

Elsevier Editorial System(tm) for Molecular Phylogenetics and Evolution  
Manuscript Draft

Manuscript Number:

Title: Influence of introgression and geological processes on phylogenetic relationships of western North American mountain suckers (*Pantosteus*, *Catostomidae*)

Article Type: Research Paper

Keywords: *Catostomus*; Cryptic species; Fossils; Molecular clock; Mitochondrial DNA

Corresponding Author: Dr Peter J Unmack, Ph.D.

Corresponding Author's Institution: Brigham Young University

First Author: Peter J Unmack, Ph.D.

Order of Authors: Peter J Unmack, Ph.D.; Thomas Dowling; Nina Laitinen; Carol Secor; Richard Mayden; Dennis Shiozawa; Gerald Smith

**Abstract:** Western North America experienced intense geological activity resulting in major topographic changes over the past 15 million years. In addition, most major rivers of western North America are composites of more ancient rivers, resulting in isolation and differential mixing between river basins over time. This, combined with the geological history of the region influenced the diversification of most of the aquatic fauna. The genus *Pantosteus* (mountain suckers, previously considered a subgenus of *Catostomus*) is one of several clades centered in the tectonically active region of western North America. The eight recognized *Pantosteus* species are widespread and common across southwestern Canada, western USA and into northern Mexico. They are typically found in medium gradient, middle-elevation reaches of rivers over rocky substrates. The goal of the present work is to (1) integrate molecular, morphological and paleontological data for the proposed species of *Pantosteus*, (2) test hypotheses of their monophyly, (3) use these data for phylogenetic inferences of sister-group relationships, and (4) estimate timing of divergence events of identified lineages. Molecular data consists of 8055 base pairs from mitochondrial DNA protein coding genes. Multiple fossils are placed within the phylogeny to provide the basis for estimating the timing of the evolution of the group. *Pantosteus* is monophyletic with the exception of *P. columbianus*, whose mtDNA is closely aligned with *Catostomus tahoensis* because of introgression. Within *Pantosteus*, several species have deep genetic divergences among allopatric sister lineages, several of these lineages are diagnosed and elevated to species, bringing to the total diversity in the group to 12 species. Conflicting molecular and morphological data are resolved when patterns of divergence are shown to be correlated with sympatry and the ability to hybridize, with evidence supporting the influence of introgression in both morphological and molecular traits.

Suggested Reviewers: Nathan Lovejoy  
lovejoy@utsc.utoronto.ca  
Good knowledge of phylogenetics and fish biogeography

Hank Bart  
hank@museum.tulane.edu  
good knowledge of phylogenetics, fish biogeography and the family *Catostomidae*

Andrew Martin  
Andrew.Martin-1@Colorado.edu  
good knowledge of phylogenetics, fishes and western North America

Wayne Stearns  
wayne.stearns@ncdenr.gov  
good knowledge of fishes, eastern Catostomids and morphological taxonomy

Opposed Reviewers: Phil Harris  
pharris@bama.ua.edu  
He is a competing researcher in western catostomid fishes

August 18, 2012

D.E. Wildman, Editor  
*Molecular Phylogenetics and Evolution*  
Wayne State University  
Ctr. for Molecular Medicine & Genetics, Michigan, USA

Dear Dr Wildman,

Please consider the following submission, “Influence of introgression and geological processes on phylogenetic relationships of western North American mountain suckers (*Pantosteus*, Catostomidae)” for *Molecular Phylogenetics and Evolution*.

In this paper we examined a widespread and common freshwater fish genus from western North America that we believe will be of interest to readers of *Molecular Phylogenetics and Evolution*. The key strength of this paper is the comprehensive taxonomic and geographic sampling within the region (multiple individuals of all species were sampled) as well as a large number of base pairs sequenced (8055) to provide a robust estimation of their phylogeny. The key highlights are:

- The genus *Pantosteus* was monophyletic, once introgression was taken into account.
- Introgression between sympatric *Pantosteus* species was common.
- *Pantosteus* has an old and complicated biogeographic history due to active geology.
- A fossil calibrated molecular clock gives an age of *Pantosteus* at 24 million years.
- Conservation of Weber, Snake and Bear River *P. virescens* populations is critical.

Please contact me if you have any questions.

Best regards,

Peter Unmack



## Highlights

- The genus *Pantosteus* was monophyletic, once introgression was taken into account.
- Introgression between sympatric *Pantosteus* species was common.
- *Pantosteus* has an old and complicated biogeographic history due to active geology.
- A fossil calibrated molecular clock gives an age of *Pantosteus* at 24 million years.
- Conservation of Weber, Snake and Bear River *P. virescens* populations is critical.

**Influence of introgression and geological processes on phylogenetic relationships of western North American mountain suckers (*Pantosteus*, *Catostomidae*)**

Peter J. Unmack<sup>1,2</sup>, Thomas E. Dowling<sup>3</sup>, Nina J. Laitinen<sup>1</sup>, Carol L. Secor<sup>3</sup>, Richard L. Mayden<sup>4</sup>, Dennis K. Shiozawa<sup>1</sup>, Gerald R. Smith<sup>5</sup>

<sup>1</sup>WIDB 401, Department of Biology, Brigham Young University, Provo UT 84602, USA

<sup>2</sup>National Evolutionary Synthesis Center, 2024 W Main Street, Suite A200, Durham NC 27705-4667, USA

<sup>3</sup>School of Life Sciences, PO Box 874501, Arizona State University, Tempe AZ 85287-4501, USA

<sup>4</sup>Department of Biology, 3507 Laclede Avenue, Saint Louis University, St. Louis MO 63103-2010, USA

<sup>5</sup>Museum of Zoology, The University of Michigan, Ann Arbor MI 48109-1079, USA

Corresponding author. Peter Unmack, National Evolutionary Synthesis Center, 2024 W Main Street, Suite A200, Durham NC 27705-4667, USA, phone: +1 (919) 668-9107, fax: +1 (919) 668-9198, email: [peter.pub@unmack.net](mailto:peter.pub@unmack.net)

## Abstract

Western North America experienced intense geological activity resulting in major topographic changes over the past 15 million years. In addition, most major rivers of western North America are composites of more ancient rivers, resulting in isolation and differential mixing between river basins over time. This, combined with the geological history of the region influenced the diversification of most of the aquatic fauna. The genus *Pantosteus* (mountain suckers, previously considered a subgenus of *Catostomus*) is one of several clades centered in the tectonically active region of western North America. The eight recognized *Pantosteus* species are widespread and common across southwestern Canada, western USA and into northern Mexico. They are typically found in medium gradient, middle-elevation reaches of rivers over rocky substrates. The goal of the present work is to (1) integrate molecular, morphological and paleontological data for the proposed species of *Pantosteus*, (2) test hypotheses of their monophyly, (3) use these data for phylogenetic inferences of sister-group relationships, and (4) estimate timing of divergence events of identified lineages. Molecular data consists of 8055 base pairs from mitochondrial DNA protein coding genes. Multiple fossils are placed within the phylogeny to provide the basis for estimating the timing of the evolution of the group. *Pantosteus* is monophyletic with the exception of *P. columbianus*, whose mtDNA is closely aligned with *Catostomus tahoensis* because of introgression. Within *Pantosteus*, several species have deep genetic divergences among allopatric sister lineages, several of these lineages are diagnosed and elevated to species, bringing to the total diversity in the group to 12 species. Conflicting molecular and morphological data are resolved when patterns of divergence are shown to be correlated with sympatry and the ability to hybridize, with evidence supporting the influence of introgression in both morphological and molecular traits.

Keywords: *Catostomus*, Cryptic species, Fossils, Molecular clock, Mitochondrial DNA

## 1. Introduction

Western North America has experienced intense geological activity resulting in major topographic changes over the past 15 million years, during which time the modern freshwater fish fauna of the region evolved (Cavender, 1986; Spencer et al., 2008). The continent had moved westward over subducting oceanic plates, the Sierra Nevada and Rocky Mountains were elevated, and the crust between them extended by as much as 100 percent, causing elongate basins 1–3 km below several hundred major north-south mountain ranges to subside (Faulds and Varga, 1998; Chamberlain et al., 2012). Changes in elevation across this terrain resulted in hundreds of short mountain streams often ending in closed-basins, with their configuration and variable connections over time influencing the diversification of lineages of many aquatic and riparian organisms (Hershler and Sada, 2002; Smith et al., 2002).

The evolution of basin and range topography has been associated with the isolation and diversification of most of the aquatic fauna (Hubbs and Miller, 1948; Taylor, 1985). Tectonism and volcanism played a major role in modifying drainage patterns as well as forming several large, long-lived lakes such as Miocene Chalk Hills Lake (Kimmel, 1975) and Pliocene Glenns Ferry Lake in southwest Idaho (Smith et al., 1982), late Miocene Hopi Lake in northwest Arizona (Dallege et al., 2008), Pleistocene Lake Bonneville in Utah (Currey, 1990; Oviatt, 2002; Sack, 2002), Pleistocene Lake Lahontan and associated lakes in Nevada (Reheis et al., 2002) and southern Oregon (Negrini, 2002), Pleistocene Tulare Lake, California (Davis, 1999), and many other lakes in the Great Basin (Hubbs and Miller, 1948; Benson, 2004). Several lakes were as large as some of the Laurentian Great Lakes and, similarly, hosted the assembly and evolution of distinctive regional faunas. Isolated individual basins were the original crucibles of local evolution of subspecies and species of aquatic organisms in the Miocene and Pliocene.

Today, the major rivers of western North America are largely composites of more ancient rivers. East of the Great Basin, the Colorado River flows southwestward across the Colorado Plateau; the Upper and Lower Colorado basins (defined as the reaches above and below the Grand Wash Cliffs, where the river leaves the Colorado Plateau and enters the Great Basin) each captured substantial parts of the Great Basin and its fish fauna. The upper and lower basins became connected through Grand Canyon in the Early Pliocene (Howard et al., 2008), creating one of the largest drainages and faunas in the west. North of the Great Basin, the Snake River flows across southern Idaho to Hells Canyon, a route established when it was captured in the late Pliocene by the Salmon River (Columbia Basin) (Wheeler and Cook, 1954; Smith et al., 2000). The Bear River flows north from the Uinta Mountains, turns west and south in southeast Idaho, where it was diverted into the Bonneville Basin about 100 Ka (Bouchard et al., 1998), which eventually caused the Bonneville Basin, Utah, to overflow and become tributary to the Upper Snake River in Idaho (the Upper and Lower Snake River are separated at Shoshone Falls). In the Miocene the Lahontan Basin, Nevada, was a tributary to the Pacific Ocean via California's Great Valley (Smith et al., 2002), and overflowing to the Upper Snake River several times in the Plio-Pleistocene (Reheis et al., 2002). Uplift of the Sierra Nevada and rise of the Cascade Mountains created climatic fluctuations across the Great Basin, causing variable environmental conditions (Retallack, 2004; Kohn and Fremd, 2008) that influenced historical drainage connections. Large lakes, sometimes trapped in closed basins in southeast California, western Nevada, western Utah and southern Oregon, for example, had fluctuating connections that created periodic sympatry and isolation of fish populations (Hubbs and Miller, 1948; Reheis et al., 2002). This complex geological and environmental setting is hypothesized to have been fundamental in causing a reticulating pattern of faunal isolation used to explain the morphological evolution of many local populations described by ichthyologists in the past 130 years (see below).

The geological and climatic history of the west also resulted in high extinction rates, resulting in a depauperate fish fauna relative to other areas of North America (Smith et al., 2010). Currently restricted basins have fewer than 25 native fishes and high relictual endemism, typically between 25 and 74% (Abell et al., 2000). The Rocky Mountains and associated mountain ranges of the continental divide (Spencer et al., 2008) have largely isolated western rivers from more speciose drainages in the tectonically stable eastern USA. Some western fish clades have Paleogene origins and represent endemic lineages whose evolution is consistent with the active geologic history of this region. Examples include the Plagopterini (Dowling et al., 2002), western clades in the Catostomini (Smith, 1992b), Cyprinidae (Dowling et al., 2002; Bufalino and Mayden, 2010), Cyprinodontidae (Echelle et al., 2005), Goodeidae (Parenti, 1981; Webb et al., 2004) and *Cottus* (Kinzinger et al., 2005). Additionally, evidence has been provided for frequent episodes of introgression of some fishes (Smith, 1966, 1992a), loss of genotypes (Dowling and DeMarais, 1993; Gerber et al., 2001) and the evolution of species of hybrid origin (DeMarais et al., 1992; Gerber et al., 2001).

The genus *Pantosteus* (mountain suckers, previously considered a subgenus of *Catostomus*) is one of several clades centered in the tectonically active part of western North America. Species of this lineage occur from southwestern Canada, western USA and into northern Mexico, inhabiting middle-elevation reaches of rivers of the Basin and Range, Coast Ranges, Cascade Mountains, Rocky Mountains, east to the Black Hills and Sierra Madres (Smith, 1966; Fig. 1). These cool-water, benthic fishes were investigated by early western American ichthyologists, from Cope (1872) to Snyder (1924). Species of *Pantosteus* were described by Baird and Girard (1854), Cope (1872, 1874), Cope and Yarrow (1875), Jordan (1878), Jordan and Gilbert (1882), Jordan and Evermann (1893), Eigenmann and Eigenmann (1893), Garman (1881), Rutter (1903), Snyder (1915, 1924) and Hubbs and Schultz (1938).

Hubbs et al. (1943) summarized the previous taxonomic history of this group and recognized the unusual frequency of hybridization in this group.

Smith (1966) evaluated previous works, added morphological and meristic data for 558 population samples across the range of *Pantosteus*, and generated a classification emphasizing relationships among populations of six polytypic species (distributions in Fig. 1). Four southern forms included: 1) *P. plebeius* in the Rio Grande Basin and its major tributary the Rio Conchos, the Rio Nazas and several Pacific basins in Mexico; 2) *P. santaanae* in the Los Angeles Basin; 3) *P. clarkii* in the Lower Colorado Basin below Grand Canyon, including the Pluvial White, Virgin, Bill Williams and Gila basins; and 4) *P. discobolus* in the Upper Colorado Basin, Upper Snake River and northern Bonneville Basin. The northern group was comprised of *P. platyrhynchus* in the Great Basin, Columbia-Snake, Fraser, upper Saskatchewan, Missouri and Green drainages; and *P. columbianus* in the Columbia-Snake drainage. All species except *P. santaanae* exhibit geographic variation, with subgroups that include morphologically differentiated populations in adjacent drainages. Some species contain variants recognized as subspecies, e.g., *P. discobolus jarrovii* in the Zuni River, New Mexico, and other drainages in northeast Arizona (Smith et al., 1983), *P. columbianus hubbsi*, in the Wood River drainage, Idaho (Smith, 1966); and *P. columbianus palousianus* in the Palouse River, Washington (Smith, 1966). An additional species, *P. nebuliferus*, from the Nazas Basin, was recognized by Miller et al. (2005).

Confusion important to generic classification surrounded *P. columbianus*, an intermediate form in the Columbia River. Eigenmann and Eigenmann (1893) described *P. columbianus* based on its several diagnostic mountain sucker traits; Hubbs and Schultz (1938) described it again as *Catostomus syncheilus* based on its *Catostomus*-like lips. Smith (1966) recognized *Pantosteus* as a subgenus of *Catostomus* when he interpreted *C. syncheilus* to be *P. columbianus* based on osteology, with this taxon exhibiting a mixture of traits

diagnostic for *Catostomus* and *Pantosteus*. This mixture of traits in *P. columbianus* was hypothesized to be the result of this taxon either being a plesiomorphic intermediate form or a product of introgressive hybridization (Smith, 1966, 1992a). Smith and Koehn (1971) conducted a phylogenetic analysis of 16 species of *Catostomus* and *Pantosteus* using both morphological and biochemical data, yielding two general groups in *Catostomus*: 1) a lineage with large, low-elevation, slow-water species in a paraphyletic subgenus *Catostomus*, and 2) a lineage of smaller mountain suckers adapted to cooler and swifter rivers at higher elevations. One study has used DNA sequences to examine broader relationships among catostomine species. Doosey et al. (2009) examined six species of *Pantosteus* and recovered five of them in a monophyletic group nested within *Catostomus*. The only species of *Pantosteus* excluded from this lineage was *P. columbianus*.

The above studies illustrate how complex distribution of traits can impact the ability to delimit taxa and interpret their evolutionary origins. This is particularly the case for *Pantosteus* and *Catostomus* as indicated by the proposed paraphyly of some taxa, potential influence of introgression on the interpretation of character evolution, highlighting how this process can make accurate reconstruction of phylogenetic relationships difficult (Smith, 1992a). The goal of the present work is to 1) integrate molecular, morphological and paleontological data for the proposed species of *Pantosteus*, 2) test hypotheses of their monophyly, 3) use these data for phylogenetic inferences of sister-group relationships, and 4) estimate timing of divergence events of identified lineages. With these insights into reconstructed evolutionary history of the group, we evaluate the variation and distribution of morphological traits found in *Pantosteus* lineages to revise the taxonomy of the genus and the impact of geological processes on biodiversity in this group.

## **2. Materials and Methods**

## 2.1. Study taxa and sampling

We conducted range-wide molecular surveys of each species and selected individuals that represented both geographic and genetic diversity within the genus (Table 1). Outgroup taxa to represent the broader diversity within Catostomidae (Smith, 1992b; Doosey et al., 2009) include: *Catostomus ardens*, *C. catostomus*, *C. commersonii* (GenBank accession AB127394.1), *C. insignis*, *C. macrocheilus*, *C. tahoensis*, *Cycleptus elongatus* (AB126082.1), *Hypentelium nigricans* (AB242169.1), *Minytrema melanops* (DQ536432.1), *Moxostoma erythrurum* and *M. poecilurum* (AB242167.1) (Table 1).

## 2.2. DNA isolation, amplification, and sequencing

Genomic DNA was extracted from muscle tissue using the DNeasy Tissue Kit (QIAGEN Inc., Chatsworth CA) or by phenol-chloroform extraction as described in Tibbets and Dowling (1996). Nine of the 13 mtDNA protein coding genes (ND1, ND2, ND4L, ND4, ND5, ND6, ATPase6/8, *cytb* and partial sequence from COIII) were amplified, representing approximately half of the mitochondrial genome (8055 bp). Single and nested PCR amplification strategies were used to obtain product for different gene combinations. Details of the primers and nesting combinations are in Supplementary Figure 1. For nested PCR the first reaction size was 10  $\mu$ L. This first PCR reaction was then diluted to 1:49, and 1  $\mu$ L of this product was added to a second 25  $\mu$ L reaction. All other single reactions were 25  $\mu$ L. Final concentrations for PCR components were as follows: 25 ng template DNA, 0.25  $\mu$ M of each primer, 0.625 units of Taq DNA polymerase, 0.1 mM of each dNTP, 2.5  $\mu$ L of 10X reaction buffer and 2.5 mM MgCl<sub>2</sub>. Amplification parameters were as follows: 94°C for 2 min followed by 35 cycles of 94°C for 30 s, 48°C for 30 s, and 72°C for 60 s (in the first nested reactions this was increased by 1 min per each thousand bp), and 72°C for 7 min. PCR products were examined on a 1% agarose gel using SYBR safe DNA gel stain (Invitrogen,

Eugene, OR, USA) and purified using a Montage PCR 96 plate (Millipore, Billerica, MA, USA). Sequences were obtained via cycle sequencing with Big Dye 3.0 dye terminator ready reaction kits using 1/16th reaction size (Applied Biosystems, Foster City, CA, USA). Sequencing reactions were run with an annealing temperature of 52°C following the ABI manufacturer's protocol), cleaned using Sephadex columns in MultiScreen 96 well assay plates (Millipore, Billerica, MA, USA), and then dried. Most sequences were obtained using an Applied Biosystems 3730 XL automated sequencer at the Brigham Young University DNA Sequencing Center. All sequences obtained in this study were deposited in GenBank, accession numbers XXXXXX–XXXXXX and the sequence alignment was deposited in Dryad, doi:XXXXX/dryad.XXXXXX.

### *2.3. Analysis of sequence data*

Sequences were edited using Chromas Lite 2.0 (Technelysium, Tewantin, Queensland, Australia) and imported into BioEdit 7.0.5.2 (Hall, 1999). Sequences coding for amino acids were aligned by eye and checked via amino acid coding in MEGA 4.0.2 (Tamura et al., 2007) to test for unexpected frame shift errors or stop codons. Editing resulted in 8055 base pairs representing the complete sequence for the nine genes, plus 22 bp of COIII. Phylogenetic analyses were performed under maximum likelihood (ML) using RAxML 7.2.8 (Stamatakis, 2006, Stamatakis et al., 2008) by bootstrapping with 1000 replicates with the final best ML tree being calculated using the GTRGAMMA model on the CIPRES cluster at the San Diego Supercomputer Center. Maximum parsimony (MP) analysis was conducted with PAUP\* 4.0b10 (Swofford, 2003) using a heuristic search with 1000 random additions and TBR branch swapping. Tree lengths reported for MP include both informative and uninformative characters. Robustness of nodes for MP was estimated by bootstrapping with 1000 replicates using a heuristic search with 10 random additions and TBR branch swapping.

The trees from DNA analyses presented in this study were deposited in TreeBASE, accession number XXX:XXXXXX, (<http://purl.org/phylo/treebase/phylovs/study/XXX:XXXXXX>).

Average between-species genetic distances were calculated based on the proportion of shared differences (p-distance) using MEGA for each lineage within *Pantosteus*, *Catostomus* as a whole, and then the remaining outgroups as a whole.

#### 2.4. Morphology and Paleontology

External anatomical and meristic characters were evaluated on specimens in the Museum of Zoology at the University of Michigan (UMMZ) as summarized in Smith (1966), using the methods of Hubbs and Lagler (2004). Osteological comparisons are based on dried skeletons and cleared-and-stained specimens in the UMMZ. Paleontological material and data are from the University of Michigan Museum of Paleontology (UMMP) and the University of Kansas Museum of Natural History (described in Smith et al., submitted). Morphological characters are somewhat biased toward features useful in diagnosing lineages identified by molecular studies. Osteological characters are selected from features useful in identifying fossil bones. Meristic characters are taken from measurements of fin-ray, scale and gill-raker counts; morphometric characters are represented by measurement of caudal peduncle depth, isthmus width and jaw width (Smith, 1966; Miller et al., 2005). Pigment differences are important to some diagnoses (Smith, 1966; Miller et al., 2005).

#### 2.5. Comparison of data sets

Molecular, morphological, and fossil data sets answer questions differently because they differ in sample sizes, character independence, and confidence estimates. Similarly, molecular data sets are so powerful because they include thousands of quasi-independent characters, measured essentially without error. The number of morphological characters are

statistically limited, but with larger sample sizes. Fossil characters are drawn from statistically inadequate numbers of fragments, limited in anatomical breadth, and with potentially large errors in estimated time and placement on trees (Spencer et al., 2008). These data sets require separate analyses. Conflicts among inferences from different data sets are analyzed by considering the nature of the genetic basis and the evidence for gene flow among populations (reviewed in Arnold, 1997); these influence characters in ways not resolvable by congruence of character code values. Inclusion in a single data matrix for total evidence analysis would obscure results unique to each kind of character history.

Calibration of the rate of evolution requires fossils to be assigned to specific branches. This is achieved by cladistic interpretation of characters for each fossil and considering it as a terminal taxon, thus allowing placement of the fossil on a specific branch (Smith et al., submitted). Applying confidence limits to estimates of times of cladistic branching has two primary sources of error: fossils are unlikely to represent a lineage soon after its divergence from its sister lineage (Marshall, 1990), and both sister lineages do not necessarily possess synapomorphies of the lineages immediately after their initial genetic isolation. Fossils of both lineages should ideally contribute to estimation of age of a node, but both are rarely available in the record. For these reasons it is not obvious which end of the stem branch of a crown group should be assigned the age represented by the fossil. We choose to correct the age of the oldest fossil bearing synapomorphies of a lineage in question with the method of Marshall (1990) which uses the density of the record of other fossils from that lineage to estimate the stem branch. The number of million-year time horizons or independent localities in which fossils of the lineage occur are applied in an equation that uses these data to estimate the probability that the oldest fossil represents the first (unobserved) appearance of the lineage. An important assumption is that the fish of interest could have lived in a depositional environment below the oldest fossil, so that the absence of fossils possessing synapomorphies

of a specific lineage can be interpreted as absence of the lineage at that time. The probability that the earliest fossil occurrence represents the first occurrence following cladogenesis is then used to correct the age estimate. The corrected age is applied to the node joining the taxa bearing the morphological synapomorphies supporting their sister-group relationship (Fig. 2).

BEAST 1.7.1 (Drummond and Rambaut, 2007) was used to estimate molecular divergence times of lineages based on Marshall fossil age estimates. We generated input files using BEAUti 1.7.1. The dataset was trimmed to single representatives per species/lineage because having a mix of within- and between-species data complicates dating owing to different processes for estimating within- versus between-species rates (Ho et al., 2008; S. Ho., pers. comm.). The analysis used an uncorrelated lognormal relaxed molecular clock with rate variation following a tree prior using the speciation birth-death process, a GTR+I+G model (identified using the AIC in Modeltest 3.7, Posada and Crandall, 1998). The topology was constrained to match the ML results.

Calibrations were based on seven time horizons with fossils (Smith et al., submitted), with age of occurrence estimated for two nodes with two date estimates in each, and a log normal prior with a standard deviation of 2. BEAST analyses were run for 50 million generations, with parameters logged every 5000 generations. Multiple runs were conducted to check for stationarity and to ensure that independent runs were converging on a similar result. The log and tree files from four runs were combined in LogCombiner 1.7.1 with a 10% burn-in. The combined logfile was examined in Tracer 1.5, while the combined treefile was summarized using TreeAnnotator 1.7.1 with the mean values placed on the maximum clade credibility tree.

### **3. Results and Discussion**

#### *3.1. Phylogenetic analyses using mtDNA sequences*

Sequence analysis of 58 specimens (Table 1) yielded 4777 invariant characters, 507 variable but parsimony uninformative characters, and 2771 parsimony informative characters. Maximum-likelihood analysis recovered one tree with a likelihood score of -60385.098771 (Fig. 2). Phylogenetic analysis of mtDNA sequences provided resolution among previously-recognized species (Smith, 1966) as well as identification of divergent lineages within four of the species. *Pantosteus* was monophyletic with the exception of *P. columbianus*, whose mtDNA was closely aligned with *C. tahoensis* a representative of the genus *Catostomus*. Both maximum likelihood and maximum parsimony methods recovered the same major lineages within *Pantosteus*, but with some different branching sequences among regional populations. In general, bootstrap analyses provide strong support (> 95%) for most of the deeper nodes (Fig. 2). In the following summary we only provide bootstrap values when support was less than strong (< 95%). Mean p-distance between these 13 major lineages within *Pantosteus* varied between 2.2% and 10.3% (Supplementary Table 1), while sister lineages varied between 2.2% and 8.3%. Outgroup lineages differed from *Pantosteus* by 11.4–19.2% pairwise sequence divergence except for the small distance between *Catostomus tahoensis* and *Pantosteus columbianus* (Supplementary Table 1).

The earliest separation within *Pantosteus* is represented by a clade (formerly *P. platyrhynchus*) comprising three groups of populations in the Columbia drainage, Missouri drainage and Lahontan Basin whose relationship is (Missouri (Lahontan, Columbia)). Sister to that lineage is a clade including all other *Pantosteus* species. *Pantosteus platyrhynchus* mtDNA is polyphyletic in the tree, with Bonneville Basin and Upper Snake River samples falling within a clade that contains *P. discobolus* (discussed below).

The next oldest lineage in *Pantosteus* mtDNA include samples of southern forms (primarily from Mexico) that extend from the Nazas Basin north into the Rio Grande Basin and its past connectives. In the ML tree, a lineage comprised of *P. nebuliferus* and *P. plebeius*

from the Miravalles drainage was sister to a well-supported group of *P. plebeius* (found in 93% of bootstrap replicates) from southern (Rios Conchos and Fuerte) and northern (Rio Grande and Guzman basins) lineages. In addition, the haplotype from the Rio Grande was part of the introgressed *P. discobolus jarrovii* populations from the upper Little Colorado River (Smith et al., 1983).

The remaining *Pantosteus* species form a clade herein identified as the *P. discobolus* group. This group consists of: *P. clarkii* from the Lower Colorado Basin, *P. santaanae* in the Los Angeles area, *P. platyrhynchus* from the Bonneville Basin and Upper Snake River and *Pantosteus discobolus* from the Upper Colorado Basin, Weber, Bear and Upper Snake rivers. The *P. clarkii*–*P. santaanae* clade is sister to *P. discobolus* and *P. platyrhynchus*. In this latter group, samples of *P. discobolus* from the Weber, Bear and Upper Snake rivers were sister to *P. discobolus* (Upper Colorado Basin) and *P. platyrhynchus* (Bonneville Basin). The high estimates of sequence divergence within *P. discobolus* suggest re-evaluation of the morphology of the Bonneville-Upper Snake populations, and recognition of these as *P. virescens* (see below).

Morphological traits of individuals from *Pantosteus* populations in the upper Green River place these populations as sister to *P. platyrhynchus* from the Bonneville and Upper Snake drainage; however, individuals also possess morphological traits of *P. discobolus* consistent with introgression between these species. The presence of *P. discobolus* mtDNA in *P. platyrhynchus* renders the mtDNA lineages of these species paraphyletic (Fig. 2). Samples of *P. platyrhynchus* from the Bonneville Basin were subdivided into well-differentiated Sevier River and northern Bonneville-Upper Snake River lineages.

Maximum parsimony analysis, with all characters weighted equally, recovered a single most parsimonious tree with a length of 11,151 (CI = 0.398, RI = 0.727). Two major topological differences exist between ML and MP trees; these are highlighted on a reduced

MP tree (inset in Fig. 2). Among *P. plebeius* and relatives, the MP tree differs from the ML tree with *P. nebuliferus* and *P. plebeius* from the Rio Miravalles sister to the northern lineage of *P. plebeius* (found in only 63% of MP bootstrap replicates), with a southern group of *P. plebeius* sister to *P. nebuliferus* and Rio Grande–Guzman Basin lineages. The second difference in the MP tree involves species from the Bonneville, Colorado and Los Angeles basins, where *P. clarkii* rather than *P. santaanae* is the sister to the clade that includes all *P. discobolus* plus *P. platyrhynchus* from the Bonneville Basin. This placement of *P. clarkii* is found in 100% of the MP bootstrap replicates.

### 3.2. Morphology of *Pantosteus*

*Pantosteus* Cope is a monophyletic group based on osteological and morphological characters. The genus is diagnosed by: dentaries with their anterior, biting distal process turned mesially at a near 90° angle from the proximal part of the bone (the angle is obtuse in *Catostomus*); the maxilla has a low dorsal flange and a prominent anterior flange; the hyomandibula has enlarged anterior and posterior flanges (much less developed in *Catostomus*); the hyomandibula also has an extensive fossa and associated processes (absent in *Catostomus*) at the post dorsal point of the bone—on the pterotic condyle—connecting with the prominent ventro-lateral process of the pterotic; the dorsolateral ridge of the pterotic is nearly vertical and usually sharp dorsally (not expanded into a broad horizontal roofing bone as in *Catostomus*); the preopercle is deep at the center, approaching half-moon shape in *Pantosteus* (but is slender in *Catostomus*); the lips are large and fleshy with large notches at the junction of the upper and lower lips, and four or more rows of papillae between the median notch of the lower jaw; and possesses a prominent cartilaginous ridge of the lower jaw. *Pantosteus columbianus* is intermediate between *Pantosteus* and *Catostomus* in many of

the above characters; *P. plebeius* and *P. nebuliferus* are intermediate in the lip characters (Smith, 1966; Smith et al., submitted).

Patterns of mtDNA variation among taxonomic lineages as previously constituted (Smith, 1966) are not consistent with the division of northern versus southern lineages, monophyly of the genera, or monophyly of species such as *P. platyrhynchus*, *P. discobolus* or *P. columbianus*. The length of internodal branches (and concomitant bootstrap) defining each of the mtDNA lineages rendering these taxa as nonmonophyletic are so large, sorting of mtDNA lineages is not a likely explanation for these incongruent results. Therefore, this conflict could be the result of convergence in morphological traits or introgressive transfer of mtDNA (Smith, 1992a). Both processes were hypothesized in previous comparative studies to explain similar patterns of morphological and biochemical characters (Smith, 1966; Smith and Koehn, 1971; Smith et al., 1983). Evaluation of these conflicts involves the use of fossils to provide historical perspective, comparison with cases of present-day introgressive hybridization and statistical evaluation of phylogenetic trees. Each lineage is analyzed below in the context of these potential lines of evidence.

Regardless of the reasons for incongruence, major revisions in the current taxonomy are required. Given the complicated nomenclatorial history of this group, we provide a summary of names that have been used as well as morphological support for these designations. This is presented in the context of mtDNA and morphological phylogenies (Figs. 2 and 4), which reveal examples of paraphyly involving *Catostomus* and *Pantosteus* as well as some of the lineages within *Pantosteus*.

### 3.3. Interpretation and diagnosis of species

Morphological and molecular results are mostly consistent for several species (*P. clarkii*, *P. santaanae*, *P. plebeius* and *P. nebuliferus*). Diagnoses for each of these species are

presented below. The remaining lineages, *P. columbianus* and those formerly referred to as *P. discobolus* and *P. platyrhynchus*, exhibit conflict between morphological and molecular traits that are considered in detail along with hypotheses to explain the complicated polymorphisms and mixtures of traits among populations.

*Pantosteus clarkii* (Baird and Girard) inhabits tributaries to the Lower Colorado Basin, including the Gila, Salt, Bill Williams, Pluvial White and Virgin rivers, which were connected to the Lower Colorado Basin below Grand Wash, but isolated from the Upper Colorado Basin and *P. discobolus* prior to the connecting through-flow (i.e., Grand Canyon) at about 5.5 Ma (Howard et al., 2008). *Pantosteus clarkii* shares most of its morphological traits with *P. santaanae*, *P. discobolus* and *P. virescens*, being distinguished from those by its large predorsal scales, usually 15–45. The predorsal scale count and caudal peduncle depth are variable among populations, leading Tanner (1932, 1942) to name *Notolepidomyzon utahensis* and *N. intermedius*, but morphological (Smith, 1966) and mtDNA differences are slight and these taxa are not recognized. MtDNA data presented here support the hypothesis that isolation of these populations is relatively recent. Examples of introgression between *P. clarkii* and *P. discobolus* and between *P. clarkii* and *P. platyrhynchus* have been found in samples from the Virgin River and Shoal Creek (Sevier drainage), respectively (Smith, 1966; Secor and Dowling, unpub. data).

*Pantosteus santaanae* Snyder is a generalized, small-bodied mountain sucker found on the Los Angeles Plain, geographically isolated from *P. clarkii* in the Lower Colorado Basin. *Pantosteus santaanae* is diagnosed by the combination of its thicker caudal peduncle (exceeding 9% of standard length), 42 post-Weberian vertebrae, 27–41 predorsal scales, and 21–28 gill rakers on the outer row of the first arch (Smith, 1966). Osteologically, it shares features with *P. clarkii* consistent the ML tree; however, these are plesiomorphic, consistent with its position in the MP tree.

*Pantosteus plebeius* (Baird and Girard) is found in the Rio Grande Basin and associated internal drainages, as well as headwaters of adjacent Pacific coastal drainages in Mexico. It and its sister species, *P. nebuliferus*, comprise the sister group to the *discobolus*–*virescens*–*clarkii*–*santaanae* clade, characterized by a robust ridge anterior to the opercular condyle and continuous with the posterior crest of the hyomandibula. *Pantosteus plebeius* shares with *P. nebuliferus* a rounded lower jaw, with lower lip deeply incised, and nine dorsal rays, but it is diagnosed from *P. nebuliferus* by three or four rows of papillae crossing the midline of the lower lip, lips not projecting from the lower profile of the head, and a narrow isthmus, its mean less than 8% of standard length (Miller et al., 2005). Analysis of allozymes also identifies another distinct lineage in the Rio Mesquital, Mexico that was not sampled here, therefore, additional sampling and characterization is required (Ferris et al., 1982).

*Pantosteus nebuliferus* (Garman) is found in the Nazas and Aguanaval basins, although mtDNA of this form is also found in the population in the Rio Miravalles of the Rio Piaxtla drainage. This suggests that *P. nebuliferus* are either more widespread than first thought, or populations in the Rio Piaxtla are introgressed. It is diagnosed from *P. plebeius* by a less deeply incised lower lip with about five rows of papillae crossing the midline, the lips distinctly projecting below the ventral profile of the head, and a broad isthmus, the mean width exceeding 8% of standard length (Miller et al., 2005).

*Pantosteus discobolus* (Cope) was described from the Green River in the Upper Colorado Basin. The divergent mtDNA data of the Bonneville Basin and Upper Snake River forms, previously included in *P. discobolus*, is consistent with some unique morphological features. These differences suggest recognition under the older name, *P. virescens* Cope. Therefore, we follow Snyder (1924) in separating *P. virescens* (Weber, Bear and Upper Snake rivers) from *P. discobolus* of the Upper Colorado Basin (see below). *Pantosteus discobolus*, *P. virescens*, *P. clarkii*, *P. santaanae* and *P. plebeius* form a clade diagnosed by

mtDNA as well as morphological traits—an extensive post-median flange that overlaps the antero-dorsal edge of the preopercle. This flange is smaller in *P. columbianus* and the *P. platyrhynchus* group, and absent in *Catostomus* (Smith et al., submitted). As constituted here, *P. discobolus* is the large mountain sucker (adults 200–350 mm standard length) of the larger rivers in the Upper Colorado Basin. It is characterized by a slender caudal peduncle, 4–8% of standard length; usually more than 50 predorsal scales; gill rakers 28–44 on the external row of the first arch and 37–60 on the internal row of the first arch; post-Weberian vertebrae usually 43–45; large jaws and lips; and much reduced swim bladder (Smith, 1966).

*Pantosteus discobolus* is replaced in headwater tributaries of the Green River by smaller fishes, less than 180 mm standard length, representing *P. platyrhynchus* of the Bonneville Basin and Upper Snake River. *Pantosteus discobolus* morphological traits and *P. discobolus* haplotypes occur in *P. platyrhynchus* of the Green River (Smith, 1966). Further sampling is required to determine patterns of mtDNA introgression in these drainages.

*Pantosteus virescens* (Cope in Cope and Yarrow, 1875) was described based on a type specimen said to be from the San Juan drainage in Colorado, but Snyder (1924) concluded that the type locality was mislabeled and *P. virescens* was identical to specimens that he collected from the Weber River, Utah, and the Bear River, Wyoming, both tributary to the northern Bonneville Basin. This species is diagnosed by the combination of small scale size, numerous gill-rakers, and offset hyomandibular condyles, which are shared with *P. discobolus*, but with a thicker caudal peduncle, 8–10% of standard length (Smith, 1966). The distribution of *P. virescens* is consistent with that of other species in the Bonneville Basin and Upper Snake River: multiple connections existed between the Bonneville Basin and the Upper Snake River in the late Pleistocene with several fish species being exchanged at that time (Hubbs and Miller, 1948). Similarly, stream captures between the upper Green River and the Upper Snake River have been postulated based on geological (Hanson, 1985; Baranek et al.,

2006) and fish evidence (Spencer et al., 2008), yet the upper Green River populations have mtDNA of *P. discobolus*, despite possible influence from *P. virescens* and *P. platyrhynchus*. *Pantosteus virescens* was formerly abundant as the largest of the mountain suckers in the Weber, Bear, and Upper Snake rivers, but it is rare or absent in most former habitats, and modern specimens are smaller in size, especially in introgressed populations. Populations of *P. virescens* in the Upper Snake River drainage morphologically introgress into *P. platyrhynchus* (p. 108 and Figure 22 of Smith, 1966). Thus, the small extant *P. virescens* populations present in the Weber and Bear rivers should be provided high conservation significance as they represent the only remaining populations of this dwindling species.

We assign *Pantosteus discobolus jarrovii* (Cope) informally to a complex of ancient Arizona and New Mexico lineages that share mixed molecular, meristic and osteological characteristics of *P. discobolus* and *P. plebeius*. They are found in headwaters of the San Juan and Little Colorado rivers in central and northeastern Arizona and southwestern New Mexico. The populations possess molecular, morphological and allozyme traits that originated in three separate but adjacent drainages, but which have been variously transferred in the Pleistocene via stream captures (Smith, 1966; Smith et al., 1983; Crabtree and Buth, 1987). The populations are diverse and threatened with extinction. Their complex patterns of introgressed molecular, morphological and osteological characters are being described elsewhere to deal with controversial but critical conservation issues (Dowling et al., unpublished data).

*Pantosteus platyrhynchus* (Cope) as previously recognized is widely distributed throughout the Lahontan and Bonneville basins, Columbia and Snake rivers, and the headwaters of the Missouri drainage and Upper Colorado Basin; however, high mtDNA divergence among several of these regions require taxonomic changes. Former populations of *P. platyrhynchus* in the Missouri drainage are again referred to *P. jordani* Evermann; those

from the Lahontan Basin are recognized as *P. lahontan* Rutter; and those from the Columbia-Lower Snake rivers represent a previously undescribed species, *P. sp.*

*Pantosteus platyrhynchus*, as here diagnosed, occupies the Bonneville drainage and its former connective, the Upper Snake river, as well as small headwater streams in the Green River. The mtDNA of *P. platyrhynchus* is derived from *P. discobolus* (Fig. 2). This reflects the non-congruence of morphological traits which are otherwise allied to northern populations in the Columbia River, Lahontan Basin, and Missouri drainage rather than *P. discobolus*. In addition, *P. platyrhynchus* contains two forms of mtDNA haplotypes diverged after acquisition of *P. discobolus* mtDNA with consistent slight morphological differences. One form occurs in the Sevier River (southern Bonneville Basin) and the other in the northern Bonneville Basin and Upper Snake River. *Pantosteus platyrhynchus (sensu stricto)* has a hyomandibula about 34% as broad as long, with an incomplete ridge anterior to the opercular condyle, 36–53 predorsal scales, usually 40–43 post-Weberian vertebrae, and 24–34 gill rakers on the outer row of the first arch.

There is evidence for introgression between *P. platyrhynchus* and sympatric *Pantosteus*. Populations in the Upper Snake and Green rivers are included in *P. platyrhynchus* based on the majority of their morphological traits; however, they have *P. virescens* and *P. discobolus* mtDNA, respectively, and mixtures of morphological traits of those species (Smith, 1966). Broader sampling across this region is required to determine whether mtDNA introgression is partial or complete. The geographic pattern of variation in the number of predorsal scales suggests introgression between populations of *P. platyrhynchus* in Shoal Creek (Bonneville Basin) and *P. clarkii* introduced from the Virgin River (Smith, 1966; Secor and Dowling, unpublished data).

*Pantosteus jordani* Evermann is a small-bodied species from the upper Missouri drainage from the Black Hills of South Dakota to western Wyoming, Montana and Alberta. It

is similar to *P. platyrhynchus*, with which it was formerly combined, but has a moderately slender hyomandibula, 34% as broad as long. It has a single post-dorsal process on the hyomandibula, 41–64 predorsal scales, usually 41–43 post-Weberian vertebrae, and 24–37 gill rakers on the outer row of the first gill arch (Smith, 1966).

*Pantosteus lahontan* Rutter occupies the Lahontan Basin of Nevada, Oregon and California. It is similar to *P. platyrhynchus*, but has a slenderer hyomandibula, 30% as broad as long, with a cup-like post dorsal process surrounding the post-dorsal fossa of the hyomandibula. It has 42–47 predorsal scales, usually 39–42 post-Weberian vertebrae, and 25–35 gill rakers on the outer row of the first arch (Smith, 1966).

*Pantosteus* sp. occurs in the Columbia drainage, based upon samples from the Willamette, Similkameen, Boise and Salmon Falls drainages. This species has formerly been referred to as either *P. jordani* or *P. platyrhynchus* (see detailed synonymy in Smith, 1966). It is diagnosed from *P. platyrhynchus* by its reciprocally monophyletic mtDNA and by its combination of a continuous but slightly emarginate ridge anterior to the opercular condyle of the hyomandibula, prominent post-dorsal fossae and process at the post-dorsal tip of the pterotic process of the hyomandibula, prominent post-mesial flange of the hyomandibula overlapping the antero-dorsal tip of the preopercle, medium width of the hyomandibula (about 34% as wide as long) and 29–37 gill rakers in the external row on the first gill arch.

*Pantosteus columbianus* Eigenmann and Eigenmann has a mixture of morphological traits of *Pantosteus virescens* and a species of *Catostomus*. Its mtDNA is most similar to that of *C. tahoensis* of the Lahontan Basin, suggesting hybrid origin involving that species. This interpretation requires immigration of *C. tahoensis* into the Snake River drainage, possibly at a time of spillover from the Lahontan Basin to the Lower Snake River proposed by Reheis et al. (2008), where it was apparently introgressed and absorbed into the populations of *P. virescens* that formerly occupied the Lower Snake River (described below). Morphologically,

*P. columbianus* is intermediate between those species in its lip shape and in many osteological traits, but shares scale counts and shapes of the jaws and suspensorium with *P. virescens* rather than *C. tahoensis* (Smith, 1966, Smith, in prep).

### 3.4. Fossil *Pantosteus*

Fossil *Pantosteus* occur in six different time divisions in the late Miocene to Pleistocene of western North America. The earliest are found in three formations in the late Miocene of southeast Oregon and southwest Idaho. Smith et al. (in prep) recovered diagnostic *Pantosteus* bones from the Juntura formation (10.4 Ma) of southeast Oregon, the Drewsey formation of southeast Oregon (8.5 Ma) and the Ellensburg formation of central Washington (10.3 Ma). Synapomorphies of these bones are shared with *P. virescens*, *P. discobolus*, *P. clarkii*, *P. santaanae* and *P. plebeius* (Figs. 2 and 4). Members of the *P. platyrhynchus* group have separate synapomorphies for this set of traits, more plesiomorphic with respect to *Catostomus*, making it difficult to assign this fossil to a specific node. It could be assigned to either end of the stem lineage for these taxa; however, we chose to be conservative and favor placement of this fossil at the oldest node within *Pantosteus* (Fig. 2). This represents first occurrence of the synapomorphies; however, this perspective requires one reversal in the hyomandibular trait of the *P. platyrhynchus* clade.

Two diagnostic samples of the second rib of the Weberian apparatus from the lower Glens Ferry formation of Castle Creek, Idaho, are similar to *P. virescens* in their extreme breadth. They are about 4.5 Ma and assigned to the node basal to that species. Fossil dentaries, a mesocorocoid and a fragment of a hyomandibulae sharing synapomorphies with *P. clarkii* (SBMNH L3313–10) are from the White Narrows formation of Clark County, Nevada. The age of those fossils is also 4.5 Ma and they are appropriately placed at the node ancestral to *P. clarkii* (Smith et al. in prep).

Many fossil dentaries of *P. columbianus* occur in the Glens Ferry formation at an estimated age of ca. 3 Ma (UMMP 55560, Miller and Smith, 1967) and ca. 2 Ma (UOMP 1413) of southwest Idaho. We do not include these fossils in the calibration because the hybrid origin of *P. columbianus* causes discordance between morphological and mtDNA topologies.

Several additional fossils are too young to be used to calibrate rates of mtDNA evolution. These include fossil bones of *P. jordani* from the Pleistocene (0.03 Ma) of northwest Nebraska, *P. plebeius* from middle Pleistocene (0.70–0.88 Ma) Hansen Bluff deposits of the Alamosa formation in south-central Colorado (Rogers et al., 1992), and a maxilla of *P. lahontan* from Plio-Pleistocene sediments of Owens Lake, California (Smith et al., 2009).

### 3.4. Dating of lineages

Fossils provided calibration points for two nodes in the tree, 11.6 Ma for the node at the base of all extant *Pantosteus* and 8.0 Ma for the node ancestral to *P. nebuliferus*–*P. plebeius* and *P. clarkii*–*P. discobolus*–*P. platyrhynchus*–*P. santaanae*–*P. virescens*. We explored a variety of subsets of our dataset to examine the effects of sample selection on divergence estimates, primarily to determine the effect of including outgroup samples from *Catostomus* on our estimates within *Pantosteus*. In all cases the results were similar (within 10%, results not shown). Sixteen individuals were selected to represent the phylogenetic diversity found within *Pantosteus* along with seven appropriate outgroup samples within *Catostomus*. The BEAST analysis was constrained to fit the ML topology (Fig. 2), with dates based on their mean and 95% highest posterior density (HPD) presented in Fig. 3; all parameter estimates had effective sample sizes greater than 2600. Each of the major nodes and associated geological events are discussed in detail below. These estimates of divergence

time are based on mtDNA; therefore, they may be influenced by introgression and must be interpreted with this caveat in mind.

The estimated mean time of generic divergence between *Pantosteus* and *Catostomus* was 23.8 Ma (95% HPD of 33.3–16.0 Ma). This corresponds to a period of high mountains throughout the area that was in the process of becoming the Basin and Range Province. The location of the earliest fossils in the proto-Columbia Basin, which corresponds with both high fossil and extant catostomid diversity in the West, suggests that the northern Great Basin is a plausible region in which initial divergence took place. The middle Miocene Nevada-Washington rift and the development of the Oregon-Idaho Graben in southern Oregon (Cummings et al., 2000) represent major tectonic structures and events possibly related to this time and place of early catostomin evolution.

The three oldest fossils of *Pantosteus*, 10.4, 10.3 and 8.5 Ma, are from separate drainages in central Washington and southeast Oregon (Smith et al., submitted), at the northern edge of the Basin and Range Province (Faulds and Varga, 1998). The estimated age of first occurrence of the clade bearing synapomorphies of these fossils (Fig. 3, mean 13.5 Ma, 95% HPD of 16.4–11.7 Ma) coincides with the extensional opening up of the Basin and Range topography. Because nearly all of the taxa in *Pantosteus* are in or adjacent to the northern and southern Great Basins (Fig. 1), it is probable that the distribution and differentiation of these fishes owes its history to habitats and barriers created in this geological province. *Pantosteus* are inhabitants of medium-gradient streams of the size that flow from the hundreds of north-south trending mountain ranges of the Basin and Range province. These mountains also form barriers that isolate populations and initiate differentiation, but this differentiation is periodically interrupted by moist-dry climatic cycles that allow dispersal through renewed aquatic connections.

The oldest branches of *Pantosteus*, based on molecular analyses, are the (1) the broad *P. plebeius*–*P. discobolus* clade, distributed south through the Great Basin to Mexico, including the upper Rio Grande Basin in the Rio Grande rift, and onto the Colorado Plateau, and (2) the *P. platyrhynchus* group, distributed in the north from the Columbia drainage and northern Great Basin to the upper Missouri drainage. Our analyses suggest that the first great pathway of dispersal and divergent evolution, following the initial split, was south through the northern to the southern Great Basin.

In the south, the *P. plebeius* and *P. discobolus* clades began diverging around 10.3 Ma (Fig. 3, 95% HPD of 12.8–8.2 Ma), a time of maximum geologic and climatic changes in the Northwest (Cummings et al., 2000). The *P. discobolus* clade differentiated into its five modern species in the northern Great Basin and Colorado Plateau about 5.2 Ma (Fig. 3, 95% HPD of 7.7–3.0 Ma). The Upper and Lower Colorado basins achieved their connection through Grand Canyon at 5.6 Ma (Howard et al., 2008), so it is likely that differentiation was occurring in pre-existing drainages in the Grand Wash area (Hunt, 1960; Howard et al., 2008; Wernicke, 2011) between 10.3 and 5.2 Ma. At this time, the Colorado Plateau was elevating but probably had less *Pantosteus* habitat than the Great Basin (Blakey and Ranney, 2008; Dallegge et al., 2003; Spencer et al., 2008). Four of the five species in the *P. discobolus* group dispersed through and differentiated in the Great Basin west of the emerging Colorado Plateau (Fig. 1).

The Bonneville and Lahontan basins were extending in the late Miocene, causing numerous north-south drainage connections. The Colorado River was in a stage preceding the modern Grand Canyon at this time although older canyons were present (Wernicke, 2011; Lucchitta, 1972, 2011). *Pantosteus* (and *Catostomus*) must have expanded their range through the Basin and Range area where the Colorado Plateau and the northern and southern

Great Basins meet (Fig. 1), to establish the southern clades in Mexico (see below) during this period.

The problem of the relationships of *P. santaanae* to either *P. clarkii* or to the remainder of this clade is not resolved by the phylogenetic analysis of mtDNA as indicated by conflict between ML and MP analyses. Morphological data are also ambiguous. An earlier date of divergence is plausible if the ancestors date back to the dispersal southward to Mexico. These estimates might be impacted, however, by the long branches induced by warm temperatures and small body size (Estabrook et al., 2007).

The map and trees (Figs. 1–4) suggest that the *P. platyrhynchus* group originated 13.5 Ma (Fig. 3, 95% HPD of 16.4–11.7 Ma) and diversified over the last 6.5–4.5 Ma (Fig. 3, 95% HPD of 9.9–2.0 Ma) into four species: *Pantosteus* sp. in the Columbia drainage, *P. jordani* in the upper Missouri drainage, *P. lahontan* in the Lahontan Basin, and *P. platyrhynchus* in the Bonneville Basin, Upper Snake River and the Green River. This diversification corresponds to a time of major volcanic and tectonic activity in the Columbia region (Tolan et al., 2009). These lineages probably began in three pre-Columbia River drainages of the northern Basin and Range Province, where the three oldest *Pantosteus* fossils were recovered. All except one differentiated in the Northern Great Basin—*P. jordani*, evolved initially in the Yellowstone region, now in the headwaters of the Missouri River, which at that time drained to Hudson Bay (Lemke et al., 1965). Later, in the Pleistocene, there were exchanges among the Bear, Snake and Green rivers in southwest Wyoming (Hanson, 1985, 1986; Smith et al., 2002; Spencer et al., 2008), which were possibly responsible for the transfer of individuals of *P. platyrhynchus* from west to east (see below).

The northern Bonneville/Upper Snake and southern Bonneville populations of *P. platyrhynchus* last shared a common ancestor 2.2 Ma (Fig. 3, 95% HPD of 3.5–0.9 Ma). This separation between the southern and northern Bonneville Basin is a common pattern shared

by two other fishes: *Snyderichthys copei* (Johnson et al., 2004) and *Rhinichthys osculus* (Billman et al., 2010). The Sevier Basin was isolated from the Salt Lake Basin by a barrier higher than the present, from at least 3 Ma until the early Pleistocene (Oviatt, 1994). The recent connection of the southern and northern halves of the Bonneville Basin via high stands of Lake Bonneville appears to have not facilitated movement between them, which may be due to these species being stream specialists that are not typically found in lakes.

The divergence of Mexican and Rio Grande Basin *Pantosteus*, between 8.3 and 1.1 Ma, suggest dispersal from Idaho southward through Nevada and Utah prior to this time (see Luchitta, 2011). The drainage changes responsible for divergence of *P. plebeius* and *P. nebuliferus* in central Mexico are not known.

### 3.5. Importance of introgressive hybridization in *Pantosteus*

Comparison of mtDNA distributions among taxa with morphological data, biogeography and paleontology reveals noncongruence that suggests either extensive morphological convergence or introgressive transfer of mtDNA among lineages in the ancient past (Fig. 4). Several considerations suggest resolution in favor of the introgressive transfer hypothesis. (1) Paleontological, morphological and biogeographical interpretations are often consistent with each other, but not with the mtDNA phylogeny. (2) The discrepancies in the mtDNA phylogeny involve taxa known to have experienced introgressive hybridization in recent times (Smith, 1966), allowing for the possibility in the past. (3) Some examples of introgression among fishes of western North America have been observed in the fossil record (Smith, 1992a) or have been verified by identification of geological evidence of stream capture consistent with the transfer leading to the introgression (Smith et al., 1983). (4) Lastly, the pattern of variation of morphological traits within individuals of an introgressed population show individually unique patterns of variation,

suggesting different degrees of penetration of different alleles, epigenetic interactions and selective regimes. That is, there are unusually variable and unique patterns of assorted polymorphisms among individuals (Hubbs et al., 1943; Hubbs, 1955; Smith, 1966, 1992a; Neff and Smith, 1979). The principal examples of introgressive hybridization in *Pantosteus* are the Zuni Sucker (*P. discobolus jarrovii*), the Bonneville Basin, Snake River and Green River *P. platyrhynchus* and *P. columbianus*, all of which demonstrate geographically patchy and non-concordant distributions of characters and alleles typical of introgression (Dowling and Moore, 1984; Dowling et al., 1989; Gerber et al., 2001). We now discuss each example of introgression below.

*Pantosteus plebeius* introgression with *P. discobolus* was documented with morphological, biogeographic and geologic data by Smith (1966), and the hypothesis tested and supported with biochemical data (Smith et al., 1984; Crabtree and Buth, 1987). The taxon, *P. discobolus jarrovii* Cope was based on the downstream cline of mixed and intermediate characters of the introgressed population, inferred to have resulted from a stream capture and fish transfer in the Pleistocene, with an estimated mean age of 1.1 Ma (Fig. 3, 95% HPD of 2.1–0.3 Ma). Because of multiple haplotypes, morphologies and conservation problems, these unusually complex populations are under separate investigation with more detailed sampling (Dowling, unpublished).

*Pantosteus discobolus* mtDNA has replaced that of *P. platyrhynchus* throughout its range as exhibited by the discordance of morphological and mtDNA characters. Morphological synapomorphies detailed by Smith (1966) and discussed above support the placement of *P. platyrhynchus* in the clade of populations, all formerly classified as *P. platyrhynchus*, distributed in the Columbia-Snake drainage, Missouri drainage, Green River and northern Great Basin. These populations are well known to hybridize with sympatric catostomins (Hubbs et al., 1943; Smith, 1966) so it is no surprise to find mtDNA of the *P.*

*discobolus* group in *P. platyrhynchus* in the Bonneville Basin, Upper Snake River and Green River. It was unexpected, however, to find all sampled individuals of Bonneville Basin and Upper Snake River *P. platyrhynchus* bearing *P. discobolus* mtDNA, dating from 3.2 Ma (Fig. 3, 95% HPD of 4.9–1.7 Ma). Introgression between these species in the Green River, between *P. platyrhynchus* and *P. virescens* in the Upper Snake River were described by Smith (1966). However, no morphological traits of *P. discobolus* were observed in *P. platyrhynchus* populations in the Bonneville Basin. The presence of *P. discobolus* mtDNA in Bonneville Basin and Upper Snake River *P. platyrhynchus* populations is especially puzzling because *P. discobolus* is allopatric to Bonneville Basin *P. platyrhynchus*; *P. virescens* is the sympatric congener (Fig. 2). *Pantosteus discobolus* and *P. virescens* are larger fish, preferring larger, downstream reaches; *P. platyrhynchus* is a headwater species, but where stream habitats overlap they are sympatric. Additional sampling is required across the range of *P. platyrhynchus* to determine if patterns of introgression are ubiquitous.

The hypothesis that the populations of *P. platyrhynchus* are actually genetic *P. discobolus* that have evolved scale, fin, pigment and osteological traits that converge on *P. platyrhynchus*, is rejected based on the diversity and independence of the morphological evidence. This assumption of independence of diverse morphological traits should be tested, however, with studies of nuclear genes responsible for morphology (Colosimo et al., 2005).

Geomorphic evidence of stream capture, possibly responsible for the transfer of fish between the Snake and Green rivers in southwest Wyoming, are noted above. Preliminary sampling indicates that the introgressed traits, which seemed more or less stabilized when observed in the Snake and Green rivers by Smith (1966, see graph, p.109), had become more prevalent in samples examined for the present study. *Pantosteus platyrhynchus* samples in the Upper Snake River exhibit *P. virescens* morphological traits and mtDNA (Fig. 2) to varying

degrees. *Pantosteus discobolus* mtDNA (Fig. 2) and morphological traits were scattered in *P. platyrhynchus* specimens in the Green River.

In particular, the *P. platyrhynchus* in Fremont Lake, in the upper Green River of Wyoming, has mtDNA that differs by a p-distance of 1.1% from *P. discobolus* collected in the upper Grand River at Rifle, Colorado (Fig. 2). The *P. discobolus* mtDNA introgressed into *P. platyrhynchus* is estimating the differentiation of *P. discobolus* in the Green River from populations in the upper Grand River (2.2 Ma, Fig. 3), not *P. platyrhynchus*.

*Pantosteus columbianus* appears to have a complex history of mixing with both *Catostomus tahoensis* and *P. virescens*. *Pantosteus columbianus* was described once in *Pantosteus* and again in *Catostomus* because of its mixture of diagnostic traits. Its lips are intermediate; most of its osteological traits are diagnostic of *Pantosteus*; its mtDNA is most similar to that of *Catostomus tahoensis*. The genetic divergence between the mtDNA of *P. columbianus* and *C. tahoensis* implies a mean age of introgression of 2.2 Ma (Fig. 3, 95% HPD of 3.7–0.9 Ma). Fossils of *P. columbianus* in the western Snake River plain between 3 and 2 Ma help constrain the time of introgression that formed *P. columbianus*. The absence of *P. virescens* below the falls of the Snake River, notwithstanding fossil evidence of its occurrence there at 4.5 Ma, implies that the immigration of *C. tahoensis* during spillover from the Lahontan Basin to the Snake River drainage in the Pleistocene (Reheis et al., 2008), resulted in complete genetic absorption of the two parental populations into a species of introgressed origin, *P. columbianus*. Further tests of this hypothesis with nuclear genes and karyotypes should be instructive. Populations of *P. columbianus* are isolated in the headwaters of the Palouse River, Washington, the Wood River, Idaho, and the Deschutes and Malheur rivers, Oregon. Their morphological differences from the central populations suggest an early widespread form of the species that was more *Pantosteus*-like, especially in

bones and lip structure. These indicate significant morphological change during their existence (Smith, 1966).

#### 4. Conclusions

The Basin and Range Province is a unique setting for the study of effects of insular isolation on diversification of plants and of aquatic and terrestrial animals (Blackwelder et al., 1948; Harper and Reveal 1978; Reheis et al., 2008). The fragmentation of a late Mesozoic mountain belt along the southwest margin of North America by crustal thinning and extension created hundreds of narrow valleys separated by hundreds of elongate mountainous remnants. The valleys were frequently isolated except when filled and externally drained (Reheis et al., 2008, Fig. 1) and the mountains were isolated except when cool climates forced transitional and boreal life zones down and out across valley floors. During late Neogene glacial/interglacial stages, fluctuations in barriers and connections led to about 20 cycles of isolation, differentiation, and repeated secondary genetic contact among populations of organisms. Only the most recent of these is well studied.

The distribution of mtDNA among populations of *Pantosteus* reflects this cyclic history. Many instances of secondary contact resulted in hybridization, and some populations experienced introgression, or even amalgamation (reviewed in Dowling et al., 1997), resulting in conflicting patterns in morphology and DNA interpretations of relationships. Most species of western catostomids (Hubbs et al., 1943) and cyprinids (Hubbs, 1955) show evidence of hybridization and all 11 species of *Pantosteus* have hybridized with sympatric relatives (Hubbs et al., 1943; Smith, 1966). Morphological and molecular evidence presented here suggests that introgression was ancient (e.g., pre-human influence) in *P. tahoensis* x *P. virescens* (= *P. columbianus*), *P. platyrhynchus* x *P. discobolus*, *P. platyrhynchus* x *P. virescens*, *P. discobolus* x *P. clarkii*, *P. discobolus* x *P. plebeius* and possibly *P. plebeius* x *P.*

*nebuliferus*. Many other hybrid individuals, e.g., *P. platyrhynchus* x *P. clarkii* and many hybrids involving *Catostomus* (Hubbs et al., 1943) show distributions and characters that suggest recent hybridization in the presence of environmental disturbance and inter-basin transfer by humans. Environmental disturbance in the form of compromised spawning sites were presumably responsible for many ancient introgressions. Lake spillovers and stream captures are circumstantially associated with many of these cases (Smith and Dowling, unpublished).

Ancient introgression causes the production of morphotype/haplotype combinations that yield misleading phylogenetic reconstructions (Smith, 1992a) and misidentification of specimens by scientists and managers. Changes in biodiversity, negative or positive, are of considerable interest (Dowling et al., 1992a, b; Dowling and Secor, 1997). Loss of species is a possibility, as in the case of dwindling populations of *P. virescens* in the Upper Snake, Bear and Weber rivers. Replacement of species, such as *C. tahoensis* and *P. virescens* by *P. columbianus* in the Columbia drainage is an interesting example. Post-glacial expansion of *P. columbianus* north to 55 degrees in Canada is evidence of an adaptable genome. Whether the added genetic variability in introgressed *P. platyrhynchus* in the Green River is adaptive, neutral, or harmful in its changing environment is not known, but compared to *P. lahontan* and *P. jordani*, some of these populations appear to be spreading the introgressed alleles while declining in numbers over the 50 years since 1960 (G. Smith, unpublished data). Comparative studies of population genetics, genomics and impact of diverse sources of genetic variation in populations of *P. discobolus jarrovii*, *P. platyrhynchus*, *P. lahontan* and *P. jordani* would be instructive.

## **Acknowledgments**

We thank the many people who assisted with all aspects of fieldwork, sample curatorial work and for providing samples and other assistance including K. Bertrand, C. Dillman, S. Gregory, G. Gunnell, D. Markle, D. Nelson, L. Schultz and E. Taylor. J. Megahan helped with figures 1 and 4 and N. Carpenter provided dates for the fossils. All collections were obtained under permit from various state fisheries agencies and research was conducted under approved institutional animal ethics guidelines. This work was supported in part by funds from the US Bureau of Reclamation to TED and CLS, by the National Evolutionary Synthesis Center (NESCent) and NSF EF-0905606 to PJU and NSF DEB 0240184, EF 0431326, DEB-0817027, DBI-0956370 and DEB-1021840 to RLM.

## References

- Abell, R., Thieme, M.L., Revenga, C., Bryeret, M., et al., 2008. Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. *Bioscience* 58, 403–414.
- Arnold, M.L., 1997. Natural hybridization and evolution. Oxford University Press, Oxford.
- Baird, S.F., Girard, C.F., 1854. Descriptions of new species of fishes collected in Texas, New Mexico and Sonora, by Mr. John H. Clark, on the U. S. and Mexican Boundary Survey, and in Texas by Capt. Stewart Van Vliet, U. S. A. *Proceedings of the Academy of Natural Sciences of Philadelphia* 7, 24–29.
- Benson, L. 2004. Western lakes. In: Gillespie, A.R., S.C. Porter, B.F. Atwater, (Eds.), *The Quaternary Period in the United States, Developments in Quaternary Science*, 1. Elsevier, p. 185–199.
- Beranek, L.P., Link, P.K., Fanning, C.M., 2006. Miocene to Holocene landscape evolution of the western Snake River Plain Region, Idaho: using the SHRIMP detrital zircon provenance record to track eastward migration of the Yellowstone hotspot. *Geological Society of America Bulletin* 118, 1027–1050.
- Blackwelder, E., Hubbs, C.L., Antevs, E., eds., 1948. The Great Basin, with emphasis on glacial and post-glacial times. *Bulletin of the University of Utah* 38. Biological Series X(7), 1–191.
- Blakey, R., Ranney, W., 2008. Ancient Landscapes of the Colorado Plateau: From Desert to Mountains. Grand Canyon Association, 176 pp. Grand Canyon, AZ.
- Bouchard, D.P., Kaufman, D.S., Hochberg, A., Quade, J., 1998, Quaternary history of the Thatcher Basin, Idaho, reconstructed from the  $^{87}\text{Sr}/^{86}\text{Sr}$  and amino acid composition of lacustrine fossils--implications for the diversion of the Bear River into the Bonneville basin: *Palaeogeography, Palaeoclimatology, Palaeoecology* 141, 95–114.

- Bufalino, A.P., Mayden, R.L., 2010. Molecular phylogenetics of North American phoxinins (Actinopterygii: Cypriniformes: Leuciscidae) based on RAG1 and S7 nuclear DNA sequence data. *Molecular Phylogenetics and Evolution* 55, 274–283.
- Burridge, C.P., Craw, D., Fletcher, D., Waters, J.M., 2008. Geological dates and molecular rates: fish DNA sheds light on time-dependency. *Molecular Biology and Evolution* 25, 624–633.
- Cavender, T.M., 1986. Review of the fossil history of North American freshwater fishes. In: Hocutt, C.H., Wiley, E.O. (Eds.), *The Zoogeography of North American Fishes*. John Wiley and Sons, New York.
- Chamberlain, C.P., Mix, H.T., Mulch, A., Hren, M.T., Kent-Corson, M.L., Davis, S.J., Horton, T.W., Graham, S.A., 2012. The Cenozoic climatic and topographic evolution of the western North American Cordillera. *American Journal of Science* 312, 213–262.
- Colosimo, P. F., Hosemann, K.E., Balabhadra, S., Villareal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R., Schluter, D., Kingsley, D.M., 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307, 1928–1933.
- Cope, E.D., 1872. Recent reptiles and fishes. Report on the reptiles and fishes obtained by the naturalists on the expedition. U.S. Geological Survey of Wyoming & Contiguous Territories (1871). U.S. Geological Survey of Wyoming & Contiguous Territories Part 4 (art. 8), 432–442.
- Cope, E.D., 1874. On the Plagopterinae and the ichthyology of Utah. *Transactions of the American Philosophical Society* 14, 129–139.
- Cope, E.D., Yarrow, H.C., 1875. Report upon the collections of fishes made in portions of Nevada, Utah, California, Colorado, New Mexico, and Arizona, during the years

- 1871, 1872, 1873, and 1874. Engineer Department, United States Army, Report upon the Geographical and Geological Explorations and Surveys west of the one hundredth Meridian v. 5 (Zoology) Chapter 6, 635–703.
- Crabtree, C.B., Buth, D.G., 1987. Biochemical systematics of the catostomid genus *Catostomus*: assessment of *C. clarki*, *C. plebeius* and *C. discobolus* including the Zuni sucker, *C. d. yarrowi*. *Copeia* 1987, 843–854.
- Cummings, M.L., Evans, J.G., Ferns, M.L., Lees, K.R., 2000. Stratigraphic and structural evolution of the middle Miocene synvolcanic Oregon-Idaho graben. *The Geological Society of America Bulletin* 112, 668–682.
- Currey, D.R., 1990. Quaternary palaeolakes in the evolution of semidesert basins, with special emphasis on Lake Bonneville and the Great Basin, U.S.A. *Palaeogeography, Palaeoclimatology, Palaeoecology* 76, 189–214.
- Dallegge, T.A., Ort, M.H., McIntosh, W.C., Perkins, M.E., 2001. Age and depositional basin morphology of the Bidahochi Formation and implications for the ancestral upper Colorado River. In: Young, R.A., Spamer, E.E., (Eds.), *Colorado River: Origin and Evolution; Proceedings of a Symposium Held at Grand Canyon National Park, June, 2000*. Grand Canyon Association Monograph 12, 47–51.
- Davis, O.K., 1999. Pollen analysis of Tulare Lake, California: Great Basin-like vegetation in central California during the full-glacial and early Holocene. *Review of Palaeobotany and Palynology* 107, 249–257.
- DeMarais, B.D., Dowling, T.E., Douglas, M.E., Minckley, W.L., Marsh, P.C., 1992. Origin of *Gila seminuda* (Teleostei: Cyprinidae) through introgressive hybridization: implications for evolution and conservation. *Proceedings of the National Academy of Sciences USA* 89, 2747–2751.

- Doosey, M.H., Bart Jr., H.L., Saitoh, K., Miya, M., 2009. Phylogenetic relationships of catostomid fishes (Actinopterygii: Cypriniformes) based on mitochondrial ND4/ND5 gene sequences, *Molecular Phylogenetics and Evolution* 54, 1028–1034.
- Dowling, T.E., DeMarais, B.D., 1993. Evolutionary significance of introgressive hybridization in cyprinid fishes. *Nature* 362, 444–446.
- Dowling, T. E., DeMarais, B.D., Minckley, W.L., Douglas, M.E., Marsh, P.C., 1992a. Use of genetic characters in conservation biology. *Conservation Biology* 6, 7–8.
- Dowling, T. E., Minckley, W.L., Douglas, M.E., Marsh, P.C., DeMarais, B.D., 1992b. Use of molecular characters in conservation biology: implications for management of the red wolf. *Conservation Biology* 6, 600–603.
- Dowling, T. E., Moore, W. S., 1984. The level of reproductive isolation between two cyprinid fishes, *Notropis cornutus* and *N. chrysocephalus*. *Copeia* 1984, 617–628.
- Dowling, T.E., Secor, C.L., 1997. The role of hybridization in the evolutionary diversification of animals. *Annual Review Ecology Systematics* 28, 593–619.
- Dowling, T.E., Smith G.R., Brown, W.M., 1989. Reproductive isolation and introgression between *Notropis cornutus* and *Notropis chrysocephalus* (family Cyprinidae): comparison of morphology, allozymes, and mitochondrial DNA. *Evolution* 43, 620-634.
- Dowling, T.E., Tibbets, C.A., Minckley, W.L., Smith, G.R., 2002. Evolutionary relationships of the plagopterins (Teleostei: Cyprinidae) from cytochrome *b* sequences. *Copeia* 2002, 665–678.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7, 214.

- Echelle, A.A., Carson, E.W., Echelle, A.F., Van Den Bussche, R.A., Dowling, T.E., Meyer, A., Wood, R.M., 2005. Historical biogeography of the new-world pupfish genus *Cyprinodon* (Teleostei: Cyprinodontidae). *Copeia* 2005, 320–339.
- Estabrook, G.F., Smith, G.R., Dowling, T.E., 2007. Body mass and temperature influence rates of mitochondrial DNA evolution in North American cyprinid fish. *Evolution* 61, 1176–1187.
- Evermann, B.W., 1894. Description of a new sucker, *Pantosteus jordani*, from the upper Missouri basin. *Bulletin of the U.S. Fish Commission* 12, 51–56.
- Eigenmann, C.H., Eigenmann, R.S., 1893. Preliminary descriptions of new fishes from the northwest. *American Naturalist* 27, 151–154.
- Faulds, J.E., Varga, R.J., 1998. The role of accommodation zones and transfer zones in the regional segmentation of extended terranes. In: Faulds, J.E., Stewart, J.H., (Eds.), *Accommodation Zones and Transfer Zones: the Regional Segmentation of the Basin and Range Province*, Geological Society of America, Special Paper 323, pp. 1–45.
- Ferris, S.D., Buth, D.G., Whitt, G.S., 1982. Substantial genetic differentiation among populations of *Catostomus plebeius*. *Copeia* 2002, 444–449.
- Garman, S., 1881. New and little-known reptiles and fishes in the museum collections. *Bulletin of the Museum of Comparative Zoology, Harvard University* 8, 85–93.
- Gerber, A.S., Tibbets, C.A., Dowling, T.E., 2001. The role of introgressive hybridization in the evolution of the *Gila robusta* complex (Teleostei: Cyprinidae). *Evolution* 55, 2028–2039.
- Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41, 95–98.
- Hansen, W.R., 1985. Drainage development of the Green River Basin in southwest Wyoming and its bearing on fish biogeography, neotectonics, and paleoclimates. *Mountain*

- Geologist 22, 192–204.
- Harper, K.T., Reveal, J.L., (Eds.). 1978. The Great Basin: a symposium. Great Basin Naturalist Memoir 2.
- Hershler, R., Sada, D.W., 2002. Biogeography of Great Basin aquatic snails of the genus *Pyrgulopsis*. In: Hershler, R., Madsen, D.B., Currey, D.R. (Eds.), Great Basin Aquatic Systems History. Smithsonian Contributions to the Earth Sciences, 33. Smithsonian Institution, Washington DC, pp. 255–276.
- Ho, S.Y.W., Larson, G., Edwards, C.J., Huepink, T.H., Lakin, K.E., Holland, P.W.H., Shapiro, B., 2008. Correlating Bayesian date estimates with climatic events and domestication using a bovine case study. *Biology Letters* 4, 370–374.
- Howard, K.A., Lundstrom, S.C., Malmon, D.V., Hook, S.J., 2008. Age, distribution, and formation of late Cenozoic paleovalleys of the lower Colorado River and their relation to river aggradation and degradation. In: Reheis, M.C., Hershler, R., Miller, D.M., (Eds.), Late Cenozoic Drainage History of the Southwestern Great Basin and Lower Colorado River Region: Geologic and Biotic Perspectives. Geological Society of America Special Paper 439, Geological Society of America, Boulder, CO, pp. 391–410.
- Hubbs, C.L., 1955. Hybridization between fish species in nature. *Syst. Zool.* 4, 1–20.
- Hubbs, C.L., Schultz, L.P., 1938. A new catostomid fish from the Columbia River. University of Washington Publications in Biology 2, 1–14.
- Hubbs, C.L., Hubbs, L.C., Johnson, R.E., 1943. Hybridization in nature between species of catostomid fishes. Contribution from the Laboratory of Vertebrate Biology, University of Michigan 22, 1–76.
- Hubbs, C.L., Miller, R.R., 1948. Correlation between fish distribution and hydrographic history in the desert basins of southwestern United States. In: The Great Basin, with

- emphasis on glacial and post-glacial times. Bulletin 38 of the University of Utah, Biological Series. 10(7), pp. 17–144.
- Hubbs, C.L., Lagler, K.F., 2004. Fishes of the Great Lakes Region, revised edition. University of Michigan Press, Ann Arbor.
- Hunt, C.B., 1960., Cenozoic Geology of the Colorado Plateau. U.S. Geological Survey Professional Paper 279, 1–99.
- Jordan, D.S., 1878. Contributions to North American ichthyology III.B. A synopsis of the family Catostomidae. Bulletin of the U.S. National Museum 12, 97–237.
- Jordan, D.S., Gilbert, C.H., 1882. Synopsis of the Fishes of North America. Bulletin of the U.S. National Museum 16, 1–1018.
- Jordan, D.S., Evermann, W.B., 1893. The Fishes of North and Middle America. Part I. Bulletin of the U.S. National Museum 47, 1–1240.
- Kimmel, P.G., 1975. Fishes of the Miocene-Pliocene Deer Butte Formation, Southeast Oregon. University of Michigan Museum of Paleontology Papers on Paleontology 14, 69–87.
- Kinziger, A.P., Wood, R.M., Neely, D.A., 2005. Molecular systematics of the genus *Cottus* (Scorpaeniformes: Cottidae). Copeia 2005, 303–311.
- Kohn, M.J. Fremd, T.J., 2008. Miocene tectonics and climate forcing of biodiversity, western United States. Geology 36, 783–786.
- Lemke, R.W., Laird, W.M., Tipton, M.J., Lindvall, R.M., 1965. Quaternary geology of the northern Great Plains. In: Wright, H.E., Frey, D.G., (Eds.), The Quaternary of the United States. Princeton N.J., Princeton University Press, pp. 15–27.
- Lucchitta, I., 1972. Early history of the Colorado River in the Basin and Range Province. Geological Society of America Bulletin 83, 1933–1948.

- Lucchitta, I., 2011. A Miocene river in northern Arizona and its implications for the Colorado River and Grand Canyon. *Geological Society of America Today* 21, 4–10.
- Marshall, C.R., 1990. Confidence intervals on stratigraphic ranges. *Paleobiology* 16, 1–10.
- Miller, R.R., Smith, G.R., 1967. New fossil fishes from Plio-Pleistocene Lake Idaho. *University of Michigan Museum of Zoology Occasional Papers* 654, 1–24.
- Miller, R.R., Minckley, W.L., Norris, S.M., 2005. *Freshwater Fishes of México*. Chicago, IL: The University of Chicago Press.
- Neff, N.A., Smith, G.R., 1979. Multivariate analysis of hybrid fishes. *Systematic Zoology* 28, 176–196.
- Negrini, R.M., 2002. Pluvial lake sizes in the northwestern Great Basin throughout the Quaternary Period. In: Hershler, R., Madsen, D.B., Currey, D.R. (Eds.), *Great Basin Aquatic Systems History*. *Smithsonian Contributions to the Earth Sciences*, 33. Smithsonian Institution, Washington DC, pp. 11–52.
- Oviatt, C.G., 1994. Review of the Quaternary geology of the Sevier and Black Rock Deserts. *Utah Geological Association Publication* 23, 97–103.
- Oviatt, C.G., 2002. Bonneville Basin lacustrine history: the contributions of G.K. Gilbert and Ernst Antevs. In: Hershler, R., Madsen, D.B., Currey, D.R., (Eds.), *Great Basin Aquatic Systems History: Smithsonian Contributions to the Earth Sciences* 33, 121–128.
- Parenti, L.R., 1981. A phylogenetic and biogeographic analysis of cyprinodontiform fishes (Teleostei, Atherinomorpha). *Bulletin of the American Museum of Natural History* 168, 335–557.
- Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.

- Reheis, M.C., Sarna-Wodjicki, A.M., Reynolds, R.L., Repenning, C.A., Mifflin, M.D., 2002. Pliocene to middle Pleistocene lakes in the western Great Basin: ages and connections. In: Hershler, R., Madsen, D.B., Curry, D.R., (Eds.), Great Basin Aquatic Systems History: Smithsonian Contributions to Earth Sciences 33, 53–108.
- Reheis, M.C., Hershler, R., Miller, D.M., eds., 2008. Late Cenozoic Drainage History of the Southwestern Great Basin and Lower Colorado River Region: Geologic and Biotic Perspectives. Geological Society of America Special Paper 439, pp. 432.
- Retallack, G.J., 2004. Late Miocene climate and life on land in Oregon within a context of global climate change. *Palaeogeography, Palaeoclimate, Palaeoecology* 214, 97–123.
- Rogers, K.L., Larson, E.E., Smith, G., Katzman, D., Smith, G.R., Cerling, T., Yang W., Baker, R.G., Lohmann, K.C., Repenning, C.A. Patterson, P., Mackie, G., 1992. Pliocene and Pleistocene geologic and climatic evolution in the San Luis Valley of south-central Colorado. *Palaeogeography, Palaeoclimatology, Palaeoecology* 94, 55–86.
- Rutter, C.M., 1903. Notes on fishes from streams and lakes of northeastern California not tributary to the Sacramento basin. *Bulletin of the U.S. Fish Commission* 22, 143–148.
- Sack, D., 2002. Fluvial linkages in Lake Bonneville subbasin integration. *Smithsonian Institution Contributions to the Earth Sciences* 33, 129–144.
- Smith, G.R., 1966. Distribution and evolution of the North American catostomid fishes of the subgenus *Pantosteus*, genus *Catostomus*. *Miscellaneous Publications, Museum of Zoology, University of Michigan* 129, 1–132.
- Smith, G.R., 1992a. Introgression in fishes: significance for paleontology, cladistics, and evolutionary rates. *Systematic Biology* 41, 41–57.
- Smith, G.R., 1992b. Phylogeny and biogeography of the Catostomidae, freshwater fishes of North America and Asia, In: Mayden, R.L., (Ed.), *Systematics, Historical Ecology,*

- and North American Freshwater Fishes. Stanford University Press, Stanford, pp. 778–826.
- Smith, G.R., Koehn, R.K., 1971. Phenetic and cladistic studies of biochemical and morphological characteristics of *Catostomus*. *Syst. Zool.* 20, 282–297.
- Smith, G.R., Swirydczuk, K. Kimmel, P.G., and Wilkinson, B.H., 1982. Fish biostratigraphy of late Miocene to Pleistocene sediments of the western Snake River Plain, Idaho. In: Bonnicksen, B., Breckenridge, R.M., (Eds.), *Cenozoic Geology of Idaho*. Idaho Bureau of Mines and Geology, Bulletin 26, pp. 519–542.
- Smith, G.R., Hall, J.G., Koehn, R.K., Innes, D.J., 1983. Taxonomic relationships of the Zuni mountain sucker, *Catostomus discobolus yarrowi*. *Copeia* 1983, 37–48.
- Smith, G.R., Morgan, N, and Gustafson, E., 2000. Fishes of the Mio-Pliocene Ringold Formation, Washington: Pliocene Capture of the Snake River by the Columbia River. *University of Michigan Museum of Paleontology Papers on Paleontology* 32, 1–47.
- Smith, G.R., Dowling, T.E., Gobalet, K., Lugaski, T., Shiozawa, D., Evans, P., 2002. Biogeography and rates of evolution of Great Basin fishes. In: Hershler, R., Madsen, D., Curry, D., (Eds.), *The Great Basin: Cenozoic Geology and Biogeography, Great Basin Aquatic Systems History*. *Smithsonian Contributions to Earth Sciences*, v 33. Smithsonian Institution, Washington, DC, pp.175–234.
- Smith, G.R., Reynolds, R.E., Serrano, R.J., 2009. Recent records of fossil fish from eastern Owens Lake, Inyo County, California. In: Reynolds, R.E., (Ed.), *Desert Symposium, 2009*. Redlands, CA, pp. 177–183.
- Smith, G.R., Badgley, C.E., Eiting, T.E., Larson, P., 2010. Species diversity gradients in relation to geological history in North American freshwater fishes. *Evolutionary Ecology Research* 12, 693–726.

- Smith, G.R., Stewart, J.D., Carpenter, N.E., submitted. Miocene-Pleistocene fossils of fishes of the genus *Pantosteus*, Western United States. University of Michigan Museum of Zoology Occasional Papers.
- Snyder, J.O., 1917. The fishes of the Lahontan system of Nevada, and northeastern California. Bulletin of the U.S. Bureau of Fisheries 35 (1915), 31–86.
- Snyder, J.O., 1924. Notes on certain catostomids of the Bonneville system, including the type of *Pantosteus virescens* Cope. Proceedings of the U.S. National Museum 64, 1–6.
- Spencer, J.E., Smith, G.R., Dowling, T.E., 2008. Middle to late Cenozoic geology, hydrography, and fish evolution in the American Southwest. In: Reheis, M.C., Hershler, R., Miller, D.M., (eds.), Late Cenozoic Drainage History of the Southwestern Great Basin and Lower Colorado River Region: Geologic and Biotic Perspectives: Geological Society of America Special Paper 439, pp. 279–299.
- Stamatakis, A., 2006. RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
- Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web-servers. Systematic Biology 75, 758–771.
- Swofford, D.L., 2003. PAUP\*. Phylogenetic Analysis Using Parsimony (\* and other methods), Version 4.0b10. Sinauer, Sunderland.
- Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596–1599.
- Tanner, V.M., 1932. A description of *Notolepidomyzon utahensis*, a new catostomid from Utah. Copeia 1932, 135–136.
- Tanner, V.M., 1942. A review of the genus *Notolepidomyzon* with a description of a new species (Pisces-Catostomidae). Great Basin Naturalist 3, 27–32.

- Taylor, D.W., 1985. Evolution of freshwater drainages and mollusks in western North America. In: Smiley, C.J., (Ed.), Late Cenozoic History of the Pacific Northwest: San Francisco, American Association for the Advancement of Science (Pacific Division), pp. 265–321.
- Tolan, T.L., Martin, B.S., Reidel, S.P., Kauffman, J.D., Garwood, D.L., Anderson, J.L., 2009. Stratigraphy and tectonics of the central and eastern portions of the Columbia River Flood-Basalt Province: an overview of our current state of knowledge. In: O'Connor, J.E., Dorsey, R.J., Madin, I.P., (Eds.), Volcanoes to Vineyards: Geologic Field Trips Through the Dynamic Landscape of the Pacific Northwest, Geological Society of America Field Guide 15, pp. 645–672.
- Webb, S.A., Graves, J.A., Macias-Garcia, C., Magurran, A.E., Foighil, D.Ó., Ritchie, M.G., 2004. Molecular phylogeny of the livebearing Goodeidae (Cyprinodontiformes), *Molecular Phylogenetics and Evolution* 30, 527–544.
- Wernicke, B., 2011. The California River and its role in carving Grand Canyon. *Geological Society of America Bulletin* 123, 1288–1316.
- Wheeler, H.E., Cook, E.F., 1954. Structural and stratigraphic significance of the Snake River capture, Idaho and Oregon. *The Journal of Geology* 62, 525–536.

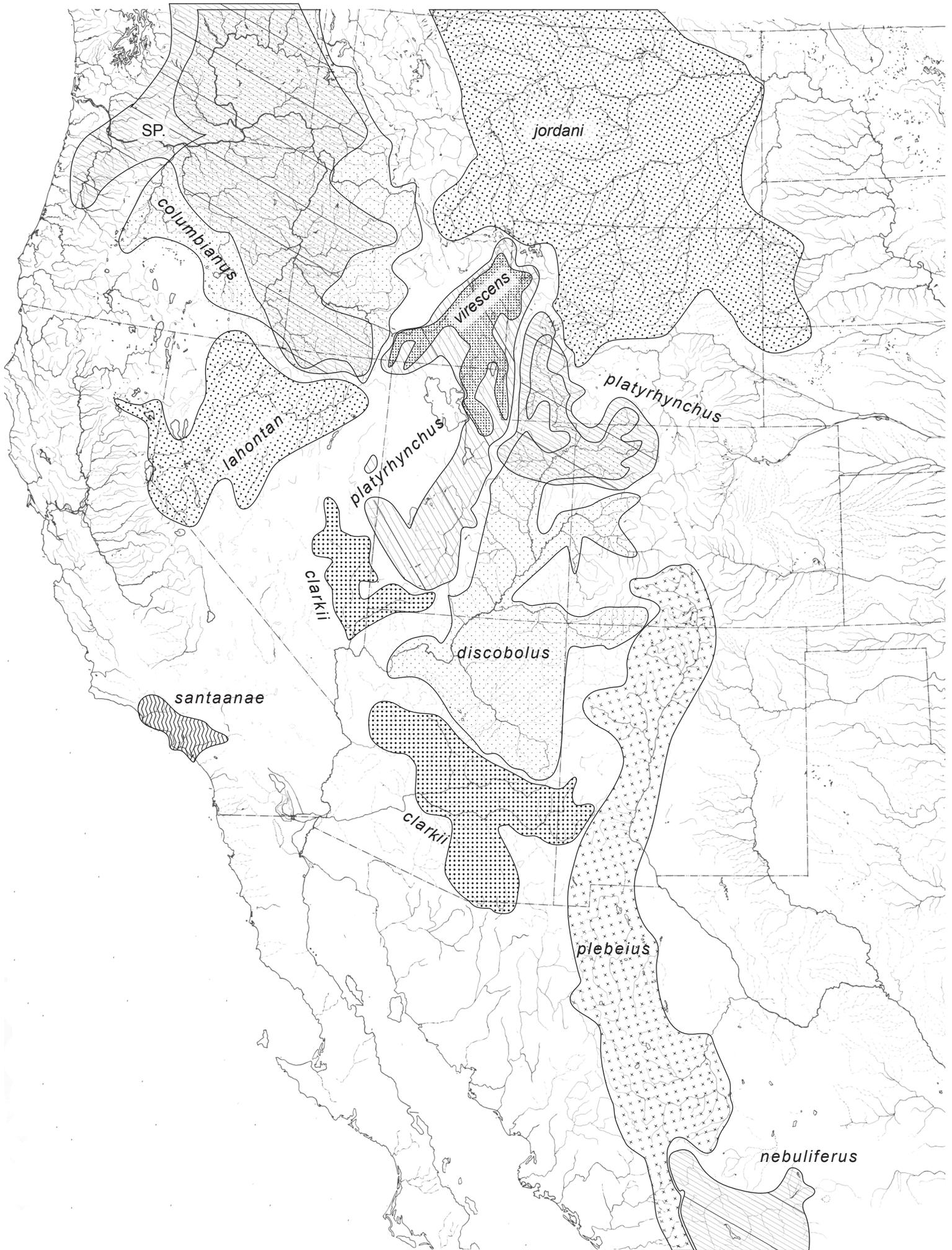
## Figure Captions

- Fig. 1.** Distribution of the 11 species of *Pantosteus*. Border species are found north and south of the margins of this map.
- Fig. 2.** Maximum likelihood (ML) tree for *Pantosteus* based on analysis of mitochondrial DNA. A collapsed maximum parsimony (MP) topology is shown in the upper left corner, the thick branches indicate differences in topology to ML. All bootstrap values are based on 1000 pseudoreplicates, a # symbol represents bootstrap values over 95 for both ML/MP. Single bootstrap values are shown for nodes which are different to MP. The tree is rooted with *Cycleptus elongatus*. Locality details are provided in Table 1.
- Fig. 3.** Bayesian molecular clock estimates for *Pantosteus* based on analysis of mitochondrial DNA. Horizontal bars represent the 95% highest posterior density ranges. Locality details are provided in Table 1, OTU order and labels match those in Fig. 2.
- Fig. 4.** Morphology-based phylogenetic tree based on synapomorphies of *Pantosteus* species, with arrows indicating direction of introgression. Timing of branch points and introgression events are not to scale.

Table 1. Locality data for all *Pantosteus* samples and outgroups.

Species	Locality	sub-basin	major basin
<i>Pantosteus clarkii</i>	White R	Pluvial White	L Colorado
<i>P. clarkii</i>	Virgin R	Virgin	L Colorado
<i>P. clarkii</i>	Francis Ck	Bill Williams	L Colorado
<i>P. clarkii</i>	Aravaipa Ck	Gila	L Colorado
<i>P. columbianus</i>	Salmon Falls Ck	L Snake	Columbia
<i>P. discobolus</i>	Rifle Ck	U Colorado (CO)	U Colorado
<i>P. discobolus</i>	Havasau Ck at mouth	U Colorado (AZ)	U Colorado
<i>P. discobolus</i>	Tsaile Ck	San Juan	U Colorado
<i>P. discobolus</i>	Wheatsfield Ck	San Juan	U Colorado
<i>P. discobolus</i>	East Clear Ck	Little Colorado	U Colorado
<i>P. discobolus</i>	Upper Little Colorado R	Little Colorado	U Colorado
<i>P. discobolus jarrovii</i>	Rio Nutria	Little Colorado	U Colorado
<i>P. jordani</i>	Whitewood Ck	Missouri (SD)	Mississippi
<i>P. jordani</i>	Lil Popo Agie R	Missouri (WY)	Mississippi
<i>P. lahontan</i>	McDermitt Ck	Quinn	Lahontan
<i>P. lahontan</i>	Lower Truckee R	Truckee	Lahontan
<i>P. lahontan</i>	E Walker R	Walker	Lahontan
<i>P. nebuliferus</i>	Arroyo Penon	Nazas	Nazas
<i>P. platyrhynchus</i>	Blackrock Ck	U Snake	U Snake
<i>P. platyrhynchus</i>	Bear R	Bear	Bonneville
<i>P. platyrhynchus</i>	Weber R	Weber	Bonneville
<i>P. platyrhynchus</i>	Soldier Ck	Spanish Fork	Bonneville
<i>P. platyrhynchus</i>	San Pitch R	Sevier	Bonneville
<i>P. platyrhynchus</i>	Salina Ck	Sevier	Bonneville
<i>P. platyrhynchus</i>	Mammoth Ck	Sevier	Bonneville
<i>P. platyrhynchus</i>	Fremont L	U Green	U Colorado
<i>P. plebeius</i>	S Fork Palomas Ck	Rio Grande	Rio Grande
<i>P. plebeius</i>	R Santa Clara	Santa Clara	Guzmán
<i>P. plebeius</i>	R Escalariado	Escalariado	Guzmán
<i>P. plebeius</i>	Arroyo Ureyna	Conchos	Conchos
<i>P. plebeius</i>	Arroyo Riito	Fuerte	Fuerte
<i>P. plebeius</i>	R Miravalles	Miravalles	Miravalles
<i>P. santaanae</i>	Big Tujunga Ck	Tujunga	Los Angeles Basin
<i>P. santaanae</i>	San Gabriel R	San Gabriel	Los Angeles Basin
<i>P. santaanae</i>	Santa Ana R	Santa Ana	Los Angeles Basin
<i>P. species</i>	Similkameen R	Columbia	Columbia
<i>P. species</i>	Willamette R	Willamette	Columbia
<i>P. species</i>	Salmon Falls Ck	L Snake	Columbia
<i>P. virescens</i>	Twin Ck	U Snake	U Snake
<i>P. virescens</i>	Tin Cup Ck	U Snake	U Snake
<i>P. virescens</i>	Goose Ck	U Snake	U Snake
<i>P. virescens</i>	Weber R	Weber	Bonneville
<i>P. virescens</i>	Echo Ck	Weber	Bonneville
<i>Catostomus ardens</i>	Strawberry Res (introduced)	Duchesne	U Colorado
<i>Catostomus catostomus</i>	Boundary waters	Lake Superior	Great Lakes
<i>Catostomus commersonii</i>	GenBank AB127394.1		
<i>Catostomus insignis</i>	Eagle Ck	Gila	L Colorado
<i>Catostomus macrocheilus</i>	Grimes Ck	L Snake	Columbia
<i>Catostomus macrocheilus</i>	Palouse R	L Snake	Columbia
<i>Catostomus tahoensis</i>	Walker R	Walker	Lahontan
<i>Cycleptus elongatus</i>	GenBank AB126082.1		
<i>Hypentelium nigricans</i>	GenBank AB242169.1		
<i>Minytrema melanops</i>	GenBank DQ536432.1		
<i>Moxostoma erythrurum</i>	Middle Fork Vermilion R	Mississippi (IL)	Mississippi
<i>Moxostoma poecilurum</i>	GenBank AB242167.1		

Figure 1



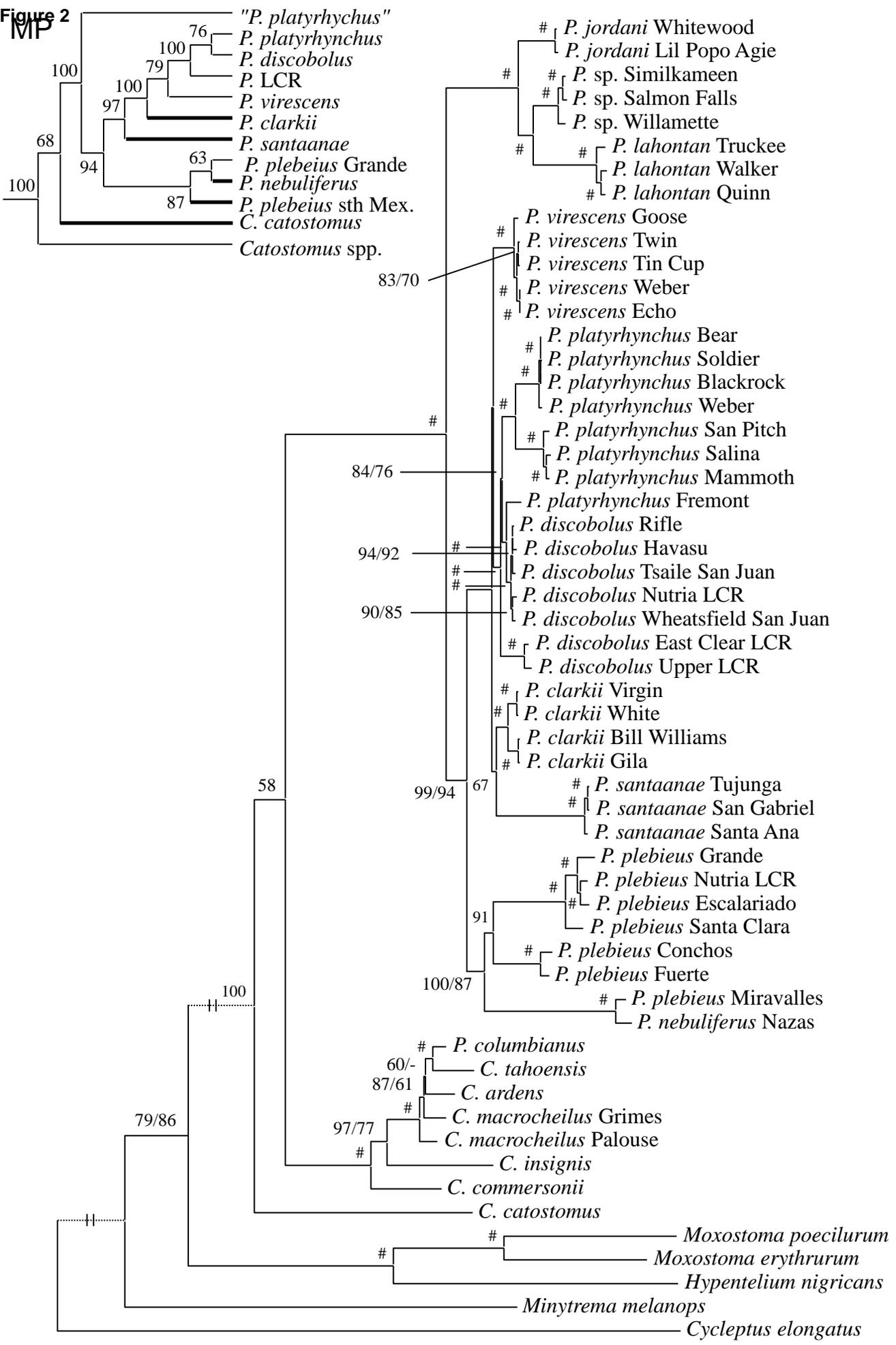


Figure 3

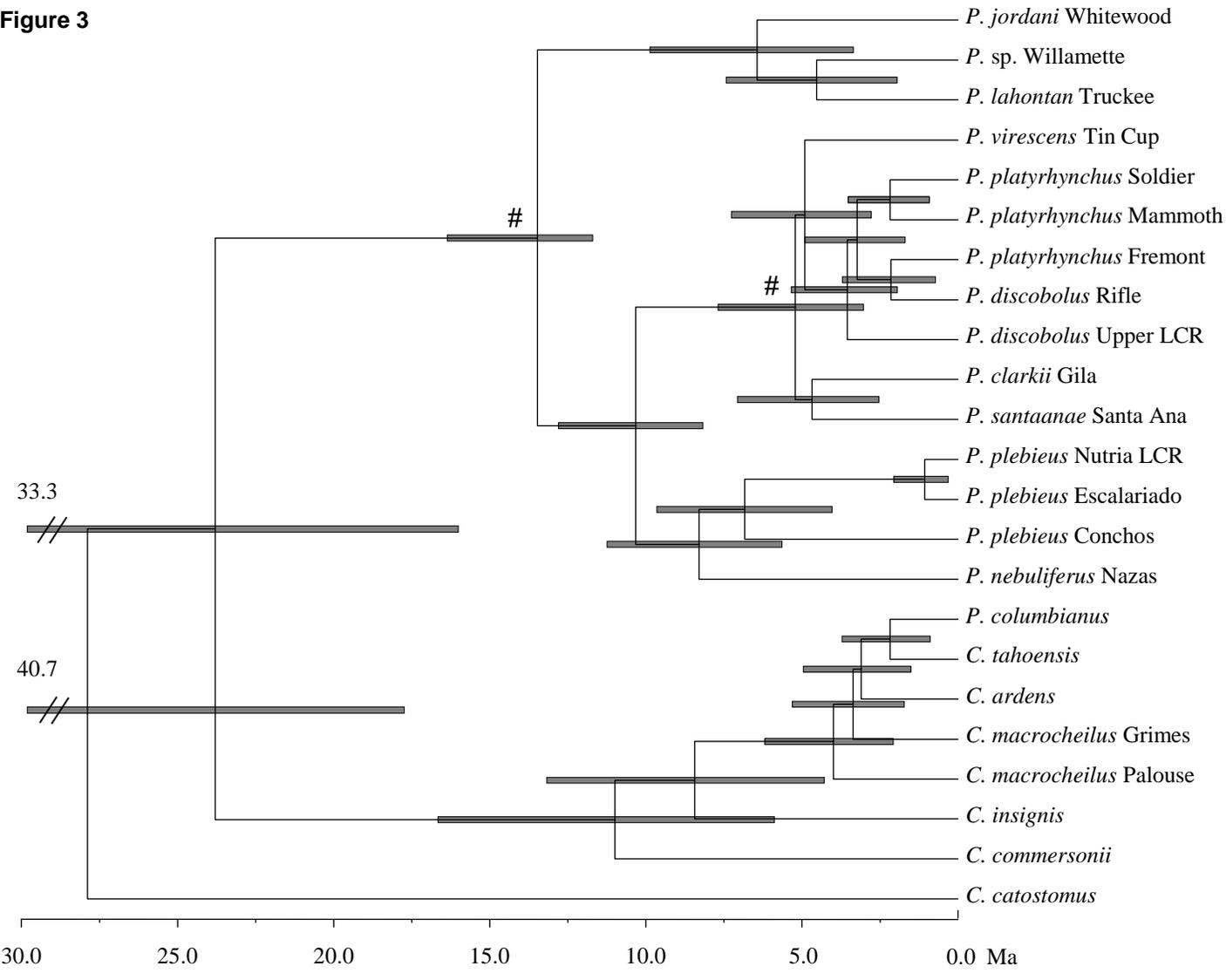
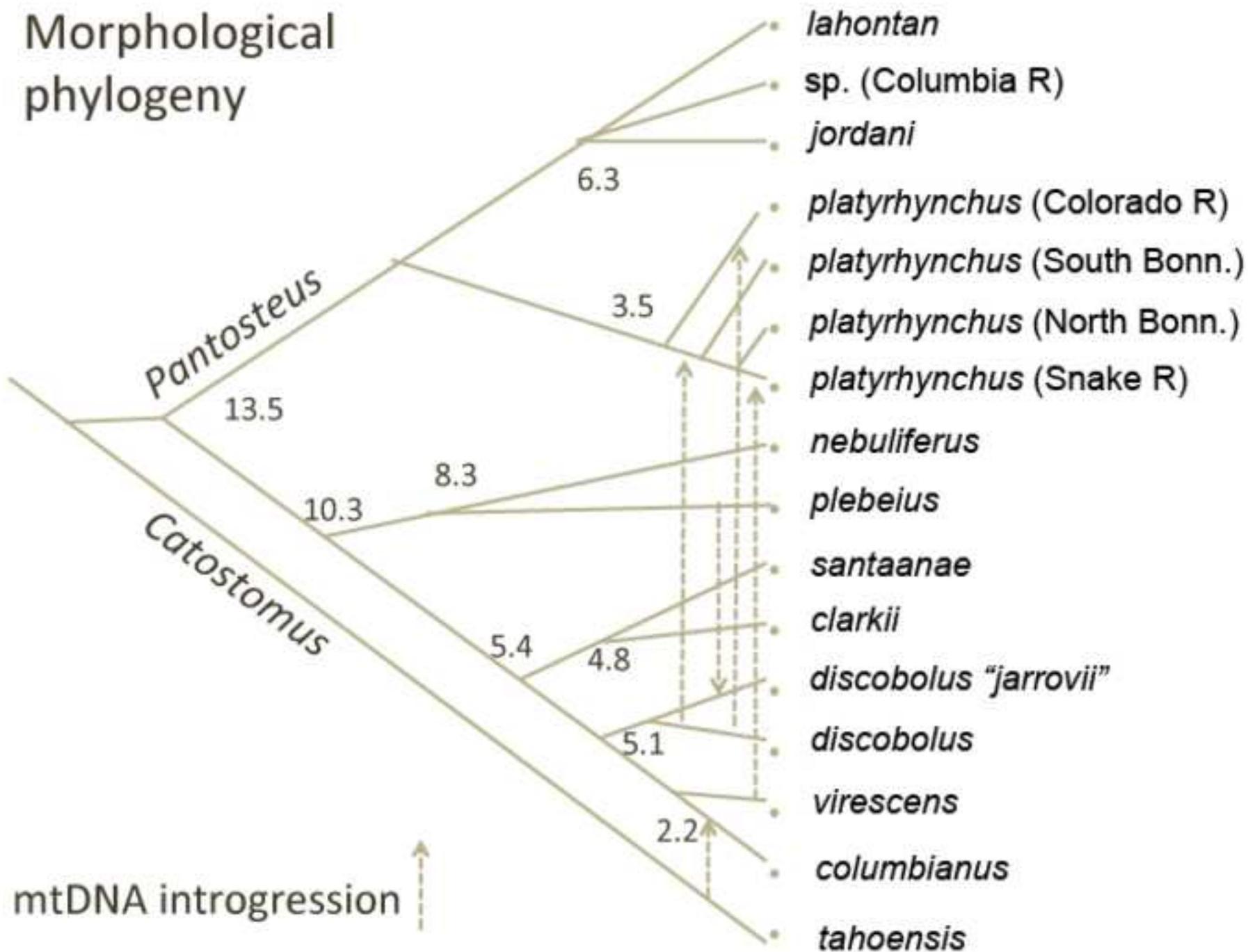


Figure 4  
[Click here to download high resolution image](#)



**Supplementary Material**

[Click here to download Supplementary Material: Supplementary Table 1.doc](#)