i.e., which we deemed candidate species, as derived below) from the mitochondrial tree. Individuals with nuclear phases that assigned to different pairs of species complexes or were intermediate to them, with mitochondrial haplotypes identical to readily delineated candidate or provisional species (see below), were assumed to be hybrids of recent origin and were excluded from further analyses. We also expected groups of mitochondrial candidate species to assign to no more than one nuclear species complex. High support but discordant assignment between trees for groups of specimens was considered evidence of older gene flow and the hybrid origin of candidate species (Soucy et al. 2015).

Next, we used the larger concatenated mitochondrial data set to delimit species using 6 approaches. First, we analyzed these data in TCS 1.21 (Clement et al. 2000) to construct 95% statistical parsimony networks (SPN). Independent networks were regarded as candidate species (Hart and Sunday 2007), although this method is conservative because separate networks at this threshold are likely to underestimate species diversity (Hart and Sunday 2007, Chen et al. 2010). Second, we used the online version (<https://bioinfo.mnhn.fr/abi/public/asap/#>) of ASAP (Assemble Species by Automatic Partitioning; Puillandre

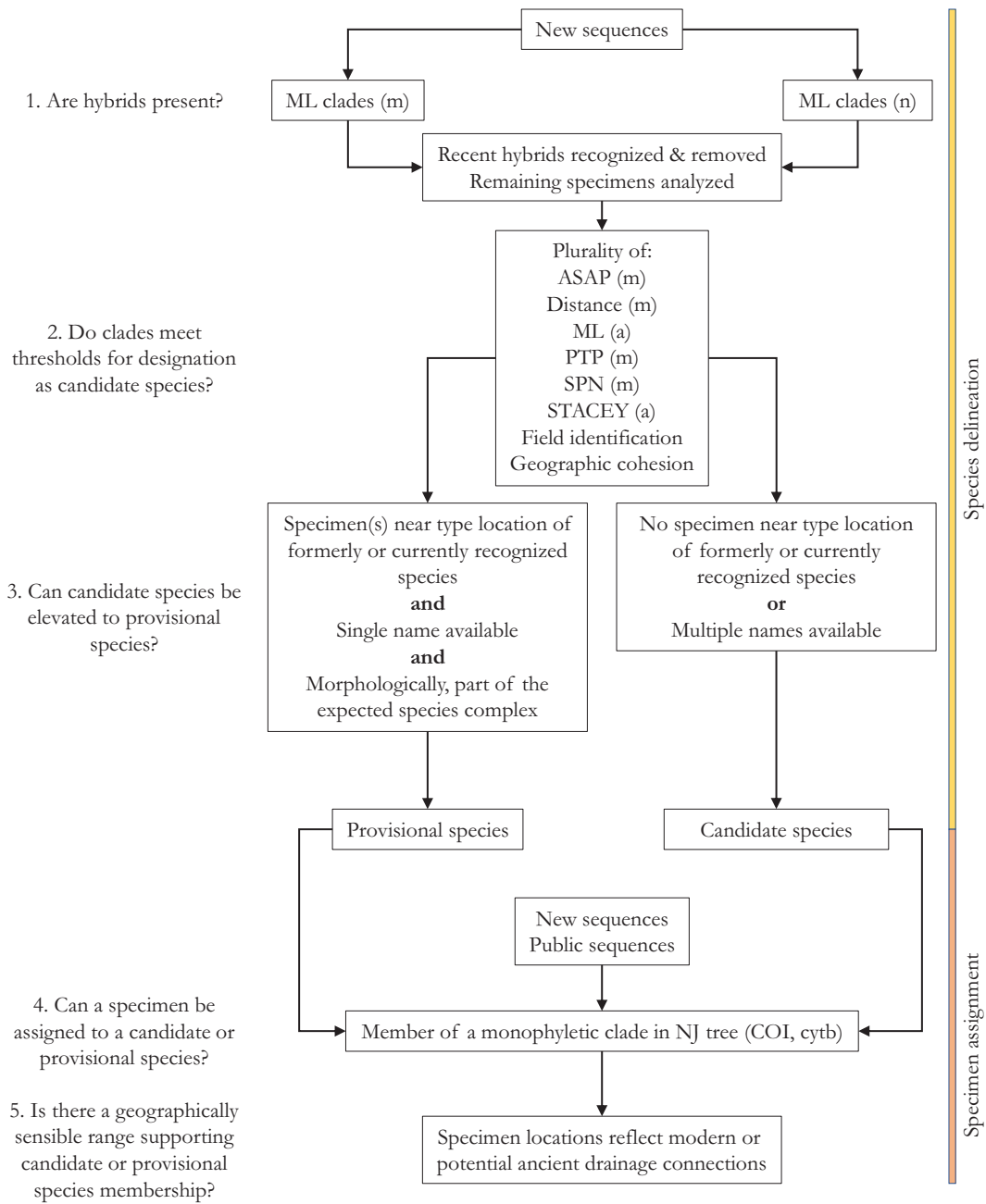


Fig. 2. Workflow for designation of candidate and provisional species and their distributions. Abbreviations (in parentheses): a = all genes, m = concatenated mitochondrial genes, n = concatenated nuclear genes.

et al. 2021), a method similar to ABGD (Automatic Bar Code Gap Discovery; Puillandre et al. 2012) in which genetic distances are used to identify the transition between intraspecific variation and interspecific divergence but which

includes a scoring system to identify the best-fitting set of partitions, i.e., candidate species. We adopted the default values and distances based on the K80 substitution model because of its similarity to the traditional distance

metric used in barcode-based analyses (Ratnasingham and Hebert 2013). The analysis was run 10 times with different initial seeds, for which the highest-scoring set of partitions did not vary. Because the approach is similar to ABGD, we assumed that species counts would also be conservative relative to other methods (Puillandre et al. 2021). Third, we used the single-threshold, maximum-likelihood implementation of Poisson tree processes (PTP; Zhang et al. 2013), which uses phylogenetic trees (from the maximum-likelihood analysis; see below) to identify the transition between species- and population-level divergence. We explored the Bayesian and multiple-threshold versions of PTP, but these analyses either failed to converge or produced unrealistic results, i.e., one-third to one-half of all sequences constituted candidate taxa. Fourth, we built another ML phylogeny on these mitochondrial data and considered strong support for candidate species to be a bootstrap value >80 for a reciprocally monophyletic clade identified by one of the aforementioned approaches. Fifth, we used MEGA 7.0 (Kumar et al. 2016) to build a pairwise distance matrix based on the absolute number of differences between sequences, and then we examined the maximum genetic distance among members of a candidate species and the minimum distance to a nonmember (Meier et al. 2008). When the latter exceeded the former and was $>1\%$ (a difference characteristic of the majority of interspecific differences among fishes; Hubert et al. 2008, Ward 2009), we considered this as strong support for a candidate species. This approach leveraged the existing osteichthyan taxonomy as a benchmark for species recognition (Galtier 2019). Sixth and finally, we considered whether a candidate species was allopatric with respect to other members of its species complex, generally pooling those that were sympatric into a single candidate species (Jenkins and Burkhead 1994).

We sought corroboration for candidate taxa by analyzing the four-gene data set using STACEY, a package within BEAST 2.5 (Bouckaert et al. 2019) which estimates a species tree and candidate taxa based on the multispecies coalescent (Jones 2017). Sequence variation related to gene flow among species that may have been of hybrid origin, however, appeared to prevent convergence in analyses

including all samples (Barley et al. 2018). Although we considered species network models that can address gene flow (e.g., PhyloNet and SpeciesNetwork; Wen et al. 2018, Zhang et al. 2018), these could not accommodate the missing data associated with indels in S7, or were too computationally intensive to run on the entire data set. Hence, we analyzed each species complex from the nuclear analyses independently in STACEY to evaluate the number of candidate species therein. To avoid overparameterization, we reduced the number of partitions to 3—the concatenated mitochondrial genes, rhodopsin, and S7—and used HKY as the substitution model for each. Parameter values were set in BEAUti 2.6.0.0. Automatic clock rates were disabled; the mean substitution rate was fixed; and site models, clock models, and trees were unlinked. For site models, substitution rates were estimated, gamma categories were set to 4, and the proportion of invariant sites was set to 0.9 (rhodopsin) or 0.5 (mitochondrial, S7) and estimated. We used a strict clock (rate 1.0, not estimated) for the mitochondrial partition, a strict clock (rate 0.1, estimated) for rhodopsin, and a relaxed exponential clock (rate 0.4, estimated) for S7. Ploidy of the mitochondrial data was set to 0.5, and the nuclear partitions were left at the default value. The prior for collapse height was 0.001; *bdcGrowthRate* ($M = 4.6$, $S = 2$) and *popPriorScale* ($M = -7$, $S = 2$) were set as log normal; and other priors were left at their default settings. Each run was set for 50 million MCMC iterations, logged every 5000 iterations with a 10% burn-in. This resulted in ESS exceeding 200 for nearly all parameters of interest. We then analyzed the file of logged trees in SpeciesDA (<http://www.indriid.com/2014/speciesDA.jar>) to delineate candidate taxa. No single collapse height in SpeciesDA produced a plausible result across all species complexes, i.e., one set of candidate species observed in a majority of all runs. Consequently, we ran the analyses with collapse heights between 0.01 and 0.001 and chose the height and candidate species set associated with the most likely outcome in light of the results of previous analyses. Larger collapse heights resulted in species estimates of 1 for all species complexes, and smaller values led to no outcome representing more than 1% of observations. We used a 10% burn-in for these analyses and

left other parameters at default settings. As a secondary multilocus analysis, we built a maximum-likelihood phylogeny using the concatenated mitochondrial and nuclear genes, adopting the evolutionary models noted above. This approach is vulnerable to producing faulty evolutionary relationships (Bryant and Hahn 2020) but was expected to be valid for assessing bootstrap support and diagnosability of candidate taxa.

The final step in species delimitation was to propose a set of species hypotheses. When a plurality of methods concurred (because consensus among all methods was unlikely; Dellicour and Flot 2018) and did not violate the phylogenetic species concept (for either the mitochondrial phylogeny or the four-gene phylogeny), we designated these as candidate species. Candidate species represented by a singleton in multigene analyses were not treated as distinct unless other specimens in single-gene analyses confirmed that they were monophyletic (Lim et al. 2012). If a specimen from at or near the type location of a currently accepted species grouped with a candidate species, the candidate species was elevated to a provisional species with that name (Supplementary Material 1). The same approach was adopted for taxa that were formerly recognized but are presently synonymized, and for which the earlier names appeared valid. In either instance, we consider provisional species to be validly named taxa. Candidate species not meeting these criteria—i.e., those for which there were no previously described species as representatives, previous names were not available, or multiple names potentially applied to a candidate taxon—were left unnamed but appear to constitute legitimate species. Lineages delineated by PTP which met at least some species delineation criteria, including representing named taxa, were labeled as forms that may warrant consideration as units of conservation.

Next, we assigned specimens to the candidate or provisional species. Data for this analysis consisted of sequences for COI (4000 new sequences and 552 public sequences with at least 531 bases) and cytb (589 new sequences and 665 public sequences with at least 875 bases; Supplementary Material 2). We retained only those public sequences cataloged as a named species from western North America, including *C. bairdii* and *C.*

cognatus, and for which geographic locations could be inferred. We initially included all public sequences of *Cottus* species from eastern North America, Europe, and Asia as well as the Lake Baikalian cottoids in Abyssocottidae, Comephoridae, and Cottocomephoridae. Their inclusion, however, did not alter assignments among the western North American specimens, and they were excluded from subsequent analyses. To assign specimens to a candidate or provisional species, we built COI or cytb neighbor-joining trees (using the number of differences) and examined clade membership. Specimens grouping with a more comprehensively evaluated candidate or provisional species were assigned to that species. Specimens forming novel clades exclusive of candidate or provisional species were not assigned. Occasionally, candidate species were the sole members of more than one mitochondrial clade; hence, they remained diagnosable. Finally, we considered whether a sample was collected within the purported or plausible range of a candidate or provisional species and whether that range made geographic sense in light of modern or paleohydrological connections, because geography is often the most diagnostic characteristic of sculpin species (Jenkins and Burkhead 1994). This information was also used to re-evaluate and refine the initial designations of candidate and provisional species.

Because the standard practice among biologists is to identify specimens in the field without genetic information, we evaluated the success of this approach by comparing field and reference library identifications based on morphology to genetic assignments. We restricted comparisons to specimens identified as a currently named species for which there was genetic support, i.e., unambiguous genetic assignment to a form, provisional species, or species complex representing a currently valid taxon. For example, we regarded genetic and phenotypic identifications as concordant if a molecularly assigned individual of candidate or provisional species 13–15 or 17–22 (see below) was identified as *C. beldingii* in the field or in reference databases. If it was identified as a different species in the field, or if a field-identified specimen of *C. beldingii* was assigned to a different candidate or provisional species, the comparison was regarded as discordant.

TABLE 1. Summary statistics for the individual genes in the concatenated data set ($n = 433$ specimens). Abbreviations: S = segregating sites, H_d = haplotype diversity, P_i = nucleotide diversity, k = mean nucleotide differences, Iss = index of substitution saturation. Critical values for significance are in parentheses.

Gene	Bases	S	Parsimony-informative sites	Haplotypes	H_d	P_i	k	Iss by position		
								1 & 2	3	3
COI	621	192	174	259	0.995	0.0429	26.7	0.037 (0.364)	0.191 (0.373)	—
cytb	1036	374	343	316	0.998	0.0487	50.5	0.209 (0.397)	0.240 (0.356)	—
Rhodopsin	472	76	65	67	0.894	0.0228	10.8	—	—	—
S7	554 ^a	112	87	114	0.967	0.0123	6.0	—	—	—
S7 (gap codes)	577 ^b	131	102	137	0.973	0.0136	6.9	—	—	—

^aOnly the 489 sites without gaps were analyzed.

^bOnly the 508 sites without gaps were analyzed.

Characteristics of the gene regions we used restricted phylogenetic inference. The indels in S7 were phylogenetically informative and emphasized species boundaries (Nagy et al. 2012), but alignments based on different methods (e.g., PRANK, MUSCLE [Edgar 2004] and various algorithms in MAFFT [Katoh et al. 2019]) often differed substantially (data not shown) and required additional manual adjustments, leading to uncertainty about the evolutionary history reflected by these sequences (Morrison 2015). Our decision to use gap coding led to issues with achieving effective samples sizes in multi-species coalescent analyses when the gap-coded segment was treated as an independent partition, but when considered as part of the S7 partition, it created a synthetic gene sequence for which the appropriate substitution model was uncertain. We assumed that the latter would identify informative nodes in the phylogeny but preclude estimates of branch lengths. Likewise, phylogenetic trees from amino acid translation of COI and rhodopsin indicated that these regions may have been under strong selection and prone to selective sweeps (Hill et al. 2019, Hill 2020) in violation of the neutral models that form the basis for dating divergence. In light of these issues, coupled with the apparent hybrid origins of some taxa and an incomplete fossil record, we did not estimate divergence dates.

RESULTS

The concatenated four-gene data set contained 367 distinct mitochondrial haplotypes and 195 nuclear haplotypes, and 479 haplotypes overall. Although the specimens sequenced at all 4 genes were biased toward those with different COI haplotypes, sequences of cytb were more variable, longer, and produced a larger number of haplotypes (Table 1). Tests of saturation of both mitochondrial genes were not significant. The nuclear genes were less variable than either

Fig. 3. (See facing page, p. 317). Phylogenies recovered from maximum-likelihood analyses of sequences of *Cottus* ($n = 433$) from the four-gene data set, with independent trees for the nuclear and mitochondrial sequences. Colors denote the 8 primary species complexes. Ultrafast bootstrap support is noted for each complex (nuclear phylogeny) or component of each complex (mitochondrial phylogeny).

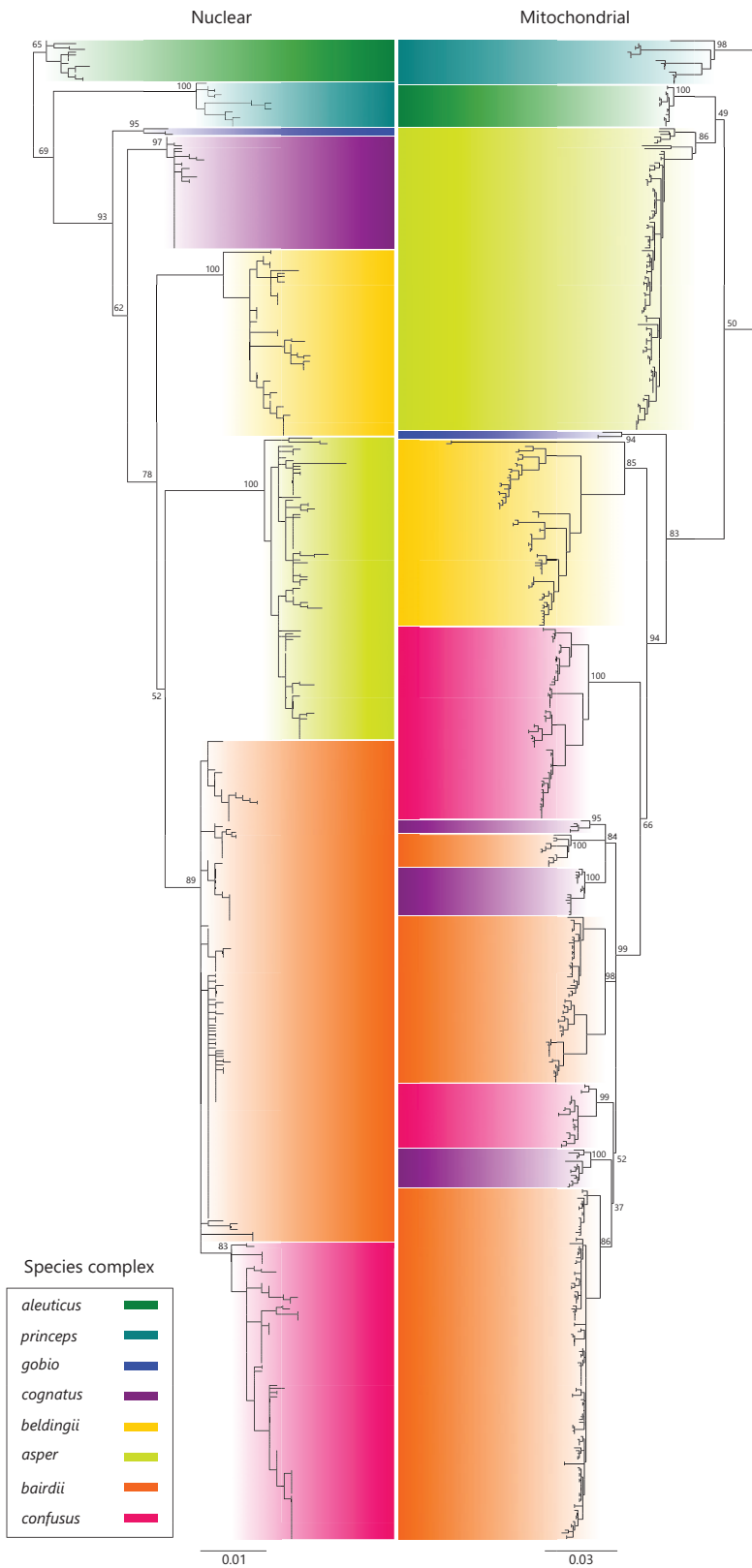


TABLE 2. Hybrid individuals or taxa among specimens collected from western North America, categorized by the species involved (nuclear \times mitochondrial). *Complex* denotes whether the specimen assigns to 1 or 2 nuclear complexes, and *form* denotes the mitochondrial form (see Supplementary Material 2). Abbreviations: WF = West Fork, EF = East Fork.

Specimen	Phenotype	Location	Category	Complex	Form
82	<i>Cottus</i> sp.	Deadman Creek, WA	<i>confusus</i> \times <i>semiscaber</i>	6	56
83	<i>Cottus</i> sp.	Deadman Creek, WA	<i>confusus</i> \times <i>semiscaber</i>	6	56
1237	<i>Cottus beldingii</i>	WF Jarbidge River, NV	<i>confusus</i> \times western <i>bairdii</i> ^a	6	59
1238	<i>Cottus beldingii</i>	WF Jarbidge River, NV	<i>confusus</i> \times western <i>bairdii</i>	6	59
1239	<i>Cottus beldingii</i>	WF Jarbidge River, NV	<i>confusus</i> \times western <i>bairdii</i>	6	59
1240	<i>Cottus beldingii</i>	WF Jarbidge River, NV	<i>confusus</i> \times western <i>bairdii</i>	6	59
1601	<i>Cottus rhotheus</i>	Minam River, OR	(<i>confusus</i> \times <i>semiscaber</i>) \times <i>semiscaber</i>	6, 8	56
1639	<i>Cottus beldingii</i>	Deer Creek, OR	(<i>rhotheus</i> \times <i>confusus</i>) \times <i>confusus</i>	6, 8	35
1721	<i>Cottus rhotheus</i>	McKenzie River, OR	(<i>rhotheus</i> \times <i>confusus</i>) \times <i>rhotheus</i>	6, 8	52
2186	<i>Cottus rhotheus</i>	Bumping River, WA	<i>semiscaber</i> \times <i>rhotheus</i>	8	54
2187	<i>Cottus rhotheus</i>	Bumping River, WA	(<i>semiscaber</i> \times <i>rhotheus</i>) \times <i>rhotheus</i>	8	54
3918	<i>Cottus</i> sp.	Catherine Creek, OR	(<i>confusus</i> \times <i>semiscaber</i>) \times <i>semiscaber</i>	6, 8	56
4550	<i>Cottus rhotheus</i>	Teanaway River, WA	(<i>semiscaber</i> \times <i>rhotheus</i>) \times <i>rhotheus</i>	8	54
4565	<i>Cottus rhotheus</i>	Mercer Creek, WA	(<i>semiscaber</i> \times <i>rhotheus</i>) \times <i>rhotheus</i>	8	54
5160	<i>Cottus</i> sp.	Anthony Creek, OR	<i>confusus</i> \times <i>semiscaber</i>	6	57
5180	<i>Cottus</i> sp.	North Powder River, OR	<i>confusus</i> \times western <i>bairdii</i>	6	59
6607	<i>Cottus rhotheus</i>	Grande Ronde River, OR	(<i>confusus</i> \times <i>semiscaber</i>) \times <i>semiscaber</i>	6, 8	56
6608	<i>Cottus rhotheus</i>	Grande Ronde River, OR	(<i>confusus</i> \times <i>semiscaber</i>) \times <i>semiscaber</i>	6, 8	56
6610	<i>Cottus rhotheus</i>	Grande Ronde River, OR	(<i>confusus</i> \times <i>semiscaber</i>) \times <i>semiscaber</i>	6, 8	56
6695	<i>Cottus</i> sp.	Flint Creek, ID	<i>confusus</i> \times <i>semiscaber</i>	6	58
6696	<i>Cottus</i> sp.	Flint Creek, ID	<i>confusus</i> \times <i>semiscaber</i>	6	58
6698	<i>Cottus</i> sp.	EF Jarbidge River, ID	<i>confusus</i> \times <i>semiscaber</i>	6	57
7670	<i>Cottus gulosus</i>	Bird Creek, CA	(<i>asper</i> \times <i>gulosus</i>) \times <i>asper</i>	3	17
7770	<i>Cottus cognatus</i>	Kettle Creek, PA	<i>cognatus</i> \times <i>bairdii</i>	7	49
7773	<i>Cottus</i> sp.	Lick Island Run, PA	<i>cognatus</i> \times <i>bairdii</i>	7	49
7775	<i>Cottus</i> sp.	Lick Island Run, PA	<i>cognatus</i> \times <i>bairdii</i>	7	49
7787	<i>Cottus</i> sp.	WF Jarbidge River, NV	<i>confusus</i> \times western <i>bairdii</i>	6	59
7988	<i>Cottus</i> sp.	EF Jarbidge River, NV	<i>confusus</i> \times <i>semiscaber</i>	6	57

^aHaplotype sharing in form 59 prevented assignment to a single species for the maternal parent of this cross.

mitochondrial gene. Although rhodopsin exhibited fewer haplotypes than did S7 (67 vs. 114), the haplotypes tended to differ by more nucleotides (k, 10.8 vs. 6.0).

The nuclear and mitochondrial phylogenies were not concordant (Fig. 3, Supplementary Material 4). The nuclear analyses delineated 8 monophyletic, readily recognizable species complexes (labeled by the first-diverging member) largely in accord with the current taxonomy of sculpins. Ordered by decreasing levels of divergence (with *Leptocottus armatus* as the outgroup), the *C. aleuticus* complex was basal to all other members, followed by complexes representing *C. princeps*, *C. gobio*, *C. cognatus*, *C. beldingii*, *C. asper*, *C. bairdii*, and *C. confusus*. The mitochondrial phylogeny identified several monophyletic groups that corresponded to those in the nuclear phylogeny, but the relationships among species complexes differed. In order of decreasing divergence from the outgroup were those represented by *C. princeps*, *C. aleuticus*, *C. asper*, *C. gobio*, *C. beldingii*,

C. confusus (in part), and a strongly supported clade of well-resolved taxa representing members of 3 nuclear complexes—*C. cognatus*, *C. bairdii*, and some members of *C. confusus*. The positions of complexes representing *C. asper*, *C. cognatus*, and *C. confusus* in one or both phylogenies invalidated the current concept of *Cottopsis* (hypothesized to consist of the *C. aleuticus*, *C. princeps*, and *C. asper* species complexes) and *Uranidea* (thought to be restricted to complexes of *C. cognatus* and *C. bairdii*) as subgenera.

Polyphyly in the terminal clade of the mitochondrial phylogeny was indicative of historical hybridization between members of the *C. bairdii* species complex and those of either the *C. cognatus* or *C. confusus* complexes, and also indicative of more recent introgression (Table 2), often with members of the *C. bairdii* complex as the maternal parent. For example, 3 specimens (7770, 7773, and 7775 from Kettle Creek and Lick Island Run, PA) in the vicinity of the type location for *C. bairdii* had nuclear

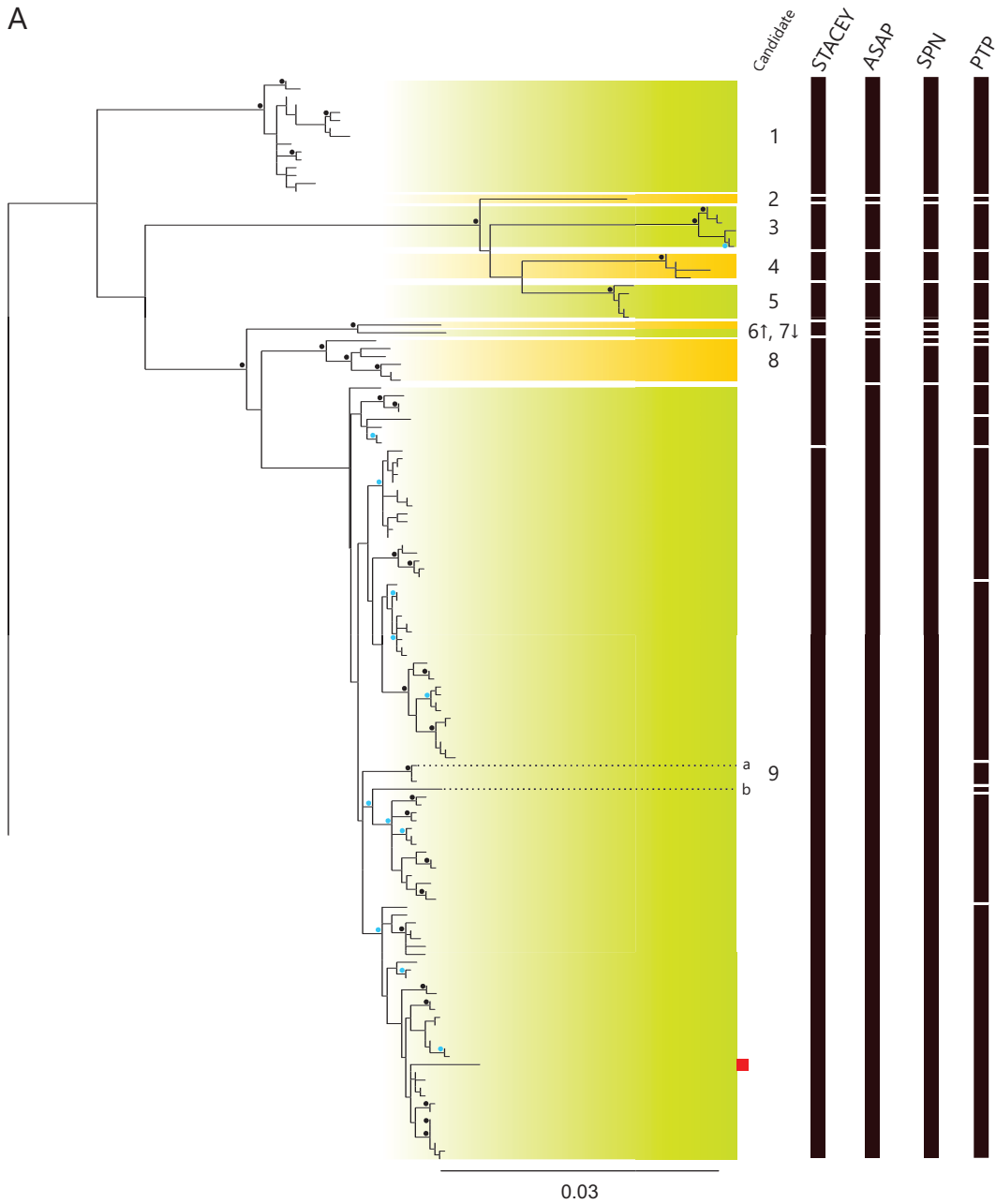


Fig. 4. Maximum-likelihood phylogeny of *Cottus* (panels A–D, pages 319–322) based on concatenated mitochondrial and nuclear sequences of the four-gene data set and the results of species delimitation analyses. Candidate and provisional species are numbered (labels are in Tables 3, Supplementary Material 2). Red rectangles denote lineages of hybrid origin and dots (blue, >90%; black, >95%) denote ultrafast bootstrap support. Line breaks in the columns denote divisions recognized in each species delimitation analysis, but some analyses pooled candidate taxa that were separated by a delimited taxon (see Table 4, Supplementary Material 2). A dotted line indicates a currently named or proposed taxon that was not delimited in these analyses: a, *C. marginatus*; b, unnamed species from Clear Lake, California; c, *C. bendirei*; d, *C. extensus*.

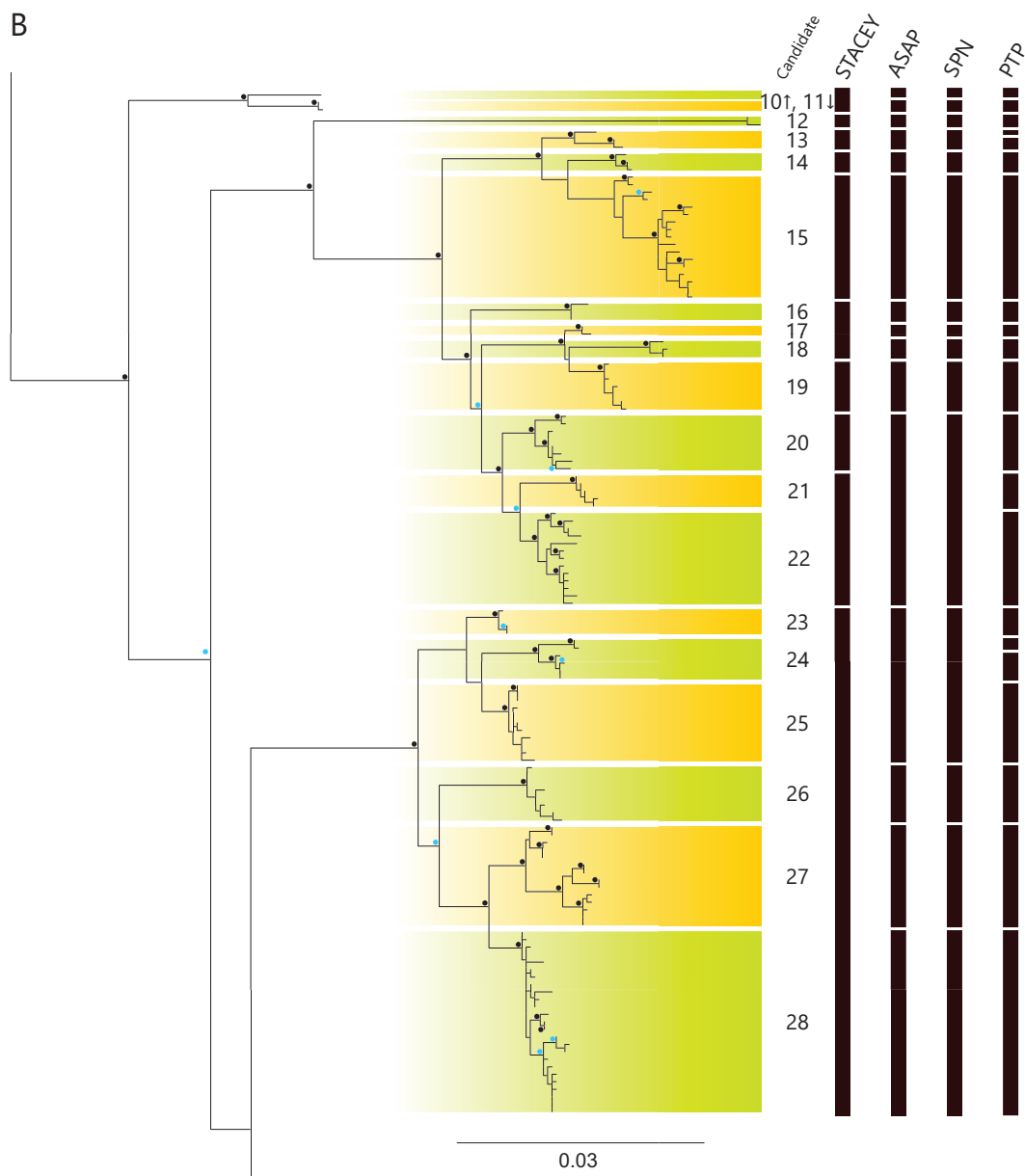


Fig. 4. Continued.

sequences of the *C. cognatus* complex but mitochondrial sequences identical (or nearly so) to topotypes of *C. bairdii* (7768, Allegheny River, PA, and GenBank accession AF549123, Coffee Run, PA; Fig. 4). Likewise, *C. schitsuumsh* was observed to have nuclear sequences of the *C. confusus* species complex but mitochondrial sequences more closely related (2.54%–2.66%) to western members of the *C. bairdii* complex

than to those of *C. confusus* (4.45%–5.04%). Left unresolved with respect to species delineation or assignment were specimens from the Jarbidge River in Idaho and Nevada to the Grande Ronde River in Oregon (and near the Kettle River in Washington) because these either tended to group with *C. schitsuumsh* found much farther north or had mitochondrial, and sometimes rhodopsin, sequences similar or

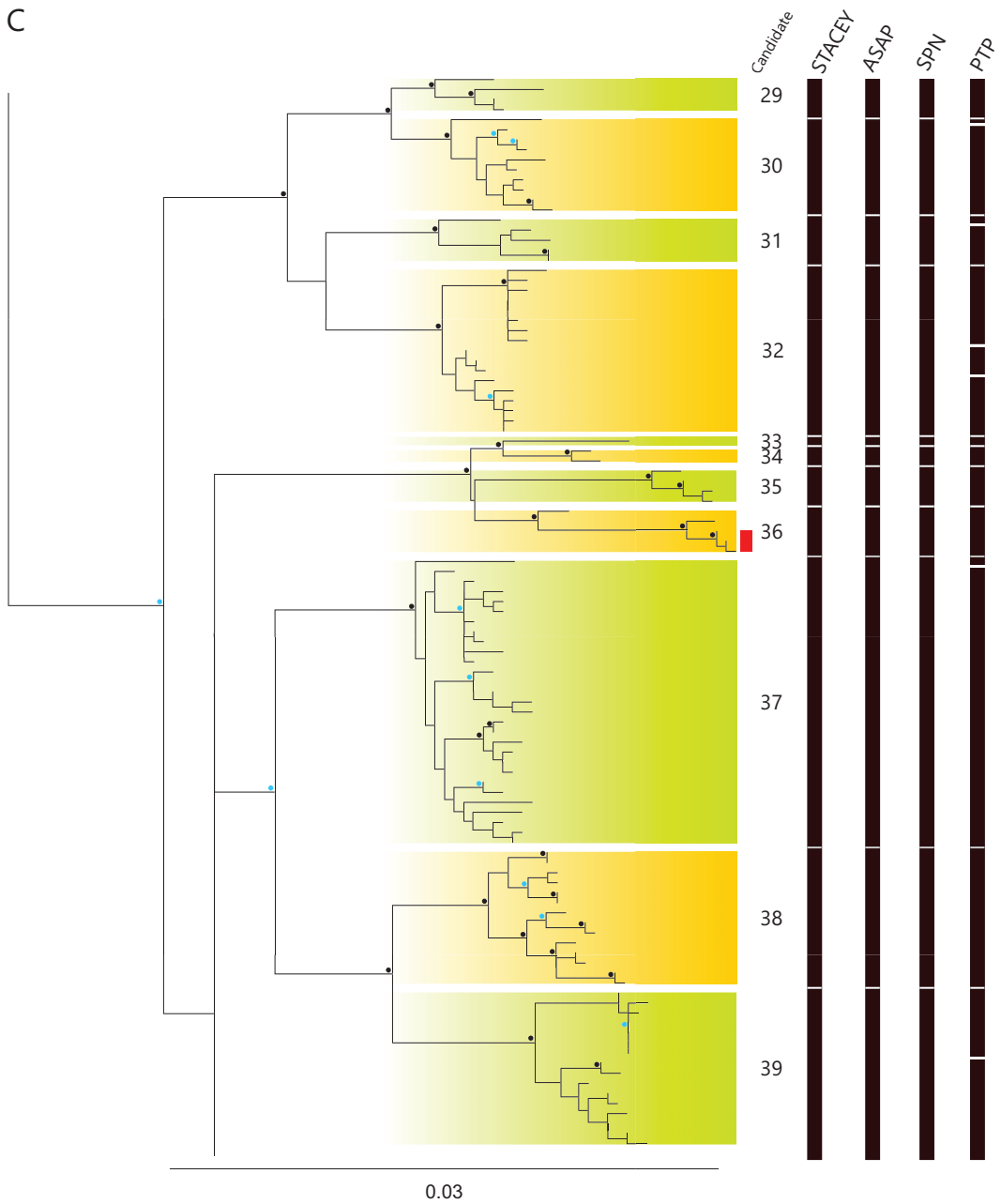


Fig. 4. Continued.

identical to *C. semiscaber* (see below), coupled with S7, and sometimes rhodopsin, sequences similar to *C. confusus*. Finally, some individuals appeared to be hybrids between members of the western *C. bairdii* complex.

There were 398 haplotypes among the 564 sequences in the concatenated mitochondrial

data set used in species delimitation. There was broad agreement among most methods with respect to the membership and number of candidate species. Counts were nearly identical between ASAP ($n = 35$) and SPN ($n = 36$), and almost all candidate species constituted strongly supported monophyletic clades

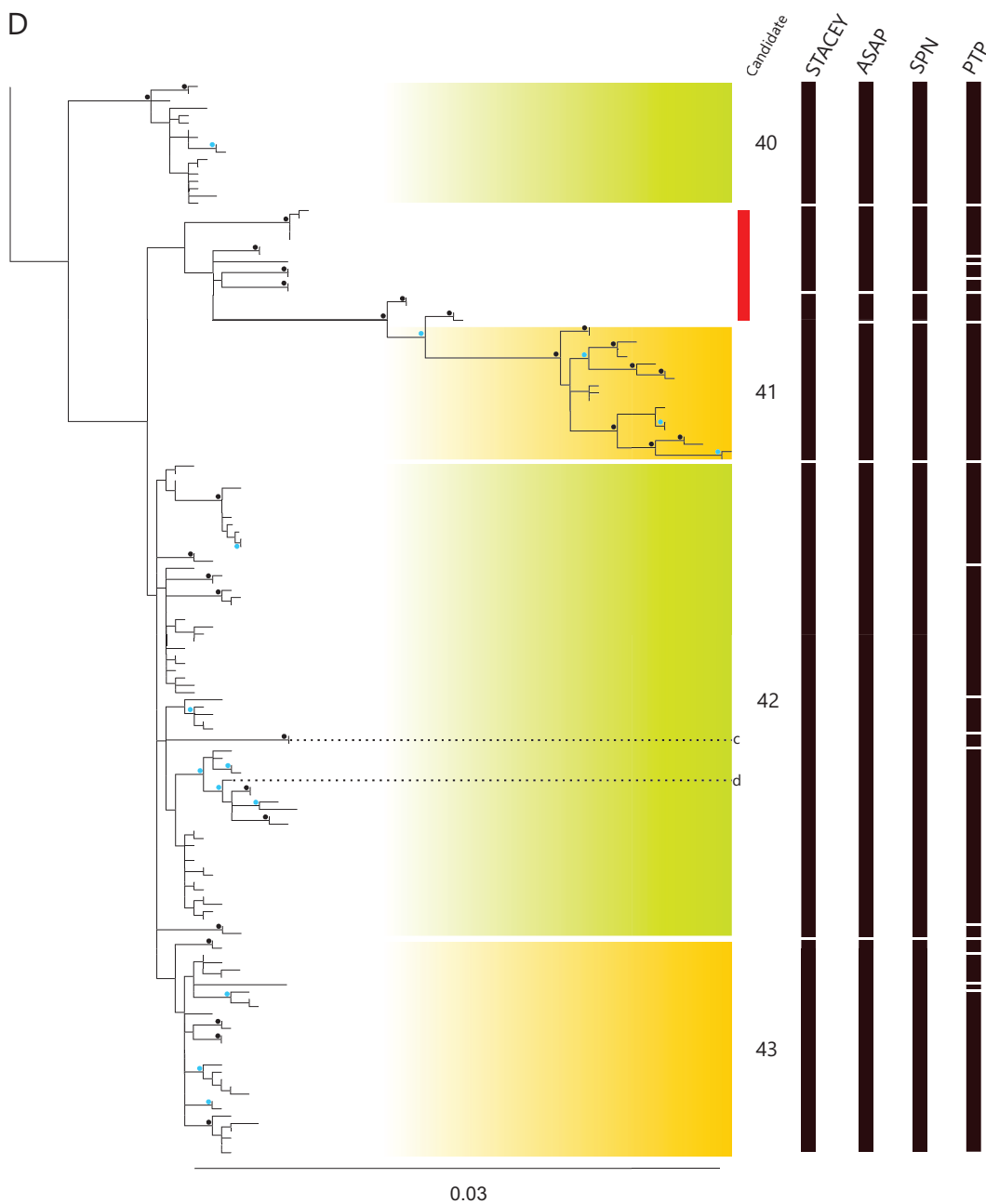


Fig. 4. Continued.

in the ML analysis and exhibited a barcode gap and differed from their nearest interspecific neighbor by $>1\%$ (Supplementary Material 2). These methods differed only in their treatment of a specimen (8194, West Fork Millcoma River, OR) as either a candidate species with an independent SPN or as part

of an ASAP group with other specimens from southern Oregon (Supplementary Material 2). Counts from PTP ($n = 66$) were about double those of other methods, and those candidate species often failed to exhibit bootstrap support, a barcode gap, or substantial divergence, but often highlighted geographically

TABLE 3. Candidate species (and provisional species where names were available) found in species delimitation analyses. Concatenated mitochondrial sequences were those used in species delimitation. Specimens assigned were all specimens from this study and public databases that could be assigned to a candidate or provisional species. Abbreviations: IS = intraspecific, NN = nearest neighbor, Y = yes, N = no. Under *species delimited*, results are in the order of STAGEY-ASAP-SPN-PTP, for which “Y” denotes that a species was delimited by that method. Monophyletic support reflects whether the candidate species was monophyletic in the analyses of the concatenated mitochondrial genes ($n = 564$ specimens) and the concatenated mitochondrial and nuclear genes ($n = 433$ specimens). Values >1 indicate ultrafast bootstrap support for the monophyletic clade represented by the candidate species; values = 1 indicate that the candidate species was monophyletic but represented by a singleton, and values <1 indicate that the candidate species was not monophyletic. Under geography, “Y” denotes that a candidate species was believed to be allopatric with respect to closely related members of its species complex, “N” that it was not, and “NA” that sampling was not sufficient for this evaluation.

Candidate	Species	Concatenated mitochondrial sequences			Assigned specimens	IS distance (max, %)	NN distance (min, %)	Species delimited	Monophyly	Geography
		Provisional	Count	Haplotypes						
1			8	8	143	0.66	3.14	Y-Y-Y	100-100	Y
2	<i>alenticus</i>		1	1	6	—	2.47	Y-Y-Y	1-1	N
3	<i>princeps</i>		4	4	9	0.66	3.38	Y-Y-Y	100-100	N
4	<i>asprerrimus</i>		5	4	9	0.48	2.11	Y-Y-Y	100-100	N
5	<i>tenis</i>		7	3	24	0.36	2.11	Y-Y-Y	100-100	N
6	<i>klamathensis</i>		2	2	125	0.18	1.99	N-X-Y	99-1	N
7	<i>pitensis</i>		1	1	83	—	1.99	N-X-Y	1-1	N
8	<i>gulosus</i>		6	6	16	1.33	2.35	N-X-Y	100-100	N
9	<i>asper</i>		77	62	1805	1.63	2.35	N-X-Y	100-100	N
10	<i>gobio</i>		1	1	2	—	1.51	N-X-Y	1-1	N
11	<i>ricet</i>		1	1	25	—	1.51	N-X-Y	1-1	Y
12	<i>greenei</i>		2	2	4	0.18	6.16	Y-Y-Y	100-100	Y
13			3	3	20	0.97	1.69	Y-Y-Y	100-100	Y
14			2	2	4	0.30	1.51	Y-Y-Y	100-100	Y
15	<i>tubulatus</i>		20	16	95	1.21	1.51	Y-Y-Y	100-100	Y
16	<i>leiopomus</i>		3	2	32	0.24	1.93	Y-Y-Y	100-100	Y
17			1	1	2	—	0.66	N-N-N-Y	1-1	Y
18			2	2	11	0.30	1.27	N-Y-Y	100-100	Y
19	<i>beldingii</i>		4	4	22	0.18	0.66	N-N-N-Y	100-100	Y
20	<i>anna</i>		8	6	80	0.78	1.03	Y-N-N-Y	100-100	Y
21			3	3	6	0.18	0.97	N-N-N-Y	100-100	Y
22			13	13	81	0.97	0.97	N-N-N-Y	92-100	Y
23			4	3	6	0.12	0.78	N-N-N-N	100-100	Y
24			5	4	18	0.60	1.03	N-N-N-Y	100-100	Y
25	<i>confusus</i>		10	8	71	0.48	0.78	N-N-N-N	100-99	Y
26			5	3	20	0.18	2.17	N-Y-Y	100-100	Y
27			15	9	58	1.09	1.21	N-X-Y	99-99	N
28			27	18	79	0.60	1.21	N-X-Y	98-98	N
29			3	3	52	0.91	0.84	Y-N-N-Y	99-99	Y
30			9	9	82	1.03	0.84	Y-N-N-Y	99-99	Y

TABLE 3. Continued.

Candidate	Species	Concatenated mitochondrial sequences			Assigned specimens	IS distance (max, %)	NN distance (min, %)	Species delimited	Monophyly	Geography
		Count	Haplotypes	Haplotypes						
31	<i>cognatus</i>	10	9	30	1.09	1.63	Y-Y-Y	94-98	N	
32		70	32	227	1.03	1.63	Y-Y-Y	100-100	N	
33	<i>philomips</i>	2	2	9	0.12	1.15	Y-Y-Y	100-1	NA	
34		2	2	8	0.24	1.03	Y-N-N-Y	100-100	NA	
35	<i>bairdii</i>	4	4	53	0.54	1.15	Y-Y-Y	100-100	NA	
36		1	1	16	—	1.03	Y-N-N-Y	100-99	NA	
37		23	22	298	1.15	1.93	Y-Y-Y	100-100	Y	
38		12	10	97	1.03	1.75	Y-Y-Y	100-100	Y	
39	<i>rhotheus</i>	21	12	48	1.21	1.75	Y-Y-Y	100-100	Y	
40		45	18	200	0.60	1.03	Y-N-N-Y	99-100	Y	
41	<i>schitsuumsh</i>	39	20	284	1.21	2.17	Y-Y-Y	100-100	N	
42	<i>semiscaber</i>	57	42	240	1.51	0.24	N-N-N-Y	0-0	Y	
43	<i>punctulatus</i>	26	20	269	0.91	0.24	N-N-N-Y	65-79	Y	

circumscribed lineages or currently described taxa not delineated by the other methods. Many candidate species identified by PTP represented subdivisions that, when viewed collectively, constituted highly supported groups in the other analyses. Counts derived from STACEY ($n = 32$) were the most conservative, and this method occasionally failed to delineate candidate species that were strongly supported in other analyses. In particular, it left the *C. confusus* complex as a single candidate species (but did delineate *C. schitsuumsh* of likely hybrid origin as distinct), whereas the other methods delineated 4–6 candidate taxa in this complex. Within most species complexes, however, there was near-consensus among methods with respect to the diversity of candidate taxa present. Using these estimates and considering additional information provided by monophyly, genetic distances, and geographic isolation, we propose 43 groups as candidate or provisional species (Table 3, Fig. 4; see discussion).

Most specimens for which there was only a single mitochondrial gene sequence assigned unambiguously to a candidate species and were useful for delineating the species' range. In the COI neighbor-joining phylogeny (Supplementary Material 5), candidate species 42 and 43 shared some haplotypes and were not diagnosable. Candidate species 31 and 32, and 34 through 36, though not monophyletic, were the sole members of separate monophyletic clades and were diagnosable. A group of specimens labeled as *C. bairdii* from Virginia and Maryland (GenBank accessions HQ557187, HQ557189; JN025004, JN025005, JN025010, JN025014, JN025016, JN025020) could not be assigned and likely represent a different species. The cytb phylogeny (Supplementary Material 6) offered greater resolution, because all candidate species and all but one form (form 61 in candidate species 42) were diagnosable. Three groups could not be assigned to a species in this analysis: specimens of *C. bairdii* from Virginia (GenBank accessions AF549125–AF549127) and Missouri (AF549162–AF549167, AY833333)—which are likely not that species—and the hybrid group associated with *C. confusus* from the middle Snake River. Only the specimens of candidate species 9 from the Russian River, California (JX484692–JX484693) could not be assigned to a previously recognized form.

TABLE 4. Comparison of morphological identification to genetic assignment. Phenotype is the name assigned in the field or to publicly cataloged specimens, and the candidate species (or forms) are those that we delimited. Abbreviations: +P/-G = specimens phenotyped as one species but genotyped (based on mitochondrial specimen identification) as a different species, +P/+G = specimens identified as the same species phenotypically and genotypically, -P/+G = specimens genotyped as this species but phenotyped as a different species. *Unidentified* indicates the number of specimens that genotyped as a species but were not identified in the field.

Phenotype	Candidate species	+P/-G	+P/+G	-P/+G	Unidentified
<i>aleuticus</i>	1	26	124	17	2
<i>princeps</i>	2	0	5	1	0
<i>asperrimus</i>	3	0	9	0	0
<i>tenuis</i>	4	0	9	0	0
<i>klamathensis</i>	5	0	21	0	0
<i>pitensis</i>	6	0	70	55	0
<i>gulosus</i>	7	208	83	0	0
<i>asper</i>	8, 9 (forms 10–13, 15–17)	13	366	582	815
<i>marginatus</i>	9 (form 14)	0	46	4	1
<i>perplexus</i>	9	381	0	0	0
<i>gobio</i>	10	0	2	0	0
<i>ricei</i>	11	0	25	0	0
<i>greenei</i>	12	0	3	0	1
<i>leiopomus</i>	16	0	18	9	5
<i>beldingii</i>	13–15, 17–22	12	106	70	145
<i>confusus</i>	23–28	81	26	21	205
<i>schitsuumsh</i>	41	0	15	2	267
<i>cognatus</i>	29–32	23	82	82	226
<i>bairdii</i> ^a	33–36	30	65	17	4
<i>rhotheus</i>	37–39	64	314	44	84
<i>punctulatus</i> ^b	40, 42, 43	49	244	29	436

^aAnalysis restricted to specimens from eastern North America.

^bAnalysis restricted to specimens from western North America currently valid as *C. bairdii*, *C. bairdii punctulatus*, *C. bairdii semiscaber*, *C. bendirei*, *C. extensus*, or *C. hubbsi*. Phenotypes and genotypes were concordant if specimens identified in the field or public databases with any of these names were genotyped as one of the specified candidate species.

In contrast, phenotypic and genotypic assignment of specimens was often discordant (Table 4). All phenotypic and genotypic assignments were in agreement for only 7 of 21 currently recognized species. For taxa such as *C. confusus*, most individuals morphologically identified as that species were genotypically identified as a different species. For taxa such as *C. asper* and *C. cognatus*, a large proportion of individuals genotypically assigned to each species were morphologically identified as something else. And for several taxa, most representatives were not identified in the field, indicative of the difficulty of recognizing species in the hand.

DISCUSSION

The biodiversity of western North American *Cottus* is a legacy of the evolutionary interplay among a dynamic continental hydrological network, broad-scale climatic patterns shifting with the arrival and departure of glaciation, and the demography, life history, and ecological specialization of individual species. Not all genes were equally informative in

revealing these patterns. Analyses of rhodopsin unambiguously delineated species complexes but offered little resolution of candidate taxa. Sequences of *S7* reinforced the species complex boundaries and occasionally delineated candidate taxa but may have altered the inferred evolutionary history of this gene as a consequence of the treatment of insertions and deletions. Sequences of *cytb* exhibited little haplotype sharing and resolved the majority of candidate and provisional taxa, whereas *COI*, though still informative about most species-level assignments, was more likely to exhibit incomplete lineage sorting and featured primarily silent mutations, with only 2 amino acids showing extensive variability in the sequence fragment we used. Collectively, however, and in concert with a near-complete inventory of most major hydrologic basins of western North America, data from these genes demonstrated that the sculpin fauna of the region is far more diverse than currently appreciated.

Below, we explore the taxonomic and phylogeographic implications of identifying

candidate species (hereafter CS) and provisional species (PS). We acknowledge that other gene regions, species concepts, or methods of species delineation may result in a different interpretation of taxonomic diversity, and that our insistence on monophyly among all species precludes recognition of what are otherwise valid taxa. Our results, though, are consistent with methods used to designate units of conservation regardless of whether these are recognized taxonomically (COSEWIC 2018, Fišer et al. 2018), and we invite further investigation to corroborate, modify, or refute the species hypotheses we pose. We also recognize the need to develop a chronology of evolutionary events and to resolve the problematic lineages that appeared to result from introgression, for which genomic tools will be necessary.

The Former Subgenus *Cottopsis*: the *Cottus aleuticus*, *princeps*, and *asper* Complexes

Our results indicate that *Cottopsis* is not a well-supported monophyletic group. In mitochondrial analyses, *C. princeps* is sister to all other *Cottus*, and branch support between the *C. aleuticus* and *C. asper* complexes is weak. In nuclear analyses, these same complexes are neither sister lineages nor monophyletic, and the *C. asper* complex forms a well-supported clade with members of *Uranidea*. Each species complex, however, constitutes a highly supported, monophyletic group with a distinctive evolutionary history.

Cottus aleuticus (PS 1, type location in the Aleutian Islands, Alaska), though not always correctly phenotypically identified, was resolved as a strongly supported, highly divergent, and monophyletic taxon distributed from California to Alaska in coastal streams that was delineated in all analysis but exhibited little phylogeographic structure (Fig. 5). Two specimens were also observed in the lower Willamette River basin, from which the species was thought to be extirpated (Bond 1963, Markle 2016). Nuclear phylogenies placed it as sister to all other species complexes in our analyses. In a more limited data set almost exclusive to the former *Cottopsis*, Baumsteiger et al. (2012) also placed this taxon as the basal group using nuclear gene sequences.

The *C. princeps* complex is strongly supported and highly divergent from all other

species complexes, and though its members—*C. princeps*, *C. asperrimus*, *C. tenuis*, and *C. klamathensis* (PS 2–5; Fig. 5B)—were often sympatric (Moyle 2002), all were unambiguously delineated as individual taxa in all analyses and almost always correctly identified in the field. Nuclear and mitochondrial analyses suggest that these taxa form a cohesive clade (cf. Baumsteiger et al. 2012) that may have diverged roughly simultaneously, making relationships among them unclear. For example, morphological comparisons align *C. asperrimus* with *C. tenuis* (Robins and Miller 1957, Moyle 2002), and amino acid sequences of rhodopsin (Supplementary Material 4) are identical between *C. princeps* and *C. klamathensis*, whereas mitochondrial sequences suggest a closer relationship between *C. tenuis* and *C. klamathensis*.

Species boundaries were more obscure within the *C. asper* species complex, which consisted of the currently recognized *C. pitensis*, *C. gulosus*, *C. asper*, and *C. perplexus*. The most divergent of these were representatives of *C. pitensis* from the headwaters of the Sacramento River basin (PS 6) and *C. gulosus* in the San Joaquin River basin (PS 7; Fig. 5C), with evidence of introgression between them in the mainstem Sacramento River that explains the discrepancies in field and molecular identification in *C. pitensis* (Baumsteiger et al. 2014). These are sister taxa and their phylogenetic placement was consistent with previous analyses (Baumsteiger et al. 2012), and we agree with those authors that neither species is found outside these basins.

Remaining members of this species complex were divided into 2 groups. The first was a widely distributed lineage (PS 9) that we regarded as *C. asper* sensu stricto (type location near the mouth of the Columbia River, Washington) present in coastal basins from southern California to southeastern Alaska and inland to California's Central Valley, Oregon's Willamette River, Washington's lower Snake River, the middle and upper Columbia River in Washington and British Columbia, and across the Continental Divide into the Peace River headwaters in Alberta (Fig. 5A, C). One specimen (7670, Bird Creek, CA) from the Pajaro River, a coastal basin over 100 km to the south of the San Francisco Delta, had nuclear sequences that grouped with *C. gulosus*. Baumsteiger et al. (2014) found similar

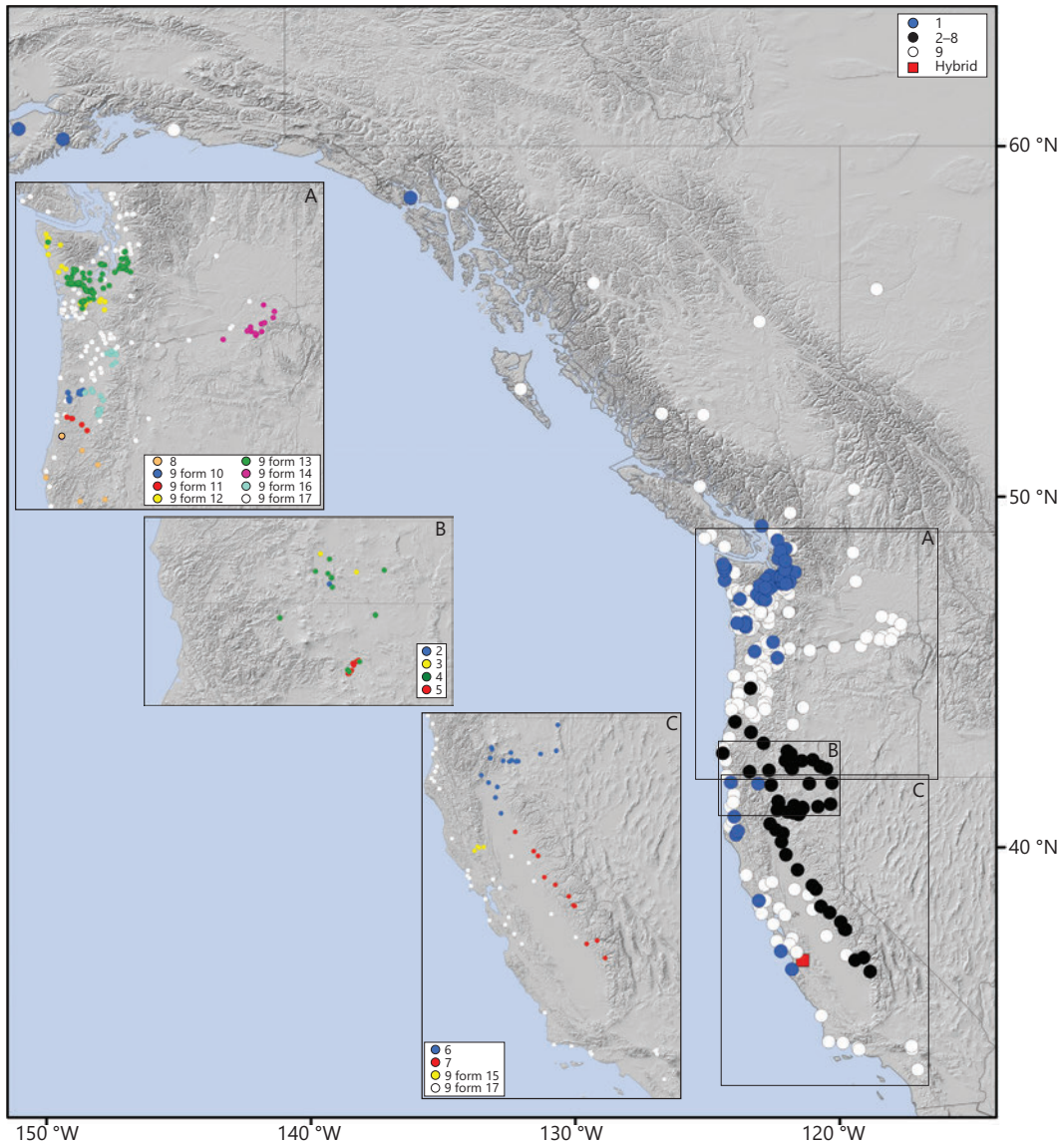


Fig. 5. Distribution of the *Cottus aleuticus* (PS 1), *C. princeps* (PS 2–5), and *C. asper* (CS/PS 6–9) species complexes. **Main map**, distribution of *C. aleuticus* (PS 1), *C. asper* (PS 9), and candidate and provisional species 2–8, plus 1 hybrid individual. **Insets: A**, distribution of CS 8 and forms of *C. asper* (PS 9), including *C. marginatus* (form 14); **B**, distribution of *C. princeps* (PS 2), *C. asperimus* (PS 3), *C. tenuis* (PS 4), and *C. klamathensis* (PS 5); **C**, distribution of *C. pitensis* (PS 6), *C. gulosus* (PS 7), and the Clear Lake (form 15) and widespread (form 17) forms of *C. asper*.

evidence of introgression with *C. gulosus* among other southern populations of *C. asper*. Whether these constitute a hybrid lineage or examples of recent hybridization is unknown.

The second group consisted of sometimes well-supported inland lineages generally restricted to one or a few basins that exhibited a latitudinal cline in divergence from *C. asper*

(Fig. 5A). One of these (CS 8) was distributed in coastal southern Oregon from the Rogue River to the Umpqua River. This lineage is sister to but divergent from the rest of the *C. asper* species complex (minimum distance, 1.09%), and specimens were labeled as *C. gulosus*, *C. perplexus* (including all those in public databases), or unknown, emphasizing

that collectors regarded them as morphologically distinct from *C. asper*. Further support for its recognition is that the freshwater fauna in this region has other endemic taxa (McPhail and Taylor 2009). One specimen (GenBank accession JX484688) in public databases assigning to this clade was reported to be collected from a tributary to the lower Willamette River basin. Other specimens from this tributary provided to us and in public databases assign to a different candidate species with strong support, so this record warrants additional review.

Delimiting other candidate species in the *C. asper* complex would have violated the phylogenetic species concept. Two of these are either recognized as or proposed to be distinct species (form 14, *C. marginatus* in northeastern Oregon and southeastern Washington; form 15, Clear Lake basin, California; Baumsteiger et al. 2016) and are monophyletic (within this clade), strongly supported, geographically isolated from all other clades of *C. asper* and may warrant such designation but would render *C. asper* paraphyletic in our data set. Farther north, although many lineages are still monophyletic and geographically restricted, their differences were more minor, and we regarded them as forms of *C. asper*. These include a group found in the Siuslaw River basin (form 11), with a single observation in the upper Umpqua River basin, and in the Alsea and Yaquina River basins (form 10), with additional collections directly to the east in the Marys River, a tributary to the Willamette River, where they are syntopic with a clade (form 16) found in the upper and lower Willamette River. Faunal exchanges between the Willamette River and these basins may explain these patterns (Markle 2019), but the extent of their overlap is unknown.

North of the Columbia River are 2 marginally distinct groups both found in many portions of the Chehalis River basin but also in coastal rivers in the northern Olympic Peninsula (form 12) or in Puget Sound (form 13). A single representative of the former was found in the Cowlitz River basin, a direct tributary to the Columbia River, and of the latter in the Bogachiel River along the Olympic Coast. Specimens from both of these forms were collected in the vicinity of the Skookumchuck River in the Chehalis River basin, the type

location for the problematic species hypothesis of *C. perplexus*. Field identification of both forms was roughly evenly divided between *C. perplexus* and *C. gulosus*. Because the latter does not occur outside California, the apparent ambiguity in differences between these species (Reimers and Bond 1967, Rowsey and Egge 2017) is instead evidence of morphological variability in putative *C. perplexus*.

We contend that *C. perplexus* does not constitute a valid taxon but represents a variant within the latitudinal cline of forms of *C. asper*. Viewed broadly, members of the *C. asper* species complex tend to exhibit 2 life history types. One of these is a widely distributed coastal lineage referable to *C. asper* that often exhibits amphidromy whereas the various inland forms, largely confined to one or a few major river basins, are usually partly sympatric with the widely distributed form but have an entirely freshwater life history and are morphologically distinct, albeit inconsistently (McAllister and Lindsey 1961, Krejsa 1967, Dennenmoser et al. 2015, Rowsey and Egge 2017). The more southerly or inland populations show greater amounts of molecular divergence from the widespread form, and have reached or are approaching thresholds associated with delimitation as separate species, e.g., *C. gulosus*, *C. pitensis*, *C. marginatus*, and candidate species in southern Oregon and the Clear Lake basin in California. Differences between coastal and inland groups farther north, especially near and in formerly glaciated areas such as the type location for *C. perplexus*, are subtle (Dennenmoser et al. 2014, 2015). Thus, *C. perplexus* appears to be a lineage in the process of diverging from *C. asper* that exhibits some morphological differences likely attributable to a wholly freshwater life history but has not yet crossed molecular thresholds for recognition as a distinct taxon.

Subgenus *Cottus*: the *Cottus gobio* Complex

We had few samples of members of this complex, but their assignment was straightforward. *Cottus ricei* (PS 11; Fig. 6) is restricted to North America from the McKenzie River basin in the north to the Saskatchewan River basin in the south, east to the Great Lakes and North Atlantic coast (Page and Burr 2011). In our analyses, it was sister to the other members of this complex from Eurasia,

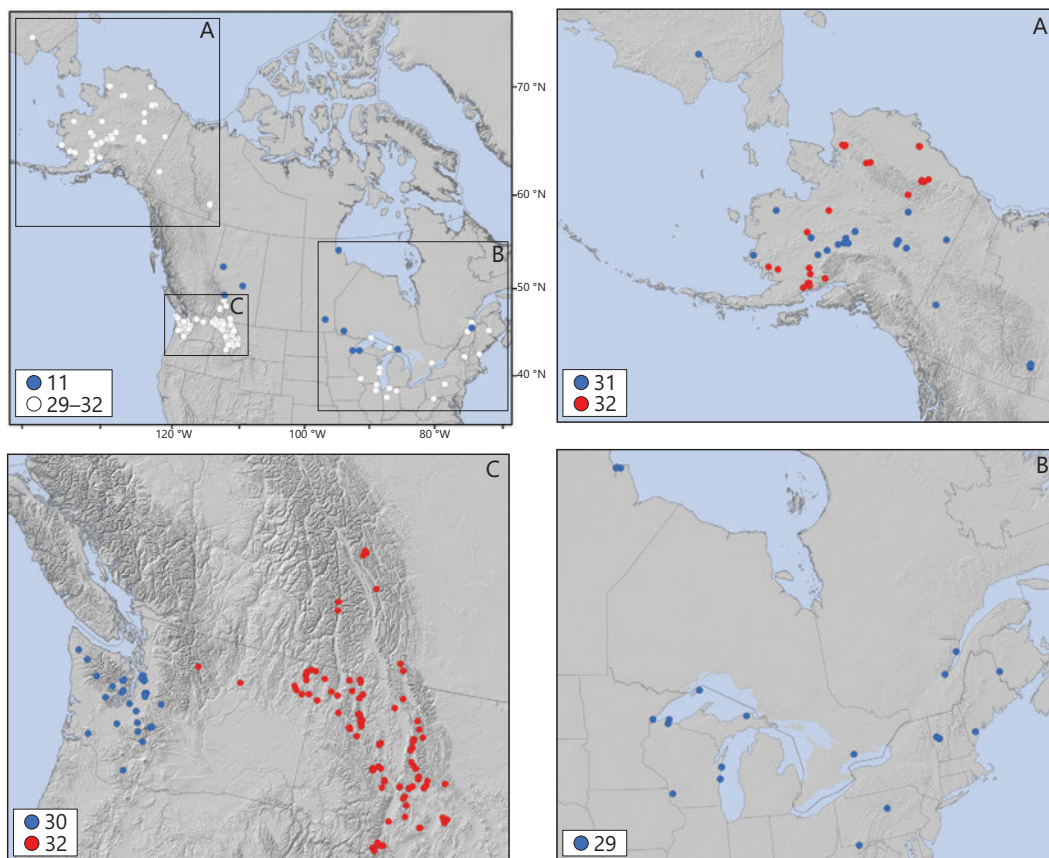


Fig. 6. Distribution of the *Cottus gobio* (in continental North America, represented by *C. ricei*, PS 11) and *C. cognatus* (CS/PS 29–32) species complexes. **Main map** (upper left), distribution of *C. ricei* (PS 11). **Insets:** **A**, distribution of *C. cognatus* (PS 31) and *C. philonips* (PS 32), in part; **B**, distribution of CS 30 and *C. philonips* (PS 32), in part; **C**, distribution of CS 29. The distribution of *C. gobio* (PS 10) is not shown.

including *C. gobio* (PS 10). Both species were supported as distinct by nearly all methods, but the treatment of *C. gobio* as representing one or many species remains unsettled (Freyhof et al. 2005, McLeish et al. 2020). The absence of *C. ricei* west of the McKenzie River basin and its relation (mean difference, 1.51%) to *C. gobio* suggests a derivation from the latter to the east, probably from populations in Greenland that were connected to North America at the onset or end of an early Pleistocene glacial interval (Dyke et al. 2003).

Subgenus *Uranidea* Part 1: the *Cottus beldingii* Complex

In previous work, members of this clade could not be assigned to a subgenus (Kinziger et al. 2005, Goto et al. 2015) because of

conflicting phylogenetic trees. In our analyses, this clade forms a highly supported group in maximum-likelihood and STACEY trees that could be considered the basal member of *Uranidea*. Likewise, we unambiguously resolved *C. greenei* (PS 12) as the most divergent member of this species complex (cf. Oh 2016), but with an enigmatic distribution (Fig. 7). It is restricted to springs adjacent to the Snake River in south-central Idaho that would have been deeply submerged during the flood triggered by the spillover of Lake Bonneville 17.4 kya (Oviatt 2015), suggesting it persisted in refugia above the floodwaters that it no longer occupies.

The present taxonomy features 2 other hypothesized species in this complex: *C. leiopomus* from the Wood River basin in Idaho

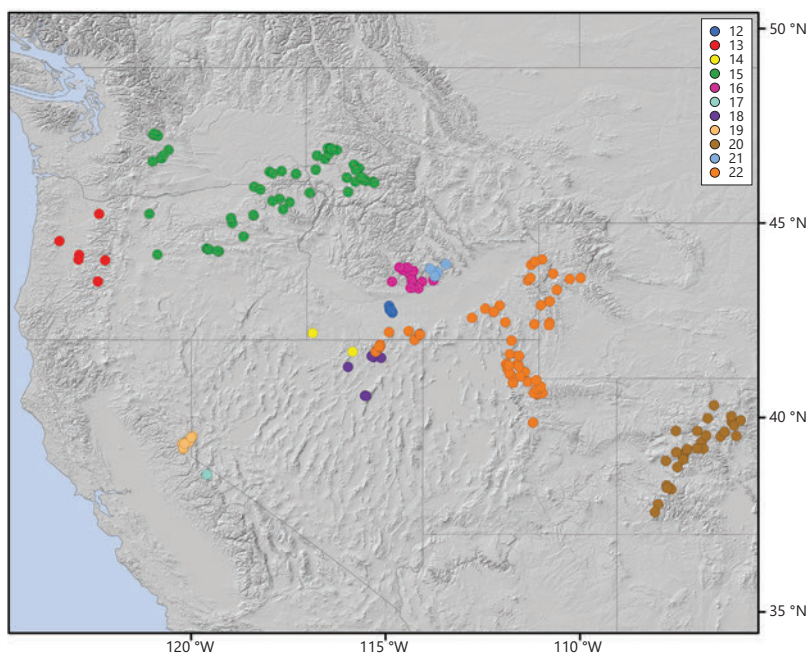


Fig. 7. Distribution of the *Cottus beldingii* species complex (CS/PS 12–22), including *C. greenii* (PS 12), *C. tubulatus* (PS 15), *C. leiopomus* (PS 16), and *C. beldingii* (PS 19).

and *C. beldingii* from the Colorado River basin (Colorado, Utah, and Wyoming), Columbia River basin (Idaho, Nevada, Oregon, Utah, Washington, and Wyoming), and Great Basin (Nevada, Oregon, and Utah). Our analyses support the current conception of *C. leiopomus* (PS 16), but we regard *C. beldingii* as a suite of 9 divergent, highly supported candidate taxa that are allopatric with respect to one another.

Three of these are from the Lahontan Basin: (1) *C. beldingii* sensu stricto (PS 19), found in many locations in the Truckee River basin (the outflow of the type location, Lake Tahoe) in California and Nevada; (2) a candidate species (CS 17) farther south in Silver King Creek, California; and (3) a candidate species (CS 18) in tributaries of the Humboldt River in northeastern Nevada. The roughly equivalent branch lengths among these well-supported groups indicate that they diverged from a common ancestor at approximately the same time, a pattern similar to that of recognized or proposed lineages of cutthroat trout—*O. c. lahontan* in the western Lahontan Basin, *O. c. seleniris* in Silver King Creek, and *O. c. humboldtensis* in the northeastern

Lahontan Basin (Trotter and Behnke 2008, Saglam et al. 2017).

A separate provisional species (PS 20) occupied the Dolores, Gunnison, and upper Colorado Rivers in Colorado. This range includes the type location for *C. annae*, the Eagle River near Gypsum, Colorado, in the headwaters of the Colorado River. Although the present taxonomy regards this taxon as a synonym of *C. beldingii*, the molecular results and its geographical isolation favor resurrecting *C. annae* as its name. Of note is that the distribution of *C. annae* is equivalent to that of the “blue” lineage of Colorado River cutthroat trout *O. c. pleuriticus* (Bestgen et al. 2019). Because both are also absent from other portions of the Colorado River basin with suitable habitat, such as the Duchesne River in Utah, the San Juan River in New Mexico, or the Green River in Colorado and Wyoming, their distribution implies that both taxa invaded at the same place and time and in a way that differed from that of 2 other candidate species distributed throughout the Colorado River basin, *C. punctulatus* (see below) and the “green” lineage of Colorado River cutthroat trout (Bestgen et al. 2019).

Sister to *C. annae* were 2 candidate taxa in the upper Snake River basin and northern Bonneville Basin. One of these (CS 21) was represented by specimens from the Big Lost River, an endorheic basin in south-central Idaho that was occasionally linked to the upper Snake River basin by Pluvial Lake Terreton during the Pleistocene (Gianniny et al. 2002). Mountain whitefish *Prosopium williamsoni* (Miller 2006) in the Sinks drainage are also derived from those in the upper Snake River and are similarly restricted to this basin, a distribution likely due to isolation of the Big Lost River by the Crater Butte lava flow 300 kya (Kuntz 2003). The final candidate species (CS 22) was found in the Snake River from Salmon Falls Creek (downstream from the traditional boundary of the upper Snake River, Shoshone Falls) to the headwaters in Wyoming, and in the northern Bonneville Basin in Utah from the Provo River north to the Bear River, with a disjunct appearance to the south in the Spanish Fork River basin. These locations were hydrologically linked at various times throughout the Pleistocene via re-routing of the Bear River between the Snake River and the Bonneville Basin (most recently 60 kya; Pedersen et al. 2016) or other trans-basin corridors (Eaton et al. 2018). A single collection in the headwaters of the Wind River in Wyoming, a tributary to the Missouri River basin, likely resulted from introductions by anglers (Joe Deromedi, Wyoming Game and Fish Department, personal communication).

A clade sister to *C. leiopomus* and the present and former *C. beldingii* consisted of 3 candidate species found downstream in the Columbia River basin. The first of these (CS 13) is in the Willamette River basin in Oregon, and the second (CS 14) was found in the Owyhee River in Idaho and Nevada, and is likely present downstream in Oregon (Bond 1963). The third (PS 15) was broadly but sporadically distributed in the Clearwater River basin in Idaho, most river basins in northern and northeastern Oregon, and the Yakima River basin in Washington. Hubbs and Schultz (1932) described *C. tubulatus* from the Potalch River basin in Idaho, which they mistakenly thought was part of the North Fork Clearwater River basin (Maughan 1974). As with *C. annae*, Bailey and Bond (1963) synonymized this species with *C. beldingii* with-

out providing justification. Here, we resurrect *C. tubulatus* as its valid name based on the grounds that it is a highly supported, divergent, and monophyletic taxon isolated from all other members of this species complex.

Subgenus *Uranidea* Part 2:
the *Cottus cognatus* Complex

Cottus cognatus, based on specimens from Great Bear Lake in the McKenzie River basin, was the first North American sculpin to be described (Richardson 1836). Although thought to be closely related to *C. bairdii* (McAllister and Lindsey 1961, Bailey and Bond 1963, Jenkins and Burkhead 1994), these 2 species complexes harbor deeply divergent nuclear sequences and are not sister lineages. The relatively close relationship among their mitochondrial phylogenies implies that introgression and mitochondrial swamping may have contributed to speciation, as well as the challenge of morphologically diagnosing them (McAllister 1964).

Furthermore, the hypothesis that *C. cognatus* represents a single taxon warrants reevaluation because we delineated 4 well-supported and divergent candidate or provisional species. The first (PS 31; Fig. 6A), which we designated as *C. cognatus* sensu stricto, was present throughout the Yukon and Kuskokwim River basins and at one location in the upper McKenzie River basin. Species occurring in the Yukon and McKenzie River systems often share a close relationship, likely facilitated by redirection of tributaries caused by the Laurentide ice sheet (Bodaly and Lindsey 1977, Duk-Rodkin et al. 2004). This candidate species also appears in the Anadyr River basin in Siberia (GenBank accession AY116364), indicating the presence of freshwater habitats that permitted dispersal across the Bering land bridge (Lindsey and McPhail 1986).

A second provisional species (PS 32; Fig. 6A,C) was found in the interior Columbia River basin in the northern United States and southern Canada, and its southernmost distribution coincided with the limit of the Cordilleran ice sheet and its periglacial lakes during the Last Glacial Maximum. Its northernmost distribution included Alaska primarily south of the Alaska Range and north of the Brooks Range, where it likely persisted in glacial refugia north and west of continental

ice (Moore et al. 2015). This northern range disjunction may also reflect the existence of a Chukchi Sea River that linked northern and southern Alaska (Lindsey and McPhail 1986). Occasional specimens also overlapped with the previous candidate species in the Yukon River basin. Because the bulk of the diversity of this lineage is in Alaska and the southern populations form a star phylogeny, we speculate that a southern Alaskan lineage colonized southward via periglacial lakes at the leading edge of the Cordilleran ice sheet, followed by northward recolonization when the ice sheet receded (McPhail and Lindsey 1986, Bernatchez and Wilson 1998).

Eigenmann and Eigenmann (1892) described *C. philonips* from Kickinghorse River in the headwaters of the Columbia River in British Columbia. This species was synonymized with *C. cognatus* by Hubbs and Schultz (1932), who claimed that the original describers had intended this name to represent a third species that had already been assigned a name, therefore rendering *C. philonips* unavailable. We interpret the original description differently (as do Fricke et al. 2021). Likewise, McAllister and Lindsey (1961) regarded *C. philonips* as available but still synonymized it with *C. cognatus*, despite observing nearly diagnostic morphological differences among specimens from 3 groups: the Columbia (which would include the Kickinghorse River) and Fraser Rivers, the McKenzie (including Great Bear Lake) and Yukon Rivers, and basins in eastern North America. These 3 groups correspond with the geographic boundaries suggested by the molecular analyses (cf. McPhail 2007). Because our samples from the Kickinghorse River are affiliated with the second candidate species, and based on the strong phylogenetic support for this clade, its unique distribution, and its potential morphological distinctiveness, we designate this provisional species as *C. philonips*.

A candidate species (CS 29; Fig. 6B) occupied portions of eastern North America, but we opted not to name it because our sampling was not intended to evaluate taxonomic diversity in that region and up to 11 synonyms might be available (McAllister and Lindsey 1961). It was sister, unexpectedly, to a fourth candidate species (CS 30; Fig. 6C) present in northwestern Washington. Despite that members of this candidate species were identified

as 7 different species by collectors, genetic analyses unambiguously assigned all of them to this taxon. Furthermore, we consider records of *C. confusus* from northwestern Washington (Lee et al. 1980, Wydoski and Whitney 2003, Page and Burr 2011) to be this species because we observed no members of the *C. confusus* species complex in this area. Morphologically, this species is distinct from *C. philonips* in eastern Washington (Wydoski and Whitney 2003). Its distribution in western Washington in and north of the Chehalis River along the Olympic Peninsula and along Puget Sound to the Snohomish River is consistent with colonization from the Chehalis Refugium (McPhail 1967) and with phylogeographic patterns in other taxa (*Novumbra hubbsi*, DeHaan et al. 2014; *Pacifastacus leniusculus*, Larson et al. 2012; CS 9 form 13, this study). Its mitochondrial divergence (mean, 1.31%) from eastern North American members suggests that this split long preceded the most recent glaciation, but how this form originally reached this location, and how it relates to other lineages in eastern North America, is unknown.

Subgenus *Uranidea* Part 3: the *Cottus bairdii* Complex

All members of this complex in eastern and western North America share a highly supported (bootstrap support 97) and invariant amino acid sequence in rhodopsin (Supplementary Material 4) that is diagnostic. Nevertheless, this group has diverged into 3 broad sets of candidate taxa: those related to *C. bairdii* in eastern North America, those related to *C. rhotheus* in the Columbia River basin (and likely more northerly basins in British Columbia; McPhail 2007), and those constituting a third group with no consistent designation but broadly distributed in the Missouri, Colorado, and Columbia Rivers and the Great Basin (Neely 2003).

The nonintrogressed members in eastern North America (Fig. 8) are divergent from one another and from all western North American members of this species complex in both mitochondrial and nuclear phylogenies, and we concur with Neely (2003) that *C. bairdii* is not present in western North America. Because our sample from the Allegheny River, Pennsylvania, was the closest to the type location of *C. bairdii* (the Mahoning River, OH,

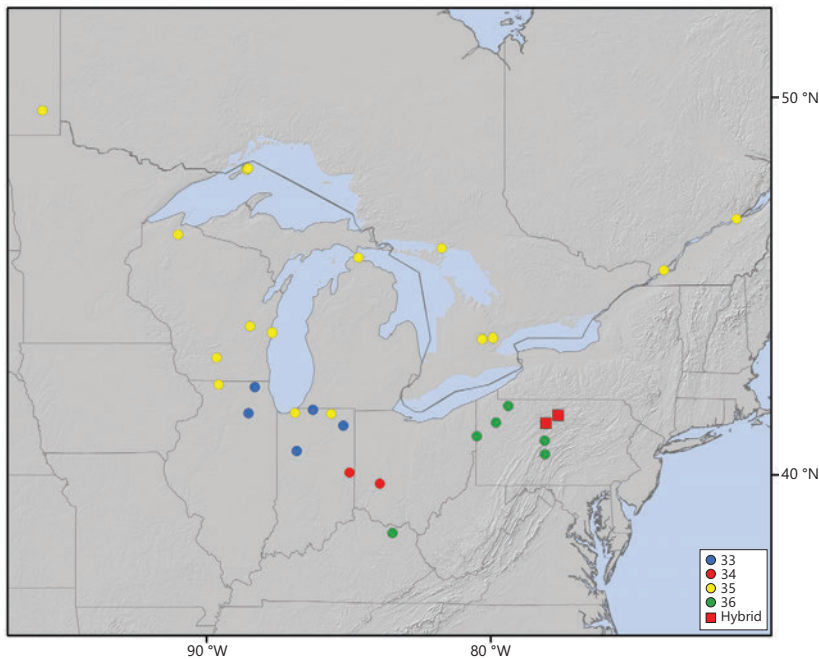


Fig. 8. Distribution of eastern North American members (CS/PS 33–36) of the *Cottus bairdii* species complex, including *C. bairdii* (PS 36) and hybrid individuals.

from which it is thought to be extinct; Neely 2003) and part of a strongly supported and divergent clade (PS 36), we consider these to be *C. bairdii* sensu stricto. This designation, however, is problematic for specimens outside the Ohio River basin. Notably, sculpins from 2 streams (Kettle Creek and Lick Island Run, Pennsylvania) in the headwaters of the Susquehanna River basin that assigned to this provisional species based on their mitochondrial sequences were late-generation hybrids or products of mitochondrial swamping because their nuclear sequences were weakly divergent from those of the *C. cognatus* complex. Moreover, several other specimens in public databases (GenBank accessions JN025091–JN025100), also from the Susquehanna River basin, have mitochondrial haplotypes of *C. bairdii* but were identified in the field as *C. cognatus*. More thorough assessments of the extent of recent or ancient introgression in this region are warranted.

Two candidate species, CS 34 (Great Miami and Whitewater River basins in the lower Ohio River basin) and CS 33 (Wabash and Fox River basins in the lower Ohio and upper Mississippi River basins) are sister taxa that may

have occupied separate refugia during glacial advances. A third, CS 35 (Great Lakes and St. Lawrence River basin, Winnipeg River basin, and upper Fox River basin in the Mississippi River basin), is broadly distributed, shows little variation, and is restricted to formerly glaciated areas, a pattern consistent with many other recent postglacial colonists (Bernatchez and Wilson 1998). Robins (1954) recognized *C. bairdii kumlieni* and *C. b. bairdii* from this region, essentially north and south of the southern Great Lakes. Although the former may correspond with our more northerly candidate species, the latter fails to accommodate the cryptic diversity observed in this area. Given this uncertainty, our limited number of specimens, and the many formerly recognized species from this region that are currently synonymized with *C. bairdii*, we did not assign names to these 3 candidate taxa.

Another group of 2 candidate taxa and 1 provisional taxon, strongly supported and monophyletic in the overall and mitochondrial phylogenies, was consistently identified in the field as *C. rhotheus* (Fig. 9). *Cottus rhotheus* sensu stricto (PS 39) was primarily found in the Kootenai and Spokane River basins (the

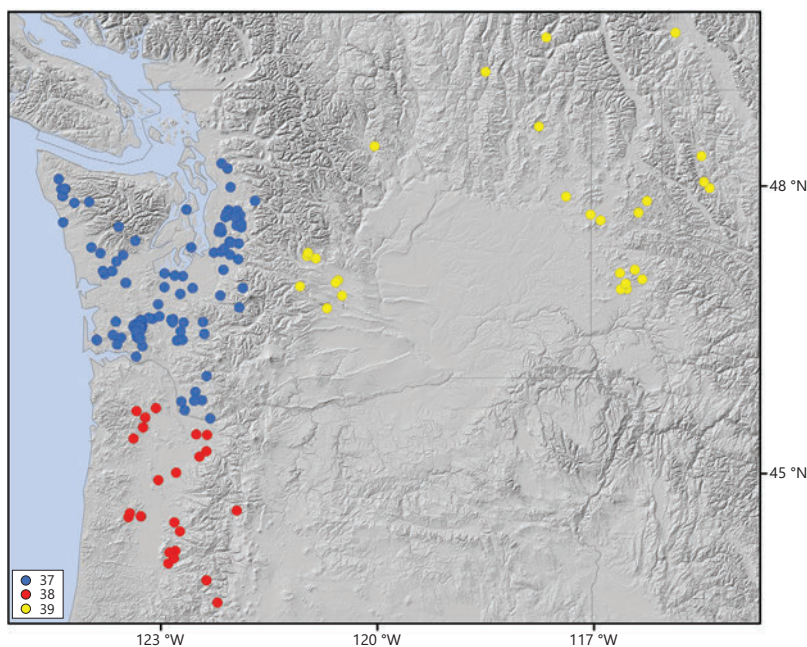


Fig. 9. Distribution of western North American members (CS/PS 37–39), in part, of the *Cottus bairdii* species complex, including *C. rhotheus* (CS 39).

latter the type location) and other portions of the interior Columbia River basin. We delimited an additional candidate species (CS 38) from many eastern and western tributaries of Oregon's Willamette River basin, and another (CS 37) that was restricted to Washington west of the Cascade Range. Coupled with this unexpected diversity in *C. rhotheus* sensu lato was morphological uncertainty. Specimens identified in the field as *C. rhotheus* from certain basins, e.g., the Clearwater River in Idaho and the Grande Ronde River in Oregon, were members of other clades, which may imply that *C. rhotheus* and its relatives have a smaller distribution than is currently believed (Wallace and Zaroban 2013, Markle 2016).

Long recognized as problematic were specimens that have often been identified as *C. bairdii* in western North America (Fig. 10). Uncertainty about their affiliations has led to a shifting taxonomy, with members variously assigned to *C. bairdii*, *C. bendirei*, *C. hubbsi*, *C. punctulatus*, *C. semiscaber*, or subspecific combinations thereof, and to 2 members of limited distribution, the extant *C. extensus* and putatively extinct *C. echinatus*. The ambiguity was also evident among the various methods

of species delimitation, for which there was little consensus on the divisions among forms and candidate taxa because of a lack of substantial and spatially coherent divergence despite their broad distribution in separate major hydrologic basins.

The candidate species with the greatest support (CS 40) was monophyletic in all gene trees and found in the upper Missouri River basin in Montana and Alberta (and likely Wyoming; Baxter and Stone 1995), suggestive of an arrival via a mid-Pleistocene Missouri River that drained to Hudson Bay (Howard 1958). It was found west of the Continental Divide throughout the North Fork Flathead River basin in Montana and British Columbia (where it was originally thought to be *C. confusus*; Peden et al. 1989) and sporadically in Montana's Swan, Blackfoot, and Clark Fork River basins. Surprisingly, members also occupied the Medicine Lodge and Beaver Creek drainages, part of the Sinks basins in east-central Idaho. It has been regarded as a distinct species on morphological, genetic, and geographical grounds (Neely 2003, McPhail 2007) and is treated as a unit of conservation in Canada (COSEWIC 2019). Its transdivide

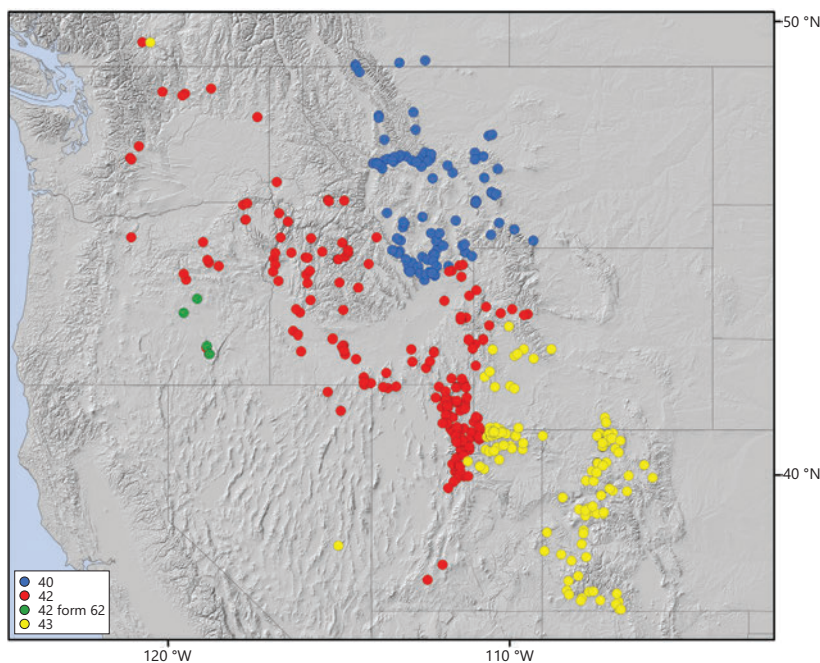


Fig. 10. Distribution of western North American members (CS/PS 40, 42, 43), in part, of the *Cottus bairdii* species complex, including *C. semiscaber* (PS 42), *C. bendirei* (CS 42 form 62), and *C. punctulatus* (PS 43).

distribution may be attributable in part to recent hydrological connections in the upper Flathead River (via Summit Lake at the headwaters of the Middle Fork Flathead and Two Medicine Rivers; Young et al. 2018), as well as its use as a baitfish in Montana until the late 20th century. Translocation by anglers may also explain its presence in the 2 isolated Sinks drainages in Idaho, but more intensive genomic inventories are required to resolve that question.

Assignment of candidate taxa to or among the remaining members of the western former *C. bairdii* is tenuous because most delimitation methods failed to recognize any divisions and only one group forms a weakly supported monophyletic clade that is geographically isolated. With respect to the latter, Gill (1862) first described a sculpin from the Colorado River basin as *Potamocottus punctulatus*, which was collected between Bridger Pass and Fort Bridger, Wyoming, and thus likely from the Little Snake or Green River basin. This group (PS 43) is monophyletic in the cytb and concatenated gene trees and was delimited in STACEY but displays more divergence among groups in the basin (0.91%) than with other

members of the species complex in other basins (0.24%). In part, the large intraspecific variation is attributable to inclusion of specimens from Butterfield Springs, Nevada, an isolated population in the pluvial White River that was once part of the lower Colorado River basin (Minckley et al. 1986).

Neely (2003) argued that specimens from the Colorado River basin should be recognized as *C. punctulatus*, morphologically diagnosing this species based on its lack of skin prickling. Gill (1876), however, released a more comprehensive description of the type specimen and indicated that prickles were likely present behind the pectoral fins. Gilbert and Evermann (1894) remarked that some specimens from the upper Snake River basin (originally described as *Cottopsis semiscaber*; Cope 1872) were smooth skinned and could not be reliably distinguished from those in the Green River basin. Nevertheless, because it is genetically and geographically circumscribed, we accept *C. punctulatus* as a provisional species.

Criteria for delimiting additional species were more difficult to meet for the remaining groups. *Cottus bendirei* (form 62) was originally

described from tributaries to Malheur Lake (Bean 1881), a closed basin in Oregon periodically connected in the Pleistocene to the Snake River or John Day River (Bisson and Bond 1971). Although Markle and Hill (2000) argued that this species (based on diagnosing individuals with respect to the extent of prickling) had a more extensive range within the Columbia River basin, we found a strongly supported monophyletic clade from the original range to be among the most highly diverged with this species complex. It was, however, sympatric with specimens currently assigned to *C. hubbsi* (form 63). For that reason and to preserve monophyly within the remainder of specimens in this lineage, we did not elevate it to a candidate species.

There was little support for further subdivision of the remaining members of this species complex, which tended to be neither monophyletic nor geographically discrete. Similarly, Oh (2016) did not retrieve specimens of this group from the Snake River, Columbia River, and Bonneville Basin as reciprocally monophyletic. Mussmann (2018) observed a similar pattern of limited divergence among speckled dace *Rhinichthys osculus* from throughout the northern Bonneville Basin and Snake and Columbia River basins, implying that similar paleohydrologic patterns may have led to the lack of phylogenetic structure in both fishes. Of note is that the distribution of this group of sculpins implies that they may have been among the last to colonize the Columbia River basin because they are largely absent upstream of waterfalls on the Palouse, Spokane, Pend Oreille, or Okanogan Rivers, which were crossed by a host of other freshwater fishes including other species of sculpins (McPhail and Lindsey 1986, Minckley et al. 1986).

Members of this group in most of the interior Columbia River basin (and some of those in the endorheic Harney Basin in Oregon) are currently valid as *C. hubbsi* (Bailey and Dimick 1949, Page and Burr 2011), with a purported upstream distribution to Shoshone Falls in the Snake River, and to major falls on other rivers in the upper Columbia River basin. Those in the upper Snake River and Bonneville Basin are currently valid as *C. b. semiscaber*. Although Neely (2003) argued for resurrecting *C. semiscaber* as a full species

and for continuing to recognize *C. hubbsi* as valid, the lack of reciprocal monophyly, the limited divergence, the overlap in geographical distribution, and the lack of diagnostic morphological traits suggest that only one name should apply to the rest of the members of this candidate species. Because the first named specimen from these 2 areas was *C. semiscaber* (PS 42), that name has precedence over *C. hubbsi*.

Recognizing *C. semiscaber* as the name for specimens from this region also renders uncertain the validity of 2 currently recognized species of sculpins presently or formerly restricted to lakes in the Bonneville Basin, *C. extensus* in Bear Lake in southeastern Idaho and northeastern Utah and *C. echinatus* from Utah Lake in Utah (Bailey and Bond 1963). Historical recognition of both species as distinct taxa relied on the extent of prickling and pre-opercular armature, both of which are variable within other species in this genus (McAllister and Lindsey 1961). Alternatively, both may represent relatively recently derived forms from adjacent fluvial populations. Fossils of *C. extensus* are common in sedimentary deposits of pluvial Lake Bonneville (Broughton and Smith 2016), demonstrating that the species once had a much broader distribution. Lake Bonneville, however, has gone through many cycles of enlargement and drying, contracting to an extent comparable to the present day as recently as 30 kya (Oviatt et al. 1999). This suggests that the species does not have a lengthy evolutionary history as a taxon distinct from stream-dwelling relatives (cf. Lucek et al. 2018). Whereas *C. extensus* is also thought to be distinctive because of its pelagic larvae, members of the *C. bairdii* complex are known to exhibit a life history shift between lacustrine populations with planktonic larvae and fluvial populations with reputedly benthic larvae (Goto et al. 2015). In freshwater fish, the dramatic variation in life history strategies within single species—or even among offspring from the same mating (Kendall et al. 2015)—suggests caution in recognizing lotic and lentic variants of single forms as distinct taxa, except in those cases when these divisions may be reinforced by long-term geographic isolation, i.e., as in Clear Lake sculpin of the *C. asper* complex (Moyle 2002). Hence, we recommend synonymizing *C. extensus* with *C. semiscaber* but continuing to recognize the

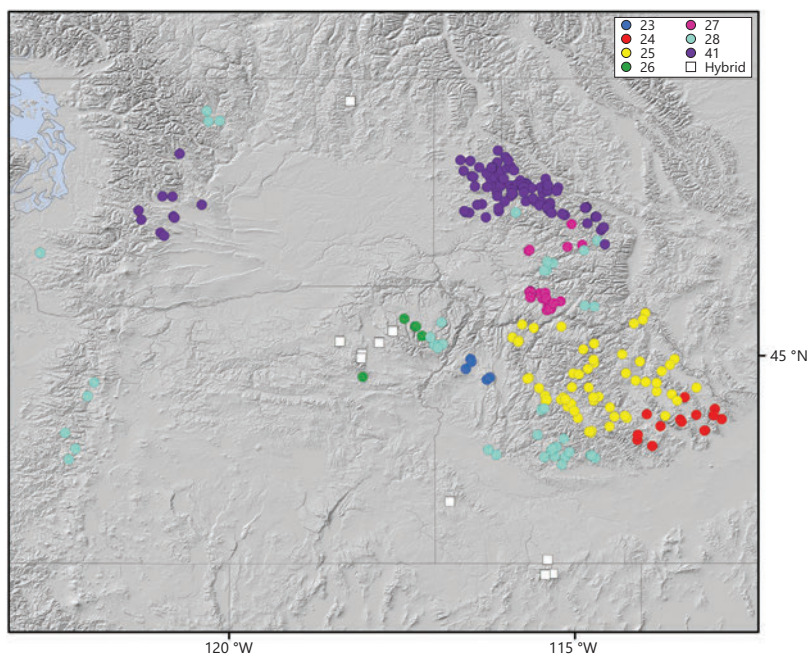


Fig. 11. Distribution of the *Cottus confusus* species complex (CS/PS 23–28, 41), including *C. confusus* (PS 25) and *C. schitsuumsh* (PS 41), and hybrid individuals.

population in Bear Lake (and other members of form 58 from the northern Bonneville Basin) as a unit of conservation. *Cottus echinatus* was thought to be extinct at the time it was described as a consequence of low water levels in Utah Lake several decades earlier (Bailey and Bond 1963). In their description, however, Bailey and Bond (1963) noted that paratypes of this species may have been collected from its major inlet, the Provo River. Our collections from that basin also assign to form 58, suggesting a shared origin with populations in Bear Lake and elsewhere in this area.

Subgenus *Uranidea* Part 4: the *Cottus confusus* Complex

Cottus confusus has been problematic since its description (Bailey and Bond 1963) as a single widely and disjunctly distributed species of the Columbia River basin. We corroborated previous analyses demonstrating that morphologically defined *C. confusus* referred to taxa in different species complexes (Peden et al. 1989, Neely 2003, LeMoine et al. 2014). Further evidence of the challenge to field workers is that, even when legitimate members of the *C. confusus* complex were

collected, they were often misidentified. Although Kinziger et al. (2005) were uncertain of the phylogenetic placement of this complex, our nuclear phylogenies indicate that it is closely related to the *C. bairdii* species complex and can be treated as a member of *Uranidea*. Elucidating candidate species was further complicated by a lack of consensus among species delimitation methods. Nevertheless, we concluded that this complex consisted of 6 highly supported, monophyletic, divergent, and generally geographically isolated groups that represented candidate taxa, a seventh species apparently of hybrid origin, and several specimens for which hybridization rendered assignment uncertain (Fig. 11).

One deeply divergent lineage was represented by 2 candidate taxa and 1 provisional taxon with restricted distributions in Idaho. The first (CS 23) was restricted to the Weiser River basin in west-central Idaho and the second (CS 24) to the Sinks drainages on the northeastern Snake River Plain in Idaho. Although the latter group has been thought to occupy the 5 drainages constituting the Sinks basin (Wallace and Zaroban 2013), we were

able to confirm its presence only in the 3 westernmost drainages, which have been isolated from drainages farther east for 1.2 My (Geslin et al. 2002). The Sinks specimens are morphologically distinguishable (Wallace and Zaroban 2013) from *C. confusus* sensu stricto (PS 25) from the Salmon River basin (type location, the Salmon River at the base of Galena Summit; Bailey and Bond 1963).

A second lineage consisted of 3 candidate species, the first 2 of which also had a limited range. One of these (CS 26) was found in the headwater tributaries of the Grande Ronde and Powder Rivers, which drain the northern and southern sides of the Blue Mountains in northeastern Oregon. The other (CS 27) comprised an assemblage of nonoverlapping, locally structured clades in different portions of the North and South Forks of the Clearwater River basin, which were never subject to Pleistocene glaciation and form part of the Rocky Mountain Refugium (Shafer et al. 2010). In striking contrast, the third candidate species (CS 28) showed little divergence but rather a broad distribution from the headwaters of the Clearwater River in central Idaho and the Boise River basin several hundred kilometers farther south, the Imnaha River basin in northeastern Oregon and the eastern tributaries of the Willamette River basin in western Oregon, and the Methow River basin on the east side and the Cowlitz River basin on the west side of the Cascade Range in Washington. The distribution of this group, somewhat akin to that of *C. tubulatus*, *C. semiscaber*, and a clade of westslope cutthroat trout (Young et al. 2018), is consistent with recent dispersal from a source in the Clearwater River to points elsewhere in the Columbia River basin via glacial floods or pooled floodwaters (Oviatt 2015).

A third lineage exhibited varying degrees of mitonuclear discord that prevented species delimitation for some, but not all, groups. *Cottus schitsuumsh* (PS 41) formed a readily diagnosed and strongly supported group consisting only of its members (mitochondrial phylogenies) or those of its species complex (nuclear phylogenies) and was delimited in all analyses. Nonetheless, it appears to be of hybrid origin because the nuclear sequences are closely related or identical to those of other members of the *C. confusus* species complex, whereas the mitochondrial

sequences are much more closely related to those of *C. semiscaber*.

Cottus schitsuumsh was originally considered to be *C. confusus*, but LeMoine et al. (2014) recognized it as a distinct taxon that ranged throughout the Spokane River basin in northern Idaho, with possibly introduced populations in the Clark Fork River basin in western Montana. Newly observed specimens from the Yakima and Wenatchee River basins in Washington exhibited novel haplotypes that suggested they were indigenous. In contrast, this species was also found at a single site in the headwaters of the Lochsa River in Idaho, but specimens had a haplotype identical to that of collections from near the type location (the Coeur d'Alene River basin, over 500 stream kilometers away), and we suspect these represent an introduced population.

Remaining forms and specimens could not be reliably assigned to separate candidate taxa. One (form 59) was found far to the south in 2 disjunct locations, the West Fork Jarbidge River in Idaho and Nevada and the North Powder River in northeastern Oregon. Although these specimens were strongly supported as sister to *C. schitsuumsh* in the mitochondrial phylogenies and shared nuclear haplotypes with that species, they were not monophyletic and co-occurred with other introgressed specimens, suggesting that their origins may be traceable to a different or ongoing hybridization event. Other specimens represented by mitochondrial forms of *C. semiscaber* and *C. punctulatus* (some members of forms 56–58) and nuclear sequences of the *C. confusus* species complex were in the East Fork Jarbidge River in Nevada, the Jordan River in Idaho, the Powder and Grande Ronde Rivers in northeastern Oregon, and the Kettle River in Washington. We did not observe heterozygous nuclear sequences indicative of recent introgression in these specimens, but that result may stem from a lack of statistical power, given the limited divergence between nuclear sequences of *C. confusus* and *C. punctulatus*. Unlike in *C. schitsuumsh*, the mitochondrial sequences of specimens in these groups showed little or no divergence from members of *C. semiscaber*. Their origins from matings between female *C. semiscaber* and male *C. confusus* and their broad distribution are intriguing, but whether these specimens constitute contemporary hybrids or taxa of recent hybrid

origin will require the genomic evaluation of a larger sample of individuals.

CONCLUSION

The taxonomy of sculpins remains one of the last major unresolved puzzles in the systematics of North American freshwater fishes. The bulk of this taxonomy has been based on morphological identification, which has proven problematic for species delimitation because morphological differences among even highly divergent taxa are often subtle or inconsistent, suggesting that niche conservatism and common responses to environmental factors may play a role in constraining phenotypic variation among taxa while promoting ecotypic variation within them (Fišer et al. 2018). Because of this, we have taken a different tack by proposing a taxonomy based on molecular and geographic data, which we regard as sufficient for delineating species hypotheses (Jörger and Schrödl 2013). Our intent was to provide a reevaluation of the recognizable biodiversity of this fauna, in part because the taxonomy has been understood as faulty almost since its inception, in part because the pace of species descriptions primarily reliant on morphology is painfully slow (mean time to description after discovery estimated to be 21 years according to Fontaine et al. [2012] and about 35 years according to Goodwin et al. [2020]), and in part because some members of this fauna may be at risk and are undergoing rapid range contractions (Adams et al. 2015, LeMoine et al. 2020). For those requiring morphological data to affirm these conclusions, we suspect that this molecular delineation will give evolutionary and geographical context to morphological comparisons among now-meaningful groups and perhaps reveal diagnostic anatomical differences among what were formerly cryptic taxa (rendering them pseudocryptic; Lajus et al. 2015). At the very least, it demonstrates the need to include genetic data to support future species descriptions in this group.

We acknowledge that our interpretation constitutes a suite of new hypotheses on evolutionary relationships among *Cottus* in this region, one subject to the limitations of the specimens, genes, and analyses used. Much remains to be done, in part because our coverage of some river basins was limited or absent,

but also because any range estimate based on sampling individuals from specific sites will underestimate the extent of occupied habitat. Widespread adoption of environmental DNA sampling would permit rapid and precise estimates of the distribution of each taxon (McKelvey et al. 2016). However, eDNA methods still require diagnostic genetic markers for each taxon derived from sequencing of individuals. Also necessary—especially for taxa of uncertain standing or apparently originating from introgression—are genomic tools that access a broader and more representative portion of the genome and can help ascertain the timing of divergence and its relation to geological events. Regardless, the forms, candidate taxa, and provisional taxa we have identified constitute fundamental units of biodiversity that are often the targets for conservation, measures of ecosystem health, and components of ecological studies; their delineation represents another step toward a more comprehensive understanding of the ecology and complexity of freshwater ecosystems in western North America.

SUPPLEMENTARY MATERIAL

Seven online-only supplementary files accompany this article. (<https://scholarsarchive.byu.edu/wnan/vol82/iss2/8>).

SUPPLEMENTARY MATERIAL 1. A synonymy of *Cottus* in western North America and the representative specimens from this study nearest the type location.

SUPPLEMENTARY MATERIAL 2. Specimens or sequences used in the analyses.

SUPPLEMENTARY MATERIAL 3. Primers, reaction volumes, and cycling conditions for gene amplification and sequencing.

SUPPLEMENTARY MATERIAL 4. Maximum-likelihood phylogeny of *Cottus* based on rhodopsin haplotypes ($n = 68$ derived from 482 sequences). Inset, amino acid phylogeny.

SUPPLEMENTARY MATERIAL 5. Diagnosis of candidate species using a neighbor-joining tree of COI sequences ($n = 4266$) from this study and public databases.

SUPPLEMENTARY MATERIAL 6. Diagnosis of candidate species using a neighbor-joining tree of cytb sequences ($n = 1152$) from this study and public databases.

SUPPLEMENTARY MATERIAL 7. Contributors of sequences, DNA, tissues, or whole specimens to the SculpinQwest project.

ACKNOWLEDGMENTS

We thank the dozens of biologists (Supplementary Material 7) whose contributions made this project possible. In particular, we thank Brian Sidlauskas and Emily Dziedzic (Oregon State University) for providing DNA samples and mitochondrial sequences for 20 Oregon specimens. Dona Horan (U.S. Forest Service) assisted with species maps. Funding was provided by the Washington Department of Fish and Wildlife, U.S. Geological Survey, U.S. Fish and Wildlife Service, Colorado Parks and Wildlife, Oregon Chapter of the American Fisheries Society, and Rocky Mountain Research Station. Funding agencies had no role in study design and interpretation or in manuscript preparation and submission.

DATA AVAILABILITY

Sequence information has been deposited in GenBank for COI (OL765312–OL769311), cytb (OL793250–OL793837), rhodopsin (OL792801–OL793249), and S7 (OL793838–OL794278).

LITERATURE CITED

- ADAMS, S.B., D.A. SCHMETTERLING, AND D.A. NEELY. 2015. Summer stream temperatures influence sculpin distributions and spatial partitioning in the upper Clark Fork River basin, Montana. *Copeia* 103:416–428.
- AHRENS, D., T. FUJISAWA, H.J. KRAMMER, J. EBERLE, S. FABRIZI, AND A.P. VOGLER. 2016. Rarity and incomplete sampling in DNA-based species delimitation. *Systematic Biology* 65:478–494.
- APRIL, J., R.L. MAYDEN, R.H. HANNER, AND L. BERNATCHEZ. 2011. Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Sciences of the United States of America* 108:10602–10607.
- BAILEY, R.M., AND C.E. BOND. 1963. Four new species of freshwater sculpins, genus *Cottus*, from western North America. *Occasional Papers of the Museum of Zoology at the University of Michigan* No. 634. University of Michigan Press, Ann Arbor, MI. 27 pp.
- BAILEY, R.M., AND M.F. DIMICK. 1949. *Cottus hubbsi*, a new cottid fish from the Columbia River system in Washington and Idaho. *Occasional Papers of the Museum of Zoology at the University of Michigan* No. 513. University of Michigan Press, Ann Arbor, MI. 18 pp.
- BARLEY, A.J., J.M. BROWN, AND R.C. THOMSON. 2018. Impact of model violations on the inference of species boundaries under the multispecies coalescent. *Systematic Biology* 67:269–284.
- BAUMSTEIGER, J., A.P. KINZIGER, AND A. AGUILAR. 2012. Life history and biogeographic diversification of an endemic western North American freshwater fish clade using a comparative species tree approach. *Molecular Phylogenetics and Evolution* 65:940–952.
- BAUMSTEIGER, J., A.P. KINZIGER, AND A. AGUILAR. 2016. Novel concordance between geographic, environmental, and genetic structure in the ecological generalist prickly sculpin (*Cottus asper*) in California. *Journal of Heredity* 107:504–517.
- BAUMSTEIGER, J., A.P. KINZIGER, S.B. REID, AND A. AGUILAR. 2014. Complex phylogeography and historical hybridization between sister taxa of freshwater sculpin (*Cottus*). *Molecular Ecology* 23:2602–2618.
- BAXTER, G.T., AND M.D. STONE. 1995. *Fishes of Wyoming*. Wyoming Game and Fish Department, Cheyenne, WY. 290 pp.
- BEAN, T.H. 1881. Descriptions of new species of fishes (*Uranidea marginata*, *Potamocottus bendirei*) and of *Myctophum crenulare* J. and G. *Proceedings of the United States National Museum* 4:26–29.
- BEHRENS-CHAPUIS, S., F. HERDER, H.R. ESMAEILI, J. FREYHOF, N.A. HAMIDAN, M. ÖZULUÇ, R. ŠANDA, AND M.F. GEIGER. 2015. Adding nuclear rhodopsin data where mitochondrial COI indicates discrepancies—can this marker help to explain conflicts in cyprinids? *DNA Barcodes* 3:187–199.
- BERGSTEN, J., D.T. BILTON, T. FUJISAWA, M. ELLIOTT, M.T. MONAGHAN, M. BALKÉ, L. HENDRICH, J. GEIJER, J. HERRMANN, G.N. FOSTER, ET AL. 2012. The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology* 61:851–869.
- BERNATCHEZ, L., AND C.C. WILSON. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology* 7:431–452.
- BESTGEN, K.R., K.B. ROGERS, AND R. GRANGER. 2019. Distinct phenotypes of native cutthroat trout emerge under a molecular model of lineage distributions. *Transactions of the American Fisheries Society* 148: 442–463.
- BETANCUR-R, R., E.O. WILEY, G. ARRATIA, A. ACERO, N. BAILLY, M. MIYA, G. LECOINTRE, AND G. ORTÍ. 2017. Phylogenetic classification of bony fishes. *BMC Evolutionary Biology* 17:162.
- BICKFORD, D., D.J. LOHMAN, N.S. SODHI, P.K. NG, R. MEIER, K. WINKER, K.K. INGRAM, AND I. DAS. 2007. Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* 22:148–155.
- BISSON, P.A., AND C.E. BOND. 1971. Origin and distribution of the fishes of Harney Basin, Oregon. *Copeia* 1971:268–281.
- BODALY, R.A., AND C.C. LINDSEY. 1977. Pleistocene watershed exchanges and the fish fauna of the Peel River Basin, Yukon Territory. *Journal of the Fisheries Research Board of Canada* 34:388–395.
- BOND, C.E. 1963. Distribution and ecology of freshwater sculpins, genus *Cottus*, in Oregon. *Doctoral dissertation*, University of Michigan, Ann Arbor, MI. 186 pp.
- BOOTH, D.B., K.G. TROOST, J.J. CLAGUE, AND R.B. WAITT. 2003. The Cordilleran ice sheet. *Developments in Quaternary Sciences* 1:17–43.
- BORCHSENIUS, F. 2009. *FastGap 1.2*. Department of Biosciences, Aarhus University, Denmark. http://www.aubot.dk/FastGap_home.htm
- BOUCKAERT, R., T.G. VAUGHAN, J. BARIDO-SOTTANI, S. DUCHÈNE, M. FOURMENT, A. GAVRYUSHKINA, J. HELED, G. JONES, D. KÜHNERT, N. DE MAIO, ET AL. 2019. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology* 15:e1006650.
- BROUGHTON, J.M., AND G.R. SMITH. 2016. The fishes of Lake Bonneville: implications for drainage history,

- biogeography, and lake levels. Pages 292–351 in C.G. Oviatt and J.F. Shroder, editors, *Lake Bonneville: a scientific update. Developments in Earth Surface Processes*, Volume 20. Elsevier, Amsterdam.
- BRYANT, D., AND M.W. HAHN. 2020. The concatenation question. Pages 3.4:1–3.4:23 in C. Scornavacca, F. Delsuc, and N. Galtier, editors, *Phylogenetics in the Genomic Era*. HAL Open Science. <https://hal.inria.fr/PGE>
- CARSTENS, B.C., T.A. PELLETIER, N.M. REID, AND J.D. SATLER. 2013. How to fail at species delimitation. *Molecular Ecology* 22:4369–4383.
- CAVENDER, T.M. 1986. Review of the fossil history of North American freshwater fishes. Pages 699–724 in C.H. Hocutt and E.O. Wiley, editors, *The zoogeography of North American freshwater fishes*. John Wiley & Sons, New York, NY.
- CHEN, H., M. STRAND, J.L. NORENBURG, S. SUN, H. KAJIHARA, A.V. CHERNYSHEV, S.A. MASLAKOVA, AND P. SUNDBERG. 2010. Statistical parsimony networks and species assemblages in cephalotrichid nemertean (Nemertea). *PLOS ONE* 5:e12885.
- CLEMENT, M., D.C. POSADA, AND K.A. CRANDALL. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9:1657–1659.
- COLLINS, R.A., AND R.H. CRUICKSHANK. 2013. The seven deadly sins of DNA barcoding. *Molecular Ecology Resources* 13:969–975.
- [COSEWIC] COMMITTEE ON THE STATUS OF ENDANGERED WILDLIFE IN CANADA. 2018. COSEWIC guidelines for recognizing designatable units. Committee on the Status of Endangered Wildlife in Canada, Ottawa; [accessed 27 January 2021]. <https://cosewic.ca/index.php/en-ca/reports/preparing-status-reports/guidelines-recognizing-designatable-units#:~:text=Designatable%20units%20should%20be%20discrete,be%20replaced%20through%20natural%20dispersion>
- [COSEWIC] COMMITTEE ON THE STATUS OF ENDANGERED WILDLIFE IN CANADA. 2019. COSEWIC assessment and status report on the Rocky Mountain sculpin *Cottus* sp., Pacific populations, Saskatchewan-Nelson River populations and Missouri River populations in Canada. Committee on the Status of Endangered Wildlife in Canada, Ottawa; [accessed 5 March 2021]. <https://www.canada.ca/en/environment-climate-change/services/species-risk-public-registry/cosewic-assessments-status-reports/rocky-mountain-sculpin-2019.html>
- COPE, E.D. 1872. Report on the recent reptiles and fishes of the survey, collected by Campbell Carrington and C.M. Dawes. Pages 467–476 in F.V. Hayden, editor, *Preliminary report of the United States Geological Survey of Montana and portions of adjacent territories*, 5th annual report.
- CUMBAA, S.L., D.E. MCALLISTER, AND R.E. MORLAN. 1981. Late Pleistocene fish fossils of *Coregonus*, *Stenodus*, *Thymallus*, *Catostomus*, *Lota*, and *Cottus* from the Old Crow Basin, northern Yukon, Canada. *Canadian Journal of Earth Sciences* 18:1740–1754.
- DEHAAN, P.W., B.A. ADAMS, R.A. TABOR, D.K. HAWKINS, AND B. THOMPSON. 2014. Historical and contemporary forces shape genetic variation in the Olympic mudminnow (*Novumbra hubbsi*), an endemic fish from Washington State, USA. *Conservation Genetics* 15:1417–1431.
- DELLICOUR, S., AND J.F. FLOT. 2018. The hitchhiker's guide to single locus species delimitation. *Molecular Ecology Resources* 18:1234–1246.
- DENNENMOSER, S., A.W. NOLTE, S.M. VAMOSI, AND S.M. ROGERS. 2015. Phylogeography of the prickly sculpin (*Cottus asper*) in north western North America reveals parallel phenotypic evolution across multiple coastal–inland colonizations. *Journal of Biogeography* 42:1626–1638.
- DENNENMOSER, S., S.M. ROGERS, AND S.M. VAMOSI. 2014. Genetic population structure in prickly sculpin (*Cottus asper*) reflects isolation-by-environment between two life-history ecotypes. *Biological Journal of the Linnean Society* 113:943–957.
- DE QUEIROZ, K. 2007. Species concepts and species delimitation. *Systematic Biology* 56:879–886.
- DUK-RODKIN, A., R.W. BARENDREGT, D.G. FROESE, F. WEBER, R. ENKIN, I.R. SMITH, G.D. ZAZULA, P. WATERS, AND R. KLASSEN. 2004. Timing and extent of Plio-Pleistocene glaciations in north-western Canada and east-central Alaska. *Developments in Quaternary Sciences* 2:313–345.
- DYKE, A.S., A. MOORE, AND L. ROBERTSON. 2003. Deglaciation of North America. Geological Survey of Canada, Open File 1574. Natural Resources Canada, Canada.
- EATON, K.R., J.L. LOXTERMAN, AND E.R. KEELEY. 2018. Connections and containers: using genetic data to understand how watershed evolution and human activities influence cutthroat trout biogeography. *PLOS ONE* 13:e0202043.
- EDGAR, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792–1797.
- EIGENMANN, C.H., AND R.S. EIGENMANN. 1892. New fishes from western Canada. *American Naturalist* 26:961–964.
- EUCLIDE, P.T., N.M. FLORES, M.J. WARGO, C.W. KILPATRICK, AND J.E. MARSDEN. 2018. Lack of genetic population structure of slimy sculpin in a large, fragmented lake. *Ecology of Freshwater Fish* 27:699–709.
- FISER, C., C.T. ROBINSON, AND F. MALARD. 2018. Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology* 27:613–635.
- FLOT, J.F. 2007. Champuru 1.0: a computer software for unraveling mixtures of two DNA sequences of unequal lengths. *Molecular Ecology Notes* 7:974–977.
- FLOT, J.F. 2010. SEQPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. *Molecular Ecology Resources* 10:162–166.
- FONTAINE, B., A. PERRARD, AND P. BOUCHET. 2012. 21 years of shelf life between discovery and description of new species. *Current Biology* 22:R943–R944.
- FONTANETO, D., J.F. FLOT, AND C.Q. TANG. 2015. Guidelines for DNA taxonomy, with a focus on the meiofauna. *Marine Biodiversity* 45:433–451.
- FREYHOF, J., M. KOTTELAT, AND A. NOLTE. 2005. Taxonomic diversity of European *Cottus* with description of eight new species (Teleostei: Cottidae). *Ichthyological Exploration of Freshwaters* 16:107–122.
- FRICKE, R., W.N. ESCHMEYER, AND R. VAN DER LAAN. 2021. Eschmeyer's catalog of fishes: genera, species, references. [Accessed 3 March 2021]. <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>
- GALTIER, N. 2019. Delineating species in the speciation continuum: a proposal. *Evolutionary Applications* 12:657–663.

- GESLIN, J.K., P.K. LINK, J.W. RIESTERER, M.A. KUNTZ, AND C.M. FANNING. 2002. Pliocene and Quaternary stratigraphic architecture and drainage systems of the Big Lost Trough, northeastern Snake River Plain, Idaho. Pages 11–26 in P.K. Link and L.L. Mink, editors, *Geology, hydrogeology, and environmental remediation: Idaho National Engineering and Environmental Laboratory, eastern Snake River Plain, Idaho*. GSA Special Papers, volume 353. Geological Society of America, Washington, DC.
- GIANNINY, G.L., G.D. THACKRAY, D.S. KAUFMAN, S.L. FORMAN, M.J. SHERBONDY, AND D. FINDEISEN. 2002. Late Quaternary highstands in the Mud Lake and Big Lost Trough subbasins of Lake Terreton, Idaho. Pages 77–90 in P.K. Link and L.L. Mink, editors, *Geology, hydrogeology, and environmental remediation: Idaho National Engineering and Environmental Laboratory, eastern Snake River Plain, Idaho*. GSA Special Papers, volume 353. Geological Society of America, Washington, DC.
- GILBERT, C.H., AND B.W. EVERMANN. 1894. A report upon investigations in the Columbia River basin, with descriptions of four new species of fishes. *Bulletin of the United States Fish Commission* 14:169–207.
- GILL, T. 1862. Observations on the genus *Cottus*, and descriptions of two new species (abridged from the forthcoming report of Capt. J.H. Simpson). *Proceedings of the Boston Society of Natural History* 8:40–42.
- GILL, T. 1876. Appendix L. Report on ichthyology. Pages 385–431 in J.H. Simpson, *Report of explorations across the Great Basin of the territory of Utah for a wagon-route from Camp Floyd to Genoa, in Carson Valley, in 1859*. Government Printing Office, Washington, DC. 518 pp.
- GOODWIN, Z.A., P. MUÑOZ-RODRÍGUEZ, D.J. HARRIS, T. WELLS, J.R. WOOD, D. FILER, AND R.W. SCOTLAND. 2020. How long does it take to discover a species? *Systematics and Biodiversity* 18:784–793.
- GOTO, A., R. YOKOYAMA, AND V.G. SIDELEVA. 2015. Evolutionary diversification in freshwater sculpins (Cottoidea): a review of two major adaptive radiations. *Environmental Biology of Fishes* 98:307–335.
- GRAY, M.A., R.A. CURRY, T.J. ARCISZEWSKI, K.R. MUNKITTRICK, AND S.M. BRASFIELD. 2018. The biology and ecology of slimy sculpin: a recipe for effective environmental monitoring. *Facets* 3:103–127.
- HAPONSKI, A.E., AND C.A. STEPIEN. 2013. Phylogenetic and biogeographical relationships of the *Sander* pikeperches (Percidae: Perciformes): patterns across North America and Eurasia. *Biological Journal of the Linnean Society* 110:156–179.
- HART, M.W., AND J. SUNDAY. 2007. Things fall apart: biological species form unconnected parsimony networks. *Biology Letters* 3:509–512.
- HEBERT, P.D., S. RATNASINGHAM, AND J.R. DE WAARD. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London, Series B: Biological Sciences* 270 (supplement 1):S96–S99.
- HIDY, A.J., J.C. GOSSE, D.G. FROESE, J.D. BOND, AND D.H. ROOD. 2013. A latest Pliocene age for the earliest and most extensive Cordilleran ice sheet in northwestern Canada. *Quaternary Science Reviews* 61:77–84.
- HILL, G.E. 2020. Genetic hitchhiking, mitonuclear coadaptation, and the origins of mt DNA barcode gaps. *Ecology and Evolution* 10:9048–9059.
- HILL, J., E.D. ENBODY, M.E. PETERSSON, C.G. SPREHN, D. BEKKEVOLD, A. FOLKVORD, L. LAIKRE, G. KLEINAU, P. SCHEERER, AND L. ANDERSSON. 2019. Recurrent convergent evolution at amino acid residue 261 in fish rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* 116:18473–18478.
- HINTZE, L.F. 1988. Geologic history of Utah. Brigham Young University Geological Studies Special Publication No. 7. Department of Geology, Brigham Young University, Provo, UT. 202 pp.
- HONEYCUTT, R.L., D.M. HILLIS, AND J.W. BICKHAM. 2010. Biodiversity discovery and its importance to conservation. Pages 1–34 in J.A. DeWoody, J.W. Bickham, C.H. Michler, K.M. Nichols, O.E. Rhodes, and K.E. Woeste, editors, *Molecular approaches in natural resource conservation and management*. Cambridge University Press, New York, NY.
- HOWARD, A.D. 1958. Drainage evolution in northeastern Montana and northwestern North Dakota. *Geological Society of America Bulletin* 69:575–588.
- HUBBS, C.L., AND L.P. SCHULTZ. 1932. *Cottus tubulatus*, a new sculpin from Idaho. *Occasional Papers of the Museum of Zoology at the University of Michigan* No. 242. University of Michigan, Ann Arbor, MI.
- HUBERT, N., R. HANNER, E. HOLM, N.E. MANDRAK, E. TAYLOR, M. BURRIDGE, D. WATKINSON, P. DUMONT, A. CURRY, P. BENTZEN, ET AL. 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PLOS ONE* 3:e2490.
- HUDY, M., AND J. SHIFLET. 2009. Movement and recolonization of Potomac sculpin in a Virginia stream. *North American Journal of Fisheries Management* 29:196–204.
- ISAAK, D.J., M.K. YOUNG, C. MCCONNELL, B.B. ROPER, E.K. ARCHER, B. STAAB, C. HIRSCH, D.E. NAGEL, M.K. SCHWARTZ, AND G.L. CHANDLER. 2018. Crowdsourced databases as essential elements for Forest Service partnerships and aquatic resource conservation. *Fisheries* 43:423–430.
- IZETT, G.A., AND R.E. WILCOX. 1982. Map showing localities and inferred distributions of the Huckleberry Ridge, Mesa Falls, and Lava Creek ash beds (Pearlette family ash beds) of Pliocene and Pleistocene age in the western United States and southern Canada. *Miscellaneous Investigations Series Map I-1325*. U.S. Geological Survey, Reston, VA.
- JENKINS, R.E., AND N.M. BURKHEAD. 1994. *Freshwater fishes of Virginia*. American Fisheries Society, Bethesda, MD. 1080 pp.
- JONES, G. 2017. Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *Journal of Mathematical Biology* 74:447–467.
- JÖRGER, K.M., AND M. SCHRÖDL. 2013. How to describe a cryptic species? Practical challenges of molecular taxonomy. *Frontiers in Zoology* 10:59.
- KATOH, K., J. ROZEWICKI, AND K.D. YAMADA. 2019. Mafft online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics* 20:1160–1166.
- KENDALL, N.W., J.R. MCMILLAN, M.R. SLOAT, T.W. BUEHRENS, T.P. QUINN, R.W. PESS, K.V. KUZISHCHIN, M.M. MCCLURE, AND R.W.I. ZABEL. 2015. Anadromy and residency in steelhead and rainbow trout (*Oncorhynchus mykiss*): a review of the processes and patterns. *Canadian Journal of Fisheries and Aquatic Sciences* 72:319–342.

- KIMMEL, P.G. 1975. Fishes of the Miocene-Pliocene Deer Butte Formation, southeast Oregon. University of Michigan Paleontology Museum: Papers on Paleontology 14:69–87.
- KINZIGER, A.P., M. HELLMAIR, S.R. FONG, D.H. GOODMAN, AND H. KELSEY. 2016. Evolution of rough sculpin (*Cottus asperimus*) genetic divergence and late Quaternary displacement on the Hat Creek fault, California, USA. *Conservation Genetics* 17:1257–1267.
- KINZIGER, A.P., R.M. WOOD, AND D.A. NEELY. 2005. Molecular systematics of the genus *Cottus* (Scorpaeniformes: Cottidae). *Copeia* 2005:303–311.
- KREJSA, R.J. 1967. The systematics of the prickly sculpin, *Cottus asper* Richardson, a polytypic species. Part II. Studies on the life history, with especial reference to migration. *Pacific Science* 21:414–422.
- KUMAR, S., G. STECHER, AND K. TAMURA. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33:1870–1874.
- KUNTZ, M.A., J.K. GESLIN, L.E. MARK, M.K.V. HODGES, M.E. KAUFFMAN, D.E. CHAMPION, M.R. LANPHERE, D.W. RODGERS, M.H. ANDERS, P.K. LINK, AND D.L. BOYACK. 2003. Geologic map of the northern and central parts of the Idaho National Engineering and Environmental Laboratory, eastern Idaho. Idaho Geological Survey Map GM-35, scale 1:50,000.
- LAJUS, D., N. SUKHIKH, AND V. ALEKSEEV. 2015. Cryptic or pseudocryptic: can morphological methods inform copepod taxonomy? An analysis of publications and a case study of the *Eurytemora affinis* species complex. *Ecology and Evolution* 5:2374–2385.
- LARSON, E.R., C.L. ABBOTT, N. USIO, N. AZUMA, K.A. WOOD, L.M. HERBORG, AND J.D. OLDEN. 2012. The signal crayfish is not a single species: cryptic diversity and invasions in the Pacific Northwest range of *Pacifastacus leniusculus*. *Freshwater Biology* 57:1823–1838.
- LEE, D.S., C.R. GILBERT, C.H. HOCUTT, R.E. JENKINS, D.E. MCALLISTER, AND J.R. STAUFFER. 1980. Atlas of North American freshwater fishes. North Carolina State Museum of Natural History, Raleigh, NC.
- LEMOINE, M.T., L.A. EBY, C.G. CLANCY, L.G. NYCE, M.J. JAKOBER, AND D.J. ISAAK. 2020. Landscape resistance mediates native fish species distribution shifts and vulnerability to climate change in riverscapes. *Global Change Biology* 26:5492–5508.
- LEMOINE, M., M.K. YOUNG, K.S. MCKELVEY, L. EBY, K.L. PILGRIM, AND M.K. SCHWARTZ. 2014. *Cottus schitsuumsh*, a new species of sculpin (Scorpaeniformes: Cottidae) in the Columbia River basin, Idaho-Montana, USA. *Zootaxa* 3755:241–258.
- LIM, G.S., M. BALKER, AND R. MEIER. 2012. Determining species boundaries in a world full of rarity: singletons, species delimitation methods. *Systematic Biology* 61:165–169.
- LINDSEY, C.C., AND J.D. MCPHAIL. 1986. Zoogeography of fishes of the Yukon and McKenzie Basins. Pages 639–674 in C.H. Hocutt and E.O. Wiley, editors, *The zoogeography of North American freshwater fishes*. John Wiley & Sons, New York, NY.
- LÖYTINOJA, A., AND N. GOLDMAN. 2010. webprank: a phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinformatics* 11:579.
- LUCEK, K., I. KELLER, A.W. NOLTE, AND O. SEEHAUSEN. 2018. Distinct colonization waves underlie the diversification of the freshwater sculpin (*Cottus gobio*) in the central European alpine region. *Journal of Evolutionary Biology* 31:1254–1267.
- LYONS, J. 1990. Distribution and morphological variation of the slimy sculpin (*Cottus cognatus*) in the north central United States. *Canadian Journal of Zoology* 68:1037–1045.
- MARKLE, D.F. 2016. A guide to freshwater fishes of Oregon. Oregon State University Press, Corvallis, OR. 140 pp.
- MARKLE, D.F. 2019. Drainage evolution and freshwater fish zoogeography in coastal Oregon and Washington. *Northwestern Naturalist* 100:71–89.
- MARKLE, D.F., AND D.L. HILL JR. 2000. Taxonomy and distribution of the Malheur mottled sculpin, *Cottus bendirei*. *Northwest Science* 74:202–211.
- MAUGHAN, O.E. 1974. A reinvestigation of the type locality of *Cottus tubulatus*. *Northwest Science* 48:105–106.
- MCALLISTER, D.E. 1964. Distinguishing characters for the sculpins *Cottus bairdii* and *C. cognatus* in eastern Canada. *Journal of the Fisheries Research Board of Canada* 21:1339–1342.
- MCALLISTER, D.E., AND C.C. LINDSEY. 1961. Systematics of the freshwater sculpins (*Cottus*) of British Columbia. *Bulletin of the National Museum of Canada* 172:66–89.
- MCCLELLAN, P.H., AND G.R. SMITH. 2020. Late Miocene fishes of the Cache Valley Member, Salt Lake Formation, Utah and Idaho. *Miscellaneous Publications No. 208*. Museum of Zoology, University of Michigan, Ann Arbor, MI.
- MCKELVEY, K.S., M.K. YOUNG, W.L. KNOTEK, K.J. CARIM, T.M. WILCOX, T.M. PADGETT STEWART, AND M.K. SCHWARTZ. 2016. Sampling large geographic areas for rare species using environmental DNA: a study of bull trout *Salvelinus confluentus* occupancy in western Montana. *Journal of Fish Biology* 88:1215–1222.
- MCLEISH, J., R.A. BRIERS, J.A. DODD, AND S. RUECKERT. 2020. First genetic evidence that invasive bullhead (*Cottus* L. 1758) in Scotland is of English origin and the difficulty of resolving the European *Cottus* species taxonomy. *Journal of Fish Biology* 96:617–630.
- MCPHAIL, J.D. 1967. Distribution of freshwater fishes in western Washington. *Northwest Science* 41:1–11.
- MCPHAIL, J.D. 2007. *The freshwater fishes of British Columbia*. University of Alberta Press, Edmonton, Canada. 620 pp.
- MCPHAIL, J.D., AND C.C. LINDSEY. 1986. Zoogeography of the freshwater fishes of Cascadia (the Columbia system and rivers north to Stikine). Pages 615–637 in C.H. Hocutt and E.O. Wiley, editors, *The zoogeography of North American freshwater fishes*. John Wiley & Sons, New York, NY.
- MCPHAIL, J.D., AND E.B. TAYLOR. 2009. Phylogeography of the longnose dace (*Rhinichthys cataractae*) species group in northwestern North America—the origin and evolution of the Umpqua and Millicoma dace. *Canadian Journal of Zoology* 87:491–497.
- MEIER, R., G. ZHANG, AND F. ALI. 2008. The use of mean instead of smallest interspecific distances exaggerates the size of the “barcoding gap” and leads to misidentification. *Systematic Biology* 57:809–813.
- MILLER, B.A. 2006. *The phylogeography of Prosopium in western North America*. Master’s thesis, Brigham Young University, Provo, UT. 150 pp.
- MINCKLEY, W.L., D.A. HENDRICKSON, AND C.E. BOND. 1986. Geography of western North American freshwater fishes: description and relationships to intra-continental tectonism. Pages 519–613 in C.H.

- Hocutt and E.O. Wiley, editors, The zoogeography of North American freshwater fishes. John Wiley & Sons, New York, NY.
- MINH, B.Q., M.A. NGUYEN, AND A. VON HAESELER. 2013. Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution* 30:1188–1195.
- MOORE, J.S., R. BAJNO, J.D. REIST, AND E.B. TAYLOR. 2015. Post glacial recolonization of the North American Arctic by Arctic char (*Salvelinus alpinus*): genetic evidence of multiple northern refugia and hybridization between glacial lineages. *Journal of Biogeography* 42:2089–2100.
- MORRISON, D.A. 2015. Is sequence alignment an art or a science? *Systematic Botany* 40:14–26.
- MOYLE, P.B. 2002. *Inland fishes of California: revised and expanded*. University of California Press, Berkeley, CA. 517 pp.
- MUSSMANN, S.M. 2018. Diversification across a dynamic landscape: phylogeography and riverscape genetics of speckled dace (*Rhinichthys osculus*) in western North America. Doctoral dissertation, University of Arkansas, Fayetteville, AR. 166 pp.
- NAGY, L.G., S. KOCSUBÉ, Z. CSANÁDI, G.M. KOVÁCS, T. PETKOVITS, C. VÁGVÖLGYI, AND T. PAPP. 2012. Re-mind the gap! Insertion–deletion data reveal neglected phylogenetic potential of the nuclear ribosomal internal transcribed spacer (ITS) of fungi. *PLOS ONE* 7: e49794.
- NEELY, D.A. 2003. A systematic and taxonomic reassessment of the mottled sculpin species complex, *Cottus bairdii* Girard (Teleostei: Cottidae). Doctoral dissertation, University of Alabama, Tuscaloosa, AL. 346 pp.
- NGUYEN, L.T., H.A. SCHMIDT, A. VON HAESELER, AND B.Q. MINH. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32: 268–274.
- NOLTE, A.W., Z. GOMPERT, AND C.A. BUERKLE. 2009. Variable patterns of introgression in two sculpin hybrid zones suggest genomic isolation differs among populations. *Molecular Ecology* 18:2615–2627.
- O'CONNOR, J.E., V.R. BAKER, R.B. WAITT, L.N. SMITH, C.M. CANNON, D.L. GEORGE, AND R.P. DENLINGER. 2020. The Missoula and Bonneville floods—a review of ice-age megafloods in the Columbia River basin. *Earth-Science Reviews* 103181.
- OH, S.Y. 2016. Phylogenetic relationships of cottids (Pisces: Cottidae) in upper Snake River basin of western North America. Master's thesis, Brigham Young University, Provo, UT. 58 pp.
- OVIATT, C.G. 2015. Chronology of Lake Bonneville, 30,000 to 10,000 yr BP. *Quaternary Science Reviews* 110: 166–171.
- OVIATT, C.G., R.S. THOMPSON, D.S. KAUFMAN, J. BRIGHT, AND R.M. FORESTER. 1999. Reinterpretation of the Burmester Core, Bonneville Basin, Utah. *Quaternary Research* 52:180–184.
- PAGE, L.M., AND B.M. BURR. 2011. *Peterson field guide to freshwater fishes of North America north of Mexico*. 2nd edition. Houghton Mifflin Harcourt, Boston, MA.
- PEDEN, A.E., G.W. HUGHES, AND W.E. ROBERTS. 1989. Morphologically distinct populations of the short-head sculpin, *Cottus confusus*, and mottled sculpin, *Cottus bairdii* (Pisces, Cottidae), near the western border of Canada and the United States. *Canadian Journal of Zoology* 67:2711–2720.
- PEDERSEN, M.W., A. RUTER, C. SCHWEGER, H. FRIEBE, R.A. STAFF, K.K. KJELDSEN, M.L. MENDOZA, A.B. BEAUDOIN, ET AL. 2016. Postglacial viability and colonization in North America's ice-free corridor. *Nature* 537:45–49.
- PIERCE, K.L., AND L.A. MORGAN. 2009. Is the track of the Yellowstone hotspot driven by a deep mantle plume?—Review of volcanism, faulting, and uplift in light of new data. *Journal of Volcanology and Geothermal Research* 188:1–25.
- PULLANDRE, N., S. BROUILLET, AND G. ACHAZ. 2021. ASAP: assemble species by automatic partitioning. *Molecular Ecology Resources* 21:609–620.
- PULLANDRE, N., A. LAMBERT, S. BROUILLET, AND G. ACHAZ. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* 21:1864–1877.
- QUIST, M., W. HUBERT, AND D. ISAAK. 2004. Factors affecting allopatric and sympatric occurrence of two sculpin species across a Rocky Mountain watershed. *Copeia* 2004:617–623.
- RATNASINGHAM, S., AND P.D. HEBERT. 2013. A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. *PLOS ONE* 8:e66213.
- REHEIS, M.C., A.M. SARNA-WOJCICKI, R.L. REYNOLDS, C.A. REPENNING, AND M.D. MIFFLIN. 2002. Pliocene to middle Pleistocene lakes in the western Great Basin—ages and connections. Pages 53–108 in R. Hershler, D. Currey, and D. Madsen, editors, *Great Basin aquatic systems history*. Smithsonian Contributions to Earth Sciences No. 33. Smithsonian Institution Press, Washington, DC.
- REIMERS, P.E., AND C.E. BOND. 1967. Distribution of fishes in tributaries of the lower Columbia River. *Copeia* 1967:541–550.
- RICHARDSON, J. 1836. *The fish. Fauna Boreali-Americana; or the zoology of the northern parts of British America: containing descriptions of the objects of natural history collected on the late northern land expeditions, under the command of Sir John Franklin, R.N.* Volume 3. John Murray, London, England.
- ROBINS, C.R. 1954. A taxonomic revision of the *Cottus bairdii* and *Cottus carolinae* species groups in eastern North America (Pisces, Cottidae). Doctoral dissertation, Cornell University, Ithaca, NY.
- ROBINS, C.R., AND R.R. MILLER. 1957. Classification, variation, and distribution of the sculpins, genus *Cottus*, inhabiting Pacific slope waters in California and southern Oregon, with a key to the species. *California Fish and Game* 43:213–233.
- ROWSEY, D.M., AND J.J. EGGE. 2017. Morphometric analysis of two enigmatic sculpin species, *Cottus gulosus* and *Cottus perplexus* (Scorpaeniformes: Cottidae). *Northwestern Naturalist* 98:190–202.
- ROZAS, J., A. FERRER-MATA, J.C. SÁNCHEZ-DELBARRIO, S. GUIRAO-RICO, P. LIBRADO, S.E. RAMOS-ONSINS, AND A. SÁNCHEZ-GRACIA. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* 34:3299–3302.
- RUDOLFSSEN, T., J.L. RUPPERT, E.B. TAYLOR, C.S. DAVIS, D.A. WATKINSON, AND M.S. POESCH. 2019. Habitat use and hybridisation between the Rocky Mountain sculpin (*Cottus* sp.) and slimy sculpin (*Cottus cognatus*). *Freshwater Biology* 64:391–404.
- SAGLAM, I.K., D.J. PRINCE, M. MEEK, O.A. ALI, M.R. MILLER, M. PEACOCK, H. NEVILLE, A. GOODBLA, C. MELLISON, W. SOMER, AND B. MAY. 2017. Genomic

- analysis reveals genetic distinctiveness of the Paiute cutthroat trout *Oncorhynchus clarkii seleniris*. *Transactions of the American Fisheries Society* 146: 1291–1302.
- SHAFFER, A.B.A., C.I. CULLINGHAM, S.D. CÔTÉ, AND D.W. COLTMAN. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology* 19: 4589–4621.
- SHELDON, A.L. 1968. Drift, growth, and mortality of juvenile sculpins in Sagehen Creek, California. *Transactions of the American Fisheries Society* 97:495–496.
- SIMMONS, M.P., AND H. OCHOTERENA. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* 49:369–381.
- SMITH, W.L., AND M.S. BUSBY. 2014. Phylogeny and taxonomy of sculpins, sandfishes, and snailfishes (Perciformes: Cottoidei) with comments on the phylogenetic significance of their early-life-history specializations. *Molecular Phylogenetics and Evolution* 79:332–352.
- SMITH, G.R., T.E. DOWLING, K.W. GOBALET, T.S. LUGASKI, D.A. SHIOZAWA, AND R.P. EVANS. 2002. Biogeography and timing of evolutionary events among Great Basin fishes. Great Basin aquatic systems history. Pages 175–234 in R. Hershler, D.B. Madsen, and D.R. Currey, editors, *Great Basin aquatic systems history*. Smithsonian Contributions to the Earth Sciences No. 33. Smithsonian Institution Press, Washington, DC.
- SMITH, G.R., K. SWIRYDCZUK, P.G. KIMMEL, AND B.H. WILKINSON. 1982. Fish biostratigraphy of late Miocene to Pleistocene sediments of the western Snake River Plain, Idaho. Pages 519–542 in B. Bonnichsen and R.M. Breckenridge, editors, *Cenozoic geology of Idaho*. Idaho Bureau of Mines and Geology Bulletin No. 26.
- SOLÍS-LEMUS, C., M. YANG, AND C. ANÉ. 2016. Inconsistency of species tree methods under gene flow. *Systematic Biology* 65:843–851.
- SOUCY, S.M., J. HUANG, AND J.P. GOGARTEN. 2015. Horizontal gene transfer: building the web of life. *Nature Reviews Genetics* 16:472–482.
- SPENCER, J.E., G.R. SMITH, AND T.E. DOWLING. 2008. Middle to late Cenozoic geology, hydrography, and fish evolution in the American Southwest. Pages 279–300 in M.C. Reheis, R. Hershler, and D.M. Miller, editors, *Late Cenozoic drainage history of the southwestern Great Basin and lower Colorado River region: geologic and biotic perspectives*. Geological Society of America Special Paper No. 439.
- STEPHENS, M., N.J. SMITH, AND P. DONNELLY. 2001. A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* 68:978–989.
- STRAUSS, R.E. 1986. Natural hybrids of the freshwater sculpins *Cottus bairdi* and *Cottus cognatus* (Pisces: Cottidae): electrophoretic and morphometric evidence. *American Midland Naturalist* 115:87–105.
- TABOR, R.A., FT. WATERSTRAT, D.W. LANTZ, H.B. BERGE, AND M.C. LIERMANN. 2017. Distribution, density, and size of migratory and fluvial sculpins in relation to barriers in Puget Sound lowland streams. *North American Journal of Fisheries Management* 37:729–742.
- TAYLOR, D.W., AND G.R. SMITH. 1981. Pliocene molluscs and fishes from northeastern California and northwestern Nevada. *University of Michigan Contributions from the Museum of Paleontology* 25:339–413.
- TROTTER, P.C., AND R.J. BEHNKE. 2008. The case for *Humboldtensis*, a subspecies name for the indigenous cutthroat trout (*Oncorhynchus clarkii*) of the Humboldt River, upper Quinn River, and Coyote Basin drainages, Nevada and Oregon. *Western North American Naturalist* 68:58–65.
- VAN TASSELL, J., AND G.R. SMITH. 2019. Keating, Always Welcome Inn, and Imbler fish paleofaunas, NE Oregon: tests of Miocene-Pliocene drainage connections. *Miscellaneous Publications of the Museum of Zoology at the University of Michigan* 204(5).
- WALLACE, R.L., AND D.W. ZAROBAN. 2013. Native fishes of Idaho. *American Fisheries Society*, Bethesda, MD. 213 pp.
- WARD, R.D. 2009. DNA barcode divergence among species and genera of birds and fishes. *Molecular Ecology Resources* 9:1077–1085.
- WEN, D., Y. YU, J. ZHU, AND L. NAKHLEH. 2018. Inferring phylogenetic networks using PhyloNet. *Systematic Biology* 67:735–740.
- WYDOSKI, R.S., AND R.R. WHITNEY. 2003. *Inland fishes of Washington*. University of Washington, Seattle, WA. 322 pp.
- XIA, X. 2018. DAMBE7: new and improved tools for data analysis in molecular biology and evolution. *Molecular Biology and Evolution* 35:1550–1552.
- YOKOYAMA, R., AND A. GOTO. 2005. Evolutionary history of freshwater sculpins, genus *Cottus* (Teleostei; Cottidae) and related taxa, as inferred from mitochondrial DNA phylogeny. *Molecular Phylogenetics and Evolution* 36:654–668.
- YOUNG, M.K., K.S. MCKELVEY, T. JENNINGS, K. CARTER, R. CRONN, E.R. KEELEY, J.L. LOXTERMAN, K.L. PILGRIM, AND M.K. SCHWARTZ. 2018. The phylogeography of Westslope cutthroat trout. Pages 261–301 in P. Trotter, P. Bisson, L. Schultz, and B. Roper, editors, *Cutthroat trout: evolutionary biology and taxonomy*. Special Publication No. 36. American Fisheries Society, Bethesda, MD.
- YOUNG, M.K., K.S. MCKELVEY, K.L. PILGRIM, AND M.K. SCHWARTZ. 2013. DNA barcoding at riverscape scales: assessing biodiversity among fishes of the genus *Cottus* (Teleostei) in northern Rocky Mountain streams. *Molecular Ecology Resources* 13:583–595.
- ZHANG, C., H.A. OGILVIE, A.J. DRUMMOND, AND T. STADLER. 2018. Bayesian inference of species networks from multilocus sequence data. *Molecular Biology and Evolution* 35:504–517.
- ZHANG, J., P. KAPLI, P. PAVLIDIS, AND A. STAMATAKIS. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29:2869–2876.
- ZIMMERMAN, E.G., AND M.C. WOOTEN. 1981. Allozymic variation and natural hybridization in sculpins, *Cottus confusus* and *Cottus cognatus*. *Biochemical Systematics and Ecology* 9:341–346.

Received 3 November 2021
 Accepted 17 December 2021
 Published online 10 June 2022