Historical Biogeography of the New-World Pupfish Genus *Cyprinodon* (Teleostei: Cyprinodontidae)


Analysis of mtDNA sequence variation (2,548 bp from ND2, cyt b, and part of the control region) indicates that the genus *Cyprinodon* began diverging in the Late Miocene from a common ancestor with *Megupsilon*, a monotypic genus on the Mesa del Norte of Mexico. The geographic pattern of mtDNA variation, with estimates of divergence time, suggests that by the end of the Miocene *Cyprinodon* occurred from the Atlantic Coast and West Indies to near the western margin of North America via ancestral Rio Grande and Colorado River systems. Phylogeographic structure within the major mtDNA complexes supports a variety of hypotheses from geology and previous phylogenetic analyses for Late Neogene connections among basins in southwestern North America now separated by formidable barriers to dispersal. Comparison of the mtDNA tree with previous phylogenetic inferences from allozymes indicates that reticulate evolution involving divergent lineages probably was important in the history of *Cyprinodon*.

FISH distributions and systematics have contributed importantly to a large body of hypotheses for the paleohydrology and associated biogeography of arid regions of North America, most of which have emphasized relatively recent geological/climatological events of Pleistocene times (Hubbs and Miller, 1948; Miller, 1948; reviews in Hocutt and Wiley, 1986). In contrast, Minckley et al. (1986) suggested, primarily from regional plate-tectonics, that the relevant time-frame for understanding the biogeography of extant western fishes might extend to Middle Tertiary. Correspondingly, a review of molecular phylogenetic studies of fishes in the Great Basin region suggested that many lineages are much older than generally appreciated (Smith et al., 2002).

In this paper, we use mitochondrial DNA variation to assess the historical biogeography of the pupfish genus *Cyprinodon*, the most widespread fish genus in a region comprising the warm deserts and southern Great Plains of southwestern North America (Miller, 1981). As a group, cyprinodontids of this region inhabit lowland waters where they are less likely than many other fishes to cross basin divides and, therefore, more likely to preserve ancient patterns of geographic variation (but see Minckley et al., 2002, for a possible exception in the Mexican highlands).

Approximately 30 of the 50 or so species of *Cyprinodon* occur in arid regions of southwestern North America where they are primarily distributed allopatrically as endemics to relatively small bodies of water, often single streams or spring-systems in isolated valley floors (Miller, 1981). The remaining 20 or so species occur as local endemics in the West Indies-Yucatan Peninsula region or as wide-ranging species with allopatric distributions in coastal or coastal-plain waters from Massachusetts to Venezuela (Smith et al., 1990; Wildekamp, 1995).
The paleohydrology of southwestern North America suggests many opportunities for secondary contact among previously allopatric fish lineages (Hubbs and Miller, 1948; Minckley et al., 1986; Smith and Miller, 1986). This and the present, primarily allopatric occurrences of pupfishes imply that, subsequent to most instances of secondary contact, the number of lineages in basins of concern was reduced to one by hybridization and genetic introgression or by extinctions leaving no genetic remnant.

Studies of reproductive compatibility among extant pupfishes suggest that ancient contact between divergent pupfishes generally would have led to genetic introgression. Laboratory experiments with three members of a small species flock on the Yucatan Peninsula revealed interfertility for all combinations and indicated pre-mating reproductive isolation for only one of the three (Strecker and Kodric-Brown, 2000). Experiments with allopatric pupfishes show little evidence of pre- or post-mating isolation between divergent species (Turner and Liu, 1977; Cokendolpher, 1980; Vilwock, 1982). In the wild, hybridization is indicated in all three morphologically well-studied instances of contact between species (Minckley, 1969; Humphries and Miller, 1981; Minckley and Minckley, 1986), and genetic introgression exists in all known situations where endemic species were exposed to anthropogenic introduction of a non-native pupfish (Echelle and Echelle, 1994, 1996; Childs et al., 1996), including one instance of introgression over more than 400 river-kilometers in less than five years (Echelle and Connor, 1989).

Hybridization confounds attempts to recover phylogenetic relationships, but it can, at the same time, inform reconstructions of biogeographic history because it might represent the only evidence of past contact between lineages (e.g., DeMarais et al., 1992). In this study, several such instances are indicated by the present geography of pupfishes and conflicting phylogenetic inferences from allozyme variation (Echelle and Echelle, 1992, 1993a, 1998) and the mtDNA results presented here.

**Materials and Methods**

**Sampling.** We obtained collections of 26 species of *Cyprinodon* from 38 localities (Material Examined, Fig. 1). We also used GenBank sequences for nine additional species not included in our collections (see below). The complete dataset includes 24 of the 29 historically extant species of *Cyprinodon* in arid regions of southwestern North America and 11 of the approximately 20 species (some undescribed) known from coastal Atlantic and Gulf of Mexico region and the West Indies. One of the excluded species (*C. salvadori*) is a recently described form from the Rio Conchos basin in Chihuahua (Lozano, 2002) that is similar to *C. eximius* and another (*C. latifasciatus*) has been extinct since the 1950s (Miller, 1964). The three remaining, aridland species excluded from the analysis (*C. ceciliae*, *C. inmemoriam*, and *C. longidorsalis*) represent a monophyletic group of four similar species (Lozano and Contreras, 1993; Echelle et al., 1995) from the Sandia Basin, Nuevo Leon, one of which, *C. veronicae*, is included here. This group was recently extirpated in the wild and is extinct except for *C. longidorsalis* and *C. veronicae*, which are being maintained in various aquarium facilities (Contreras and Lozano, 1996; pers. comm.).

The remaining species not examined are considered close relatives of *C. variegatus* (Miller, 1962; Smith et al., 1990; Lozano and Contreras, 1999): *C. bobmilleri* from a Gulf of Mexico drainage in Nuevo Leon, *C. hubbsi* from inland Flori-
da, two undescribed species from Cuba (M. Smith, pers. comm.), C. laciniatus and an undescribed species flock (Holtmeier, 2001) from the Bahamas, and C. higleyi and C. nicholls from Hispaniola. Surveys of mtDNA variation in pupfish from the Bahamas, Hispaniola, other sites in the West Indies, and the Atlantic and Gulf of Mexico coasts (Bunt, 2001; L. Fuselier et al., unpubl. data) indicate that our collections capture the major clades of mtDNA variation in this region.

Except for C. dearborni from Venezuela, species of *Cyprinodon* with large or fragmented geographic ranges were collected from multiple localities: six for the wide-ranging *C. variegatus* (three each from the Atlantic and Gulf of Mexico coasts), three encompassing the extremes of the range of *C. eximius*, two for the coastal form on the Yucatan Peninsula (*C. artifrons*), and one for each of two isolated populations of *C. rubrofluviatilis* (Brazos and Red rivers), *C. elegans* (Phantom Lake and San Solomon springs), and *C. nazas* (Rio Nazas and Rio Guatimapé). A form Miller (1976) considered a third population of *C. nazas* is treated as a separate species (*C*. sp., Rio Aguavatal) on the basis of a phylogenetic analysis of allozymes (Echelle and Echelle, 1998) and results presented herein. Finally, for the species from Cuatro Cièèegas, *C. atrorus* and *C. bifasciatus*, we used three specimens from the two most divergent mtDNA lineages detected in a survey of this basin (E. Carson, unpubl. data).

The phylogenetic analysis included three monotypic genera: *Jordanella*, a species restricted to fresh and brackish waters in Florida, and *Cualac* and *Megupsilon*, each of which is endemic to a separate, springfed system on the Mexican Plateau. Phylogenetic analyses of allozymes (Echelle and Echelle, 1999b), mtDNA (Parker and Kornfield, 1995), and morphology (Costa, 1997) indicate that *Cyprinodon* is closely related to these genera, with allozymes and mtDNA indicating a sister relationship to *Megupsilon*.

**Data collection.**—DNA was extracted from muscle or liver tissue by the methods of either Longmire et al. (1997) or Hillis et al. (1990). We used restriction fragment length polymorphism (RFLP) to identify variant mtDNA haplotypes in six collections of the wide-ranging *C. variegatus* (*n* = 12 each), each of two collections of *C. rubrofluviatilis* (*n* = 6 each), one of the two collections of *C. elegans* (*n* = 6), and the collections of *C. bovinus* and *C. pecosensis* (*n* = 6 each). After digesting each sample with 14 six-base recognizing restriction enzymes as described by Echelle and Dowling (1992), the fragments were end-labeled with α²³¹P, electrophoresed in 0.8% agarose gels, and visualized by autoradiography. This identified 19 composite haplotypes: *C. variegatus* = 9, *C. elegans* = 2, *C. rubrofluviatilis* = 4, *C. pecosensis* = 2, and *C. bovinus* = 2. We sequenced one sample of each composite RFLP haplotype, and, for the remaining collections, two specimens each except the following (*n* = 1): one of the two collections of *C. artifrons*, one of the three collections of *C. eximius*, two pupfishes of the western clade (*C. macularius* and *C. nevadensis*), the two Hispaniolan species, and the three genera, *Megupsilon*, *Cualac*, and *Jordanelle*.

A total of 65 specimens were sequenced for the entire NADH dehydrogenase subunit 2 (ND2; 1047 bp) and cytochrome *b* (*cytb*; 1140 bp) genes and a 361-bp section of the control region adjacent to tRNA<sup>Ami</sup>. The primers ND2B-L and ND2F-H (Broughton and Gold, 2000) were used to amplify ND2 in 50 µl reactions under the following conditions: 93 C for 3 min; 30 cycles of 94 C for 1 min, 48 C for 1 min, and 72 C for 2 min; 72 C for 2 min. We used those primers to sequence the amplified fragment with a Perkin-Elmer ABI Prism 377 automated sequencer. Haplotypes from the same collection locality were compared and all discrepancies were either verified by eye in reassessments of the computer-generated sequence histograms or by re-sequencing. In some instances, we also used degenerate internal primers (forward, 5′-CTTAAARTTGNNCTTYGCYCCCTT-3'; backward, 5′-RATNGARGTGAGKGRGGGCC-3') to obtain sequences originating around ND2 base positions 330 (forward sequence) and 700 (reverse sequence). Cytochrome *b* was amplified with primers LA and HA (Schmidt et al., 1998) in 50 µl reactions with the following thermal profile (30 cycles): 94 C for 1 min, 48 C for 1 min, and 72 C for 2 min. The same primers were used to sequence the amplified gene with an ABI Prism 377 automated sequencer. Haplotypes from the same sample were then compared and all discrepancies were verified as noted for ND2. For the control region, primers L15926 (Kocher et al., 1989) and H16498 (Meyer et al., 1990) were used in both amplification and sequencing. Eleven control-region sequences (seven *C. variegatus*, two *C. rubrofluviatilis*, and two *C. pecosensis*) were obtained with methods described by Meyer et al. (1994); the remaining sequences were obtained with the amplification and sequencing conditions described for ND2.

To allow re-assessment of relationships in a clade of ten species referred to as the western pupfishes (Echelle and Echelle, 1992), including three species not sequenced in this study (*C. diabolis*, *C. eremus*, and *C. salinus*), we used 14
ND2/control-region sequences from GenBank: 12 from Duvernell and Turner’s (1998) survey of the four species of the Death Valley System (AF028281–91, AF028293, AF028300–9), and two from a survey (Echelle et al., 2000) of C. iberus (AF199001, AF199002, AF198983, and AF198984).

Finally, to more thoroughly represent the pupfishes of the Yucatan Peninsula, we used 14 control-region sequences (GenBank accession numbers AY006375–6, AY006378–86, AY007978–9, L37113) representing all haplotypes detected in 17 specimens of C. artifrons and 5–12 fish from each of six species in a species flock in Laguna Chichancanab (Strecker et al., 1996; Strecker, 2002). The monophyly of the Yucatan pupfishes has not been assessed in an analysis that included pupfishes from other regions.

Phylogenetic analyses.—Sequence alignment was performed with CLUSTAL W (Thompson et al., 1994). Indels were treated as missing data. Unless otherwise mentioned, all analyses were performed with PAUP* (vers. 4.0b8w, D. L. Swofford, PAUP*: phylogenetic analysis using parsimony [*and other methods], Sinauer, Sunderland, MA, 2002). The complete dataset, including the GenBank sequences (control-region for the Yucatan pupfishes and ND2/control region for the western pupfishes), was used in a heuristic, maximum-parsimony analysis of relationships. For this analysis, we used TBR branch swapping, 20 random addition replicates, and equal weighting of characters. To avoid excessive computer time resulting primarily from low resolution for the Yucatan sequences, maximum number of trees for branch-swapping was held to 100 and branches with minimum lengths of zero were collapsed. Using the same restrictions, nonparametric bootstrapping (Felsenstein, 1985) with 500 replicates (ten random addition sequences per replicate) was used to assess support for clades. To compare the fast-parsimony method used here with a more thorough search, we subjected a dataset comprising only the ND2/cytb/control region sequences we detected to the same analyses except that branches of length zero were not collapsed and there was no limit on number of trees held for branch swapping.

For a second approach to phylogenetic analysis, we subjected the composite ND2, cyt b, and control-region haplotypes detected in this study to the Bayesian method provided by MrBayes 2.0 (Husonbeck and Ronquist, 2001). ModelTest 3.04 (Posada and Crandall, 1998) indicated that the base-substitution model best fitting the data was the general-time-reversible model with some sites assumed to be invariable and with variable sites following a discrete gamma distribution (GTR + I + Γ). With this model, we first performed an analysis in which the nucleotide substitution parameters were treated as unknown variables with uniform prior values and estimated as part of the analysis. The analysis began with random starting trees and was run for 1 × 10⁶ generations, with the initial 1,000 cycles discarded as burn-in. In a separate Bayesian analysis, we partitioned the characters into seven classes (control region and base positions 1, 2, and 3 for both ND2 and cyt b) and used a site-specific gamma model (GTR + SS Γ). The analysis was run for 1 × 10⁶ generations, with the initial 3,600 cycles discarded as burn-in. To assure stationarity, we graphically monitored the fluctuating value of the likelihood in both Bayesian analyses and ran the GTR + I + Γ analysis three times. Bayesian analysis produces, for all tree nodes, true posterior probabilities under the assumed substitution model (Rannala and Yang, 1996). We considered clades significantly supported when probabilities were > 0.95.

Estimating divergence times.—To assess the robustness of our estimates of divergence time, we used two different approaches, one based on a uniform ND2 clock developed independently of our data, and another that is based on our ND2 data and allows rate variation in different parts of the tree. In the first approach, we used a molecular clock for the genus Aphanius, which is sometimes mentioned as the Old World equivalent of Cyprinodon (e.g., Villwock, 1982). Both groups include extremely hardy, arid-land or coastal cyprinodontids existing in a highly fragmented aquatic landscape. Two reasonably well-understood vicariant events are postulated for Aphanius: (1) the development, 5.5 million years ago (Mya), of the Strait of Gibraltar, which apparently initiated divergence of A. iberus and A. baeticus (Perdices et al., 2001; Doadrio et al., 2002) and (2) the transgression, 13 Mya, of the Red Sea into the Wadi Sirhan of Jordan, resulting in divergence producing the sister species A. sirhani and A. dispar (Hrbek and Meyer, 2003). To calibrate an ND2 clock for Aphanius, we first produced a minimum-evolution tree for the 49 sequences detected in 16 species of the genus (Hrbek and Meyer, 2003; GenBank accession numbers AF449287–335), with Valencia letourneuxi as the outgroup (accession numbers AF449336–7). The tree was based on gamma-corrected, Tamura-Nei distances (maximum-likelihood estimate of α = 0.37). To adjust for inconsistent substitution rates detected in rela-
Fig. 2. Strict consensus of maximum-parsimony trees. Terminal nodes with locality numbers in parentheses represent haplotypes generated in the present study. Taxa in boxes represent GenBank sequences.

Weakly divergent (uncorrected $p_{50.001\pm0.009}$) terminal nodes are collapsed to one except *C. nevadensis*, for which all are retained to show paraphyly with respect to *C. diabolis*, and the Lake Chichancanab/*C. artifrons* node, a polytomy of shared haplotypes. Numbers above branches are bootstrap support. Asterisks = Bayesian probabilities 95% (model = GTR + SS). Nodes uptree of arrows include GenBank sequences not subjected to Bayesian analysis. The node with a dashed line received Bayesian and bootstrap support but was not retained in the strict consensus of the shortest maximum-parsimony trees.

To preserve examples of mtDNA paraphyly, three sequences from *C. v. variegatus*, four from *C. eximius*, and two from *C. nevadensis*. To determine if the ND2 tree is consistent with the total-information topology in Figure 2, we performed the more thorough maximum-parsimony analyses described above on the subset of 38 sequences used in estimates of time. The results were completely consistent with Figure 2 except that fewer nodes received bootstrap support.

The multidivtime procedure requires some biologically plausible prior estimates for certain parameters, but they can be vague (hence they are given large error estimates) because the Markov chain Monte Carlo procedure (MCMC) employed in the analysis is relatively robust to
their variation (Weigmann et al., 2003). We used 0.10 (SD = 0.10) as the prior estimate of the mean for rate of substitution per time unit (1 unit = 10 million years) at the ingroup root and, following rationale of Weigmann et al. (2003), 1.0 as a log-scale estimate of the variance in rate of change per unit time. As the most plausible estimate of time from the present to the root of the ingroup (Cualac, Megapus-silon, and Cyprinodon), we used 11 Mya (1.1 time units; SD = 1.1) based on results from the molecular clock developed with Aphanus. One might view this as a violation of the intent to use two independent approaches. However, our experiments with values of 5 Mya, 11 Mya, and 20 Mya, indicated that the analysis was relatively insensitive to this prior. Given the internal time constraints described below, the posterior solutions were, respectively, 11.7 Mya, 13.4 Mya, and 14.6 Mya, and the results elsewhere in the tree had no effect on biogeographic conclusions. For the remaining priors, we followed Weigmann et al. (2003) and used the default values in multidivtime. The analysis included a burn-in period of 100,000 cycles of the MCMC procedure followed by sampling of parameters every 100 cycles until 10,000 samples were collected. This design appeared adequate because means of divergence time and confidence intervals from different runs were similar to the third significant figure.

For internal time constraints we restricted two nodes as follows: (1) the lower and upper limits for time of divergence between C. macularius and C. evenus were set at 1.1 Mya and 1.7 Mya, representing the timing of the Pinacate volcanic eruptions (Lynch, 1981) assumed to have initiated divergence of those two species (Turner, 1983; Echelle et al., 2000), and (2) the upper limit for divergence of C. rubiflavis/tiatisis from species in the Pecos River was set at 4.5 Mya. The latter represents the 4–5 Mya cessation of Ogallala deposition in the southern Great Plains (Chapin and Cather, 1994). As described later, this led to formation of the upper Pecos River, which, in turn, helps explain the zoogeography of pupfish, other fishes (Echelle and Echelle, 1978; Smith and Miller, 1986), and snails (Hershler et al., 2002).

Reticulate evolution.—We searched for evidence of potential reticulate evolution by comparing phylogenetic inferences from mtDNA with those from allozymes. Conclusions are tentative because of low resolution in the allozyme results. Pupfishes exhibit high levels of allele sharing for allozymes, with few fixed differences among groups (Turner, 1974; Echelle and Echelle, 1992, 1998). Consequently, the phylogenetic trees are based on allele-frequency parsimony (Berlocher and Swofford, 1997), and clades often are supported by frequencies of alleles shared with other clades. To reduce errors from poor resolution, we consider only those conflicts supported by at least one synapomorphy allozyme allele not shared with other nodes.

Results

We detected 58 mtDNA haplotypes among the 141 specimens surveyed (by RFLP or directly by sequencing). All sequences are deposited in GenBank (accession numbers AY901992–AY902165). The control-region alignments had eight indels; none occurred in ND2 or cyt b. Sequences were identical in six instances where two specimens were sequenced from the same collection: C. dearborni, C. artifrons from Cancun, C. tularosa, C. macroilepis, Aguanaval pupfish (C. sp.), and C. albivels. With one exception (C. eaximus from Rio Florida; uncorrected p = 0.057), level of divergence among haplotypes from the same collection was low, ranging from 0.001 to 0.013. Maximum uncorrected p-values were 0.176 and 0.223 for, respectively, cyt b and ND2 (for both genes, fondanella vs. a species of Cyprinodon) and 0.194 for the control region (Cualac vs. three samples of Cyprinodon). There were 777 parsimony informative characters representing 26% (96/361), 28% (322/1140), and 34% (359/1047), respectively, of the control region, cyt b, and ND2 bases.

Phylogenetic analyses.—The faster approach to maximum-parsimony analysis of the complete dataset of 58 composite ND2/cyt b/control-region haplotypes and 28 partial haplotypes obtained from GenBank produced a topology (Fig. 2) that, for the samples in common, was identical to the topology obtained using a more thorough parsimony approach with only the 58 haplotypes we detected, and differences in levels of bootstrap support were negligible. Thus, only results from the complete dataset are described here. This analysis produced 446 trees (length = 2,715; CI = 0.42; RI = 0.80) differing only in topology within terminal clades, primarily the one consisting of 15 haplotypes from the Yucatan Peninsula. All nodes in the strict consensus tree received greater than 50% bootstrap support except for two deep nodes within Cyprinodon (Fig. 2).

The results for the western pupfish group (clade 5, Fig. 2) were, except for C. radious, consistent with previous mtDNA-based analyses (Echelle and Dowling, 1992; Duvernell and
Turner, 1998). Those studies supported, albeit weakly, a sister relationship between \textit{C. radiosus} and the \textit{C. macularius-\textit{C. eremus}} clade, whereas we found weak bootstrap support (56\%) for \textit{C. radiosus} as the basal member of the western pupfishes, a result consistent with lack of Bayesian support for relationships among the four basal nodes of the clade (Fig. 2).

The maximum parsimony bootstrap analysis supported (98\%) mtDNA monophyly for the Yucatan Peninsula pupfishes, \textit{C. artifrons}, and the Lake Chichancanab species flock (Fig. 2). The 14 haplotypes previously detected in this region and an additional haplotype we found in \textit{C. artifrons} were weakly divergent (uncorrected \( p = 0.000-0.014 \)), and relationships were unresolved excepting 65\% support for a cluster of two of the nine haplotypes known only in \textit{C. artifrons}.

The two models of substitution (GTR + I + \Gamma and GTR + SS \Gamma) used in the Bayesian analysis of the 59 ND2/cytb/control-region haplotypes resulted in identical topologies, and the nodes with probabilities > 0.95 were the same. With one exception, the Bayesian and maximum parsimony cladograms were identical when collapsed to nodes having, respectively, > 0.95 Bayesian probabilities or >50\% maximum parsimony bootstrap support (Fig. 2). The exception is the maximum-parsimony placement (67\% bootstrap support) of clade 1 (Yucatan pupfishes) as the basal lineage of \textit{Cyprinodon}, an inference not receiving Bayesian support.

None of the nodes connecting the five primary mtDNA clades in \textit{Cyprinodon} received Bayesian or maximum parsimony support except the just mentioned conflict for clade 1. Correspondingly, the optimal trees from the two approaches showed no areas of congruence for relationships among those clades. Thus, relationships among the five basal clades within \textit{Cyprinodon} are unresolved.

Bayesian analysis and maximum-parsimony resulted in nearly identical, well-supported topologies within each of the seven major mtDNA clades of \textit{Cyprinodon} (Fig. 2), all of which, except for a minor difference within clade 2A, were consistent with the minimum-evolution tree (Fig. 3) to which the molecular clock from \textit{Aphanius} was applied. The one disparity involved the Red and Brazos river forms of \textit{C. rubrofluviatilis}. The strict-consensus, maximum-parsimony tree (Fig. 2) supported monophyly of this species, but with no bootstrap support, whereas the Bayesian tree indicated paraphyly with respect to the \textit{C. pocosensis-\textit{C. bovinus-\textit{C. elegans}}} clade, also with no statistical support. Thus, we consider the question of monophyly for \textit{C. rubrofluviatilis} to be unresolved.

Hereafter in this paper, we refer to the seven major clades with names indicating region of primary occurrence (Figs. 3 and 4A): clade 1 = Yucatan (Laguna Chichancanab and coastal areas on the Yucatan Peninsula), 2A = maritime (coastal Atlantic, Gulf of Mexico, and Hispaniola; Tularosa Basin of the northern Chihuahuan Desert is the only occurrence in the interior of North America), 2B = southern Great Plains-northern Chihuahuan Desert (Red, Brazos, and Pecos river basins), 3 = Sandia-Potosi (Sandia and Potosi basins, Nuevo Leon, Mexico), 4A = Rio Conchos-middle Rio Grande (Rio Conchos, Alamito Creek, and Devils River; all in the middle Rio Grande Basin), 4B = Old Río Nazas (rios Nazas, Aguanaval and Mezquital, Laguna Santiaguillo, and Cuatro Cienegas Basin), and 5 = western (Guzmán Basin, lower Colorado River, and Death Valley System).

\textit{Divergence times.—}The two methods of estimating divergence times gave similar results throughout the tree (Table 1). The largest absolute differences were at the base of the tree, with differences of 2.5 Myr and 1.7 Myr in estimated times of divergence between \textit{Cyprinodon} and, respectively, \textit{Cuaci} and \textit{Megupsilon}. The differences were effectively negligible within \textit{Cyprinodon}. The major discrepancy was the time (1.9 Myr vs. 2.6 Myr) separating the Aguanaval pupfish (\textit{C. sp.}) from the Cuatro Cienegas pupfishes (\textit{C. atrorus} and \textit{C. bifasciatus}), and none of the differences resulted in conflicting biogeographic hypotheses.

\textit{Mitochondrial DNA/alloyzme conflicts.—}The synthesis of the various phylogenetic studies of mtDNA and allozyme variation (Fig. 5) indicates six instances of potential conflict, five supported by synapomorphic allozyme alleles. The sixth conflict involves a polymorphism in the Río Florido population of \textit{C. eximius} for two markedly divergent mtDNAs (uncorrected divergent = 5.7\%), one being only 1.3\% divergent from the mtDNA detected in \textit{C. macrolepis}. Thorough geographic surveys of mtDNA variation in \textit{C. bifasciatus} (E. Carson, unpubl. data) and \textit{C. tularosa} (Stockwell et al., 1998) indicate fixation for the mtDNA lineages involved in the mtDNA/alloyzme conflicts for those species. Similar fixation in the other instances of conflict seems likely despite small sample sizes: \textit{C. elegans} was sampled from two localities in its small range \((n = 2\) and 6\)); \textit{C. pachycephalus} \((n = 2\) is restricted to an extremely small spring system; and \textit{C. radiosus} \((n = 2\), which has un-
Fig. 3. Estimated timing of cladogenetic events. Ages indicated are as follows (20.3 Mya in Early Miocene) and the start of each of the following (from Berggren et al., 1995): Late Miocene (11.2), Pliocene (5.2), and Pleistocene (2.5 Mya). Based on the ultrametric, minimum-evolution tree, with timing based on the clock (2.8%/Mya) developed for *Aphanius*. Timing closely resembles that from the multidivtime approach (Table 1), but that approach does not produce an ultrametric tree in which all terminal nodes are at time zero.

**DISCUSSION**

Estimating divergence time from molecular divergence is fraught with difficulties (Doyle and Donoghue, 1993; Bermingham et al., 1997), and our estimates have wide confidence intervals (Table 1). Nonetheless, the close agreement reported here for divergence times derived by two extremely divergent approaches indicates that the means for the estimates are sufficiently robust to provide reasonable hypotheses of event times. We also note that the estimates for the western pupfish group (clade 5) are similar to those obtained from still another approach, where sequence divergences estimated from restriction-site data for whole mtDNA (Echelle and Dowling, 1992) were used with a molecular clock based primarily on the goodeid fossil record and Webb’s (1998) mtDNA sequence data for goodeids (Smith et al., 2002).

Our results support the monotypic genus *Megupsilon* as the sister-group of a monophyletic *Coryphodon*, in agreement with results from allozymes for a diversity of pupfishes (Echelle and Echelle, 1993b), as well as those from analyses...
Fig. 4. Geography and vicariance history of mtDNA clades in *Cyprinodon*: (A) Geographic distribution of the seven primary mtDNA-clades. 1 = Yucatan, 2A = maritime, 2B = southern Great Plains-northern Chihuahuan Desert, 3 = Sandia-Potosi, 4A = Rio Conchos-middle Rio Grande, 4B = Old Rio Nazas, 5 = western. Clade 2A is the maritime group, with a basal *C. variegatus ovinus-C. bondi* clade, and a clade with the following topology: (C. *tularosa* [(C. *dearborni* [C. *variegatus*, C. sp. from Lake Enriquillo]]). (B) Some hypothetical paleosystems (arrows) and estimated times of vicariance for pupfish in each system. Basin abbreviations as in Fig. 1. Lettering in boxes = age of associated vicariance (average of the estimates in Table 1) and clades involved (labeled as in Fig. 2).

of mtDNA (Parker and Kornfeld, 1995) and morphology (Costa, 1997) that included, respectively, two and three species of *Cyprinodon*. Estimated divergence time suggests that the genera began diverging in Late Miocene about 7–9 Mya. Branch lengths and unresolved relationships among the five most basal mtDNA clades of *Cyprinodon* indicate origins rather closely together in time ~6–7 Mya (Table 1; Fig. 3). This is consistent with the estimated age of the only fossil assigned to the genus *Cyprinodon*, a specimen from Death Valley of Late Miocene or Early Pliocene age (Miller, 1945, 1981).

The mtDNA results provide little support for the suggestion (Echelle and Echelle, 1992, 1998) that various inland lineages might have arisen independently from coastal forms. This hypothesis predicts that inland clades on the mainland will be intercalated among coastal clades (Lovejoy and Collette, 2001), and, with one exception, there is no evidence for such a pattern. The exception is *C. tularosa* of the Tularosa Basin in southcentral New Mexico, whose mtDNA arose rather recently (~1.6–1.9 Mya) within the maritime clade. One or more of the three major inland clades might have a similar maritime origin but unresolved relationships among these clades preclude testing this hypothesis.

**Atlantic and Gulf Coasts and West Indies.**—The pupfishes from this region represent two deeply divergent mtDNA clades: the Yucatan clade from the Yucatan Peninsula (clade 1, Fig. 2) and the wide-ranging maritime clade (2A), which, with one exception, *C. tularosa* from southcentral New Mexico, occurs in coastal mainland pupfishes of North and South America and islands of the West Indies. The monophyly of the Yucatan clade is consistent with the suggestion from genetic similarities that the coastal species, *C. artifrons*, was ancestral to the Laguna Chichancaclanab species flock, a complex of six ecologically diverse taxa that, based on past lake levels (Covich and Stuiver, 1974) and low levels of genetic divergence, might be only 8,000 years old (Humphries, 1984; Strecker et al., 1996).

Because of limited sampling, we cannot exclude the possibility that the Yucatan mtDNA clade occurs outside of the Yucatan Peninsula or that *C. artifrons*, a wide-ranging pupfish on the peninsula and nearby islands, does not carry representatives of other mtDNA clades detected in this study. However, surveys of pupfishes in the Bahamas (Bunt, 2001) and on Hispaniola, Grand Cayman, and the Florida Keys (L. Fuselier et al., unpubl. data) so far reveal haplotypes belonging only to the maritime clade.

**Guzmán Basin, Lower Colorado River, and Death Valley System.**—ND2 divergence among basal lineages of the western pupfish clade indicates that, by ~3.6–3.8 Mya, pupfishes occurred in the regions now occupied by the Death Valley System (Owens Valley and the Amargosa River-Death Valley region), lower Colorado River, and Guzmán Basin of northwestern Chihuahua (Fig. 4B).
There are two competing hypotheses for the lower Colorado River-Death Valley connection: (1) a Miocene-Early Pliocene route via an Amargosa-Colorado paleoriver that preceded development of the Death Valley trough (Howard, 1996), or (2) a Plio-Pleistocene route via connections among successively lower basins between the Death Valley System and the lower Colorado River (Miller, 1948, 1981; Smith et al., 2002). The older Amargosa-Colorado paleoriver seems more compatible with the mtDNA estimate of divergence time. However, as discussed below, the possibility of a subsequent invasion of the Death Valley System would help explain other genetic aspects of the western pupfish clade.

Branch lengths in the mtDNA tree (Fig. 3) suggest a Pliocene connection between Guzmán Basin and the lower Colorado River at about the time of the Colorado River-Death Valley connection. At that time Guzmán Basin was occupied by Lake Cabeza de Vaca, an immense Miocene to Late Pleistocene complex of basins extending from southwestern Chihuahua into southern New Mexico (Strain, 1966; Seager et al., 1984). Topography and geological data, including a Miocene valley-basalt flow, indicate potential connections between this system and the Gila River (Miller, 1981), which flows west- erly to the lower Colorado River. The Gila River was integrated with the lower Colorado River by 5 Mya and probably earlier (Howard, 1996). The period from 5 Mya through the Pleistocene saw accelerated tectonic activity in the associated Rio Grande Rift (Seager et al., 1984), and such activity might have severed westward connections between the Lake Cabeza de Vaca complex and the lower Colorado River.

A problematic portion of the western-pupfish mtDNA clade is the indicated sister relationship
### Table 1. Comparison of mtDNA Divergence Times Based on Results Obtained from the Multidivtime Analysis and the Molecular Clock for Aphanius.

Except where otherwise noted, the label for each node consists of two taxa, one for each of the two descendant clades, some with locality numbers. Values given are the mean and 95% confidence interval from multidivtime and, for the Aphanius clock, the estimates based on the two calibrations for gamma-corrected Tamura-Nei divergence (2.4% and 3.1% per Mya) and the average of the two (2.8%).

<table>
<thead>
<tr>
<th>Node</th>
<th>Divergence time (Mya)</th>
<th>Multidivtime</th>
<th>Aphanius clock</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(18.4–23.8)</td>
</tr>
<tr>
<td>Jordanella/Cyprinodon</td>
<td>—</td>
<td>20.3</td>
<td>(9.8–12.7)</td>
</tr>
<tr>
<td>Cualac/Cyprinodon</td>
<td>13.4</td>
<td>10.9</td>
<td>(5.3–12.6)</td>
</tr>
<tr>
<td>Megupsilon/Cyprinodon</td>
<td>8.5</td>
<td>6.8</td>
<td>(6.1–7.9)</td>
</tr>
<tr>
<td>Among clades 1–5 of Cyprinodon</td>
<td>5.5–7.2</td>
<td>5.5–5.8</td>
<td>(6.3; 5.5–10.2)*</td>
</tr>
</tbody>
</table>

#### Clade 2
- **Clade 2A/2B**
  - C. bondi vs C. ovinus
    - Divergence time (Mya): 2.5 (1.3–3.9)
    - Aphanius clock: 2.0 (1.8–2.4)
  - C. bondi/C. variegatus
    - Divergence time (Mya): 2.9 (1.6–4.4)
    - Aphanius clock: 2.4 (2.2–2.8)
  - C. tulansoa/C. variegatus
    - Divergence time (Mya): 1.9 (1.0–3.2)
    - Aphanius clock: 1.6 (1.4–1.8)
  - C. rubrofluviatilis/C. pecosensis
    - Divergence time (Mya): 3.7 (2.4–4.5)
    - Aphanius clock: 3.6 (3.3–4.2)
  - C. rubrofluviatilis (1 & 2)
    - Divergence time (Mya): 3.5
    - Aphanius clock: 3.4
  - C. pecosensis/C. elegans
    - Divergence time (Mya): 2.6 (1.5–3.8)
    - Aphanius clock: 2.3 (2.1–2.7)
  - C. pecosensis/C. bovinus
    - Divergence time (Mya): 0.6 (0.2–1.2)
    - Aphanius clock: 0.5 (0.5–0.6)

#### Clade 3
- **C. veronicae/C. alvarezi**
  - Divergence time (Mya): 2.1 (1.0–3.6)
  - Aphanius clock: 2.0 (1.8–3.3)

#### Clade 4
- **Clade 4A/4B**
  - C. eximius (7) / C. eximius (8)
    - Divergence time (Mya): 0.1 (0.05–0.4)
    - Aphanius clock: 0.4 (0.4–0.5)
  - C. eximius (7) / C. eximius (9)
    - Divergence time (Mya): 0.4
    - Aphanius clock: 0.4
  - C. eximius (7) / C. macrolepis
    - Divergence time (Mya): 2.9 (1.6–4.9)
    - Aphanius clock: 3.4 (3.0–3.8)
  - C. macrolepis/C. pachycephalus
    - Divergence time (Mya): 3.0 (1.6–4.9)
    - Aphanius clock: 3.4 (3.0–3.7)
  - C. meeki/C. nazas
    - Divergence time (Mya): 3.0
    - Aphanius clock: 3.4
  - C. meeki/C. atrorus
    - Divergence time (Mya): 4.1 (1.6–4.7)
    - Aphanius clock: 4.8 (3.1–4.0)
  - C. atrorus/C. sp. (Aguanaval)
    - Divergence time (Mya): 1.9 (0.9–3.3)
    - Aphanius clock: 2.6 (2.4–3.1)

#### Clade 5
- **C. radiosus/C. macularius**
  - Divergence time (Mya): 3.8 (2.3–5.0)
  - Aphanius clock: 3.6 (3.5–4.5)
- **C. macularius/C. eremus**
  - Divergence time (Mya): 1.3 (1.1–1.6)
  - Aphanius clock: 1.5–1.9
- **C. macularius/C. diabolis**
  - Divergence time (Mya): 3.5
  - Aphanius clock: 3.9
between *C. fontinalis* of Guzmán Basin and the Amargosa River-Death Valley pupfishes (*C. nevadensis*, *C. diabolis*, and *C. salinus*). This is enigmatic because the suggested sister groups are separated by the lower Colorado River region, which is inhabited by *C. macularius*. Allozymes support, albeit weakly, the geographically more plausible hypothesis that the Amargosa River-Death Valley pupfishes are part of a monophyletic complex in the Death Valley System that includes *C. radiosus* in nearby Owens Valley. This conflict is discussed in a later section.

**Table 1.** Continued.

<table>
<thead>
<tr>
<th>Node</th>
<th>Divergence time (Mya)</th>
<th>Multidivtime</th>
<th>Aphanius clock</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. fontinalis/C. diabolis</em></td>
<td>2.7</td>
<td>(1.5–4.2)</td>
<td>(2.8–3.7)</td>
</tr>
<tr>
<td><em>C. nevadensis/C. salinus</em></td>
<td>1.0</td>
<td>(0.4–1.9)</td>
<td>(0.9–1.2)</td>
</tr>
<tr>
<td><em>C. nevadensis/C. diabolis</em></td>
<td>0.5–0.7$^d$</td>
<td>(0.1–1.4)</td>
<td>(0.5–0.8)</td>
</tr>
<tr>
<td><em>C. nevadensis/2–3 haplotypes</em></td>
<td>0.7</td>
<td>(0.2–1.4)</td>
<td>(0.6–0.8)</td>
</tr>
</tbody>
</table>

$^a$ In parentheses, average divergence time among the five clades and the range from the lower 95% confidence limit for the smallest interclade value (5.5) to the higher limit for the largest (7.2).

$^b$ Timing constrained to have an upper limit of 4.5 Mya.

$^c$ Timing constrained to be between 1.1 Mya and 1.7 Mya.

$^d$ Times shown are between *C. diabolis* and the two haplotypes of *C. nevadensis* included in the analysis.

$^e$ See Materials and Methods for haplotypes included.

Ancestral Rio Grande and Mesa del Norte.—The three Guzmán Basin members of the western pupfish clade (*C. albivelis*, *C. fontinalis*, and *C. pisteri*) and most other southwestern pupfishes occur in remnant streams and endorheic basins of Smith and Miller’s (1986) ancestral Rio Grande System. In the Miocene and Pliocene, the upper Rio Grande occupied a developing north-south succession of fault basins in the Rio Grande Rift from Colorado to southwestern New Mexico. By 3.4–4.5 Mya, basins of south-central New Mexico, including parts of the Lake Cabeza de Vaca complex, probably formed the terminus of an ancestral upper Rio Grande (Seager et al., 1984; Repenning and May, 1986; Mack et al., 1993). By ~2 Mya, the river was emptying into the Tularosa-Hueco complex of basins at Fillmore Pass 40 km north of the present city of El Paso (Mack et al., 1996). From there, it probably flowed to the Gulf of Mexico in roughly the present course (Hawley et al., 1976).

The Tularosa Basin-lower Rio Grande connection ~2 Mya is supported by the mtDNA estimate for the age of *C. tularosa* (~1.6–1.9 Myr) and that of a springsnail *Juturnia tularosae* (2–3 Myr; Hershler et al., 2002). The closest mtDNA relatives of these endemics to the Tularosa Basin are, respectively, the coastally distributed *C. v. variegatus-C. dearborni* clade and *J. kosteri* in the Pecos River. Early Pleistocene uplifting and westward tilting of the San Andres-Franklin mountains diverted flow away from the Tularosa-Hueco Basin and to the west of El Paso into Mesilla Basin of the Lake Cabeza de Vaca complex (Strain, 1966; Hawley et al., 1976). This contributed to the isolation of Tularosa Basin and temporarily severed the connection between the upper and lower Rio Grande.

Approximately 0.8 Mya (Mack et al., 1993), the upper and lower Rio Grande basins were re-integrated via a more southern connection when the upper basin was diverted from the Lake Cabeza de Vaca complex by basin overfill and canyon cutting near El Paso (Hawley et al., 1976; Seager et al., 1984). The history of connections between the lower Rio Grande and its major tributaries is not well understood, but a minimum age for the lower Pecos River is ~13 Myr based on dating of Gatuña formation sinkhole-fill, near Orla, Reeves County, Texas, and the connection between the Río Conchos and the Rio Grande might be as recent as early Pliocene (J. W. Hawley, pers. comm.). The Río Conchos crosses ancient mountain ranges separated by valley floors partially filled with thick alluvium of Late Tertiary or Quaternary age, indicating that drainage was not continually through-flowing to the Rio Grande (Smith and Miller, 1986).

The pupfishes on the Mesa del Norte of Mexico east and south of Guzmán Basin carry a well-supported mtDNA complex (clade 4A–B) that, early in the evolution of *Cyprinodon* (~5 Mya), split into two major clades, one now in the Río Conchos-Middle Rio Grande and one supporting recognition of the Old Rio Nazas, a pa-
lesosystem originally inferred from fish and herpetofaunal distributions (Meek, 1904; Conant, 1963). The Old Río Nazas would have extended from the western highlands of Mexico, across what is now the Chihuahuan Desert, to the Río Grande. The progenitor of the Old Río Nazas clade might have evolved in isolation on the Mesa del Norte, starting in early Pliocene time.

The mtDNA tree corroborates the indication from allozymes (Echelle and Echelle, 1998) of an Old Río Nazas clade (clade 4B) comprising the pupfishes of Durango and Zacatecas (rios Nazas, Aguanaval, and Mezquital, and Laguna Santiaguito) and one (allozymes) or both (mtDNA) pupfishes in the Cuatro Ciénegas Basin. This argues for one of three hypotheses (Arellano, 1951; Conant, 1963) for the connection of the Old Río Nazas with the Río Grande: namely via Río Salado (Fig. 4B), the Pleistocene outlet for Cuatro Ciénegas Basin. The ND2 divergence between the Río Aguanaval pupfish (C. sp.) and the Cuatro Ciénegas pupfishes indicates severance of a connection across the Chihuahuan Desert ~1.9–2.6 Mya.

The only fishes known from the small, endorheic Sandia and Potosí basins are the five species of the C. alvarezi-C. veronicae clade and the monotypic genus Megupsilon, suggesting that these are relics of an ancient fauna isolated in this region of the Mesa del Norte. Resolution of C. alvarezi-C. veronicae as one of the basal mtDNA clades in Cyprinodon and inclusion of the Sandia and Potosí basins in Smith and Miller’s (1986) Río Grande System indicates early isolation from the ancestral Río Grande (Fig. 3).

Southern Great Plains-Northern Chihuahuan Desert.—Divergence among pupfishes carrying the southern Great Plains-northern Chihuahuan Desert mtDNA clade might have involved fragmentation of a system (Gustavson and Finley, 1985; Conner and Suttkus, 1986) extending eastward from the southern Rocky Mountains to the Gulf of Mexico. This ancestral southern Great Plains system (Fig. 4B) would have comprised a complex of low-gradient streams occationally interconnected by cross-grading as they meandered over the relatively flat surface of the Miocene-to-Pliocene Ogallala formation (Conner and Suttkus, 1986). Vicariance and mtDNA lineage sorting from a polymorphic ancestor in this system might explain the short internodal branches (Fig. 3) among the three markedly divergent, basal clades of this complex, two representing the Red and Brazos river forms of C. rubrofluviatilis. The lack of support for a monophyletic C. rubrofluviatilis (see Results) also occurred in a phylogenetic analysis of allozymes (Echelle and Echelle, 1992).

Range fragmentation leading to the present geography of the southern Great Plains-northern Chihuahuan Desert mtDNA clade might have been triggered by cessation of Ogallala deposition, an event that occurred 4–5 Mya (Chapin and Cather, 1994) and was associated with entrenchment of the Red, Brazos, and upper Pecos rivers (Gustavson and Finley, 1985). For the Pecos River, dissolution of Permian salt beds and resultant landscape subsidence led to ponding of portions of eastwardly flowing streams, including the Brazos River, in a trough forming the proto-Pecos River Valley in eastern New Mexico. Subsequent integration of the upper and lower Pecos occurred through headward subsidence and erosion of the much older lower Pecos River. Estimated divergence time (~3.6 Mya) for basal members of the southern Great Plains-northern Chihuahuan Desert clade is consistent with this model of river evolution.

Caveats and reticulate evolution.—The estimated ages for some recently evolved populations/taxa may be substantial overestimates because of mtDNA polymorphism in ancestral species (see Avise, 2000). The size of this error relative to estimated age generally should be more pronounced for younger taxa. For example, the estimated age of C. diabolis (~0.5 Myr) may be an order of magnitude too great. Earlier workers suggested an age of only 10,000 to 20,000 years (Hubbs and Miller, 1948; Miller, 1981), which seems reasonable from geology. Devils Hole, the fault cavern supporting C. diabolis, apparently did not open to the surface until about 50,000 years ago (Winograd et al., 1988), and this almost certainly preceded isolation of C. diabolis from C. nevadensis.

If mtDNA diversity in ancestral C. nevadensis-C. diabolis was near that of existing populations of C. nevadensis, then C. diabolis would be much younger than the 0.5 Myr minimum indicated for its mtDNA. Both C. n. nevadensis and C. n. miniceps are polymorphic for shared mtDNA lineages that have been diverging from each other and from the mtDNA of C. diabolis for ~0.5–0.6 Myr. The lack of reciprocal monophyly probably reflects incomplete lineage sorting since isolation of C. diabolis. Such isolation might have occurred anytime from the last few thousand years to about 500,000 years ago, and the former seems likely given the geological history of Devils Hole.

Various mtDNA/allozyme conflicts in phylogenetic inferences (Fig. 5) suggest that the biogeographic history of Cyprinodon is more com-
plex than indicated by mtDNA alone. Such conflicts might reflect random lineage sorting from polymorphic ancestors, genetic introgression, selection, or poor resolution with one or the other dataset (Avise, 1997). Without further data, selection is not a useful hypothesis to explain the observed conflicts, leaving lineage sorting and introgression as potential evolutionary factors. As explained earlier, interpretations are compromised by relatively low confidence in phylogenetic resolution for allozyme variation in pupfish. Nonetheless, the major clades indicated by allozymes generally are consistent with the mtDNA tree, and most conflicts suggest hypotheses of secondary contact and genetic introgression that are geographically plausible.

The Rio Florido population of *C. eximius* clearly illustrates reticulate evolution. Past hybridization with *C. macrolepis*, a nearby, now-isolated spring-dweller, is indicated by geographic distributions of both allozymes (Echelle and Echelle, 1998) and mtDNA lineages (Fig. 5). With lineage sorting over time, the Rio Florido population should become fixed for one or the other of the divergent mtDNA lineages it now carries, a process occurring more rapidly for mtDNA than for nuclear genes because of the effectively haploid nature and maternal transmission of vertebrate mtDNA (Avise, 1994).

Such fixation following introgressive hybridization between divergent pupfishes might explain other mtDNA/allozyme conflicts indicated in Fig. 5. Three of the five remaining conflicts readily suggest reasonable hypotheses of genetic introgression. Two involve locally co-occurring species (*C. bifasciatus* and *C. aterius; C. pachycephalus* and *C. eximius*) in the only known situations of natural contact and hybridization between pupfish species in southwestern North America (Minckley, 1969; Minckley and Minckley, 1986). The third situation suggests contact and introgressive hybridization between *C. elegans*, which is endemic to isolated springs in the relatively old lower Pecos River, and a *C. pecosenis-C. bovinus* progenitor that was in the upper Pecos River prior to the proposed integration of the upper and lower basins.

The mtDNA/allozyme conflict in placement of *C. tularosa* is more difficult to explain. It potentially involved hybridization between a pupfish of the southern Great Plains-northern Chihuahuan Desert clade and a member of the maritime clade that, by at least Early Pleistocene, gained access to Tularosa Basin via the Rio Grande (Fig. 5). Additional studies of nuclear DNA will be needed to evaluate this hypothesis.

The previously mentioned mtDNA/allozyme conflict for the western pupfishes (clade 5, Fig. 5) might be explained by mtDNA-lineage sorting, either from an original ancestor in the Guzman Basin-lower Colorado River-Death Valley region, or subsequent to secondary contact and introgressive hybridization between divergent pupfishes (Echelle and Echelle, 1993a). Regarding the latter alternative, multiple instances of secondary contact would explain the curious lack of allozyme loci showing fixed differences among eight of the nine species, the exception being the peripherally distributed *C. radiosus* (Echelle and Echelle, 1993a). Otherwise, the only known examples of such a lack of fixed allozyme differences in *Cyprinodon* are the sympatric species in Lake Chichancanab (Humphries, 1984), the coastal species *C. variatus* and *C. artifrons* (Darling, 1976), and *C. macrolepis* and *C. eximius*, which, as mentioned earlier, were in contact relatively recently.

**Comparisons with other organisms.**—The allopatty of the primary mtDNA clades of *Cyprinodon* might reflect major vicariant events for the aquatic biota of southwestern North America. This predicts congruence in timing of cladogenetic events in pupfishes with those of other aquatic organisms in the region.

Mitochondrial DNA surveys of two fishes, *Cyprinella lutrensis* and *Fundulus zebrinus*, geographically overlap the southern Great Plains-northern Chihuahuan Desert mtDNA-clade for pupfishes. The cyprinid mirrored the marked divergence (uncorrected = 6.2%) between the Red and Brazos river forms of *C. rubrofuscatilis*, with 8.4% divergence (restriction-fragment-length analysis) between a clade in the Brazos River and another in more northern populations (Richardson and Gold, 1995). However, the fundulid phylogeography seems more recent, with a shallow clade (~2% divergence) of populations from the Red, Brazos, and Pecos rivers (Kriesser, 2001; Kriesser et al., 2001).

Relationships among the basal mtDNA clades in *Cyprinodon* are unresolved, but the areas occupied by those in the ancestral Rio Grande System correspond well to those for four of Mayden’s (1989) morphologically recognized clades of the *Cyprinella formosa-lepida* species complex: Guzmán Basin, Gulf Coastal Plain, Rio Conchos-Middle Rio Grande, and Cuatro Ciénegas-Lower Rio Grande Basin. The results for pupfish and the *C. formosa-lepida* group show little congruence with phylogenetic inferences for other fishes in the ancestral Rio Grande System, perhaps because of differing tendencies to occupy upland streams where dispersal or transport across basin divides is more likely. Pupfishes and
the C. formosa-lepida group primarily occupy lowlands and valley floors, whereas other groups examined phylogenetically, including the cyprinid genus Dionda (Mayden et al., 1992) and western catostomids (Smith, 1992a), are more common in upland streams. The results for Dionda do, however, parallel those for pupfish by being consistent with recognition of an Old Rio Nazas-Lower Rio Grande system and by supporting a sister relationship between an ancestral form in that system and species in the Rio Conchos-Middle Rio Grande.

Hydrobid snails often are restricted to isolated spring systems in southwestern North America and their phylogenetic pattern should be especially informative regarding patterns of vicariance. Studies by Hershler et al. (1999a, 1999b, 2002) and Hurt (2004) indicate a more ancient history than that of Cyprinodon, but both groups support (1) Early Pliocene connections between Guzmán Basin, the Colorado River, and the Death Valley System, (2) similar timing for connections between eastward flowing southern Great Plains streams and the Pecos River (Cyprinodon, 3.6 Mya; hydrobids, 4–6 Mya), and (3) the contact ~2 Mya between the Tularosa Basin and the Rio Grande. Further comparisons await additional studies of hydrobids, particularly those on the Mexican Plateau.

Concluding comments.—The mtDNA tree for Cyprinodon seems to preserve signals of vicariance in the Late Neogene biogeographic history of the aquatic fauna of southwestern North America. It appears, however, that mtDNA alone portrays an incomplete picture of the historical biogeography of Cyprinodon. The conflicting phylogenetic inferences from mtDNA and allozymes indicate that reticulate evolution involving divergent lineages were important in the genetic history of the group, and this probably is not unusual, at least for teleost fishes. Hybridization is common among and within extant genera of fishes (Hubbs, 1955; Schwartz, 1972), and both the fossil record (Smith, 1992b) and biochemical analyses (Dowling and Hoeh, 1991; DeMarais et al., 1992) indicate that introgression was not uncommon in the past. The mtDNA/allozyme conflicts in phylogenetic inferences suggest explicit hypotheses of vicariance and secondary contact and introgression that can be tested with additional nuclear markers for Cyprinodon, phylogenetic analyses of other organisms in the region, and geological studies of past surface-water connections among basins.

Collection localities for sequences generated in this study. Parentheses show locality numbers from Fig. 1, followed by number of specimens surveyed for variation (RFLP or sequencing) and number sequenced. Brackets = catalog numbers for vouchers at Oklahoma State University Collection of Vertebrates (OSUS): Cyprinodon rubrofuscus (1:6/2) Prairie Dog Town Fork of Red River, 20 km N Childress, Childress Co., Texas; C. pecosensis (3:6/2) Salt Creek, at Highway 285 bridge, Reeves Co., Texas; C. elegans (4:6/2) Giffin Spring canal at Toyahvale, Reeves Co., Texas [18246], (4b/2/2) captive stock at Uvalde National Fish Hatchery derived from Phantom Lake Spring, Jeff Davis Co., Texas; C. bovinus (5:6/2) Dexter National Fish Hatchery stock derived from Diamond Y Draw, Pecos Co., Texas; C. tularosa (6/2) Lost River at Holloman Air Force Base, Otero Co., New Mexico; C. eximus (7:2/2) Devils River at Paffords Crossing, Val Verde Co., Texas, (8:1/1) Alamito Creek, 5 km SE Presidio, Presidio Co., Texas, (9:2/2) Rio Florido, 1 km SE Villa Lopez, Chihuahua [18240]; C. macrolepis (10:2/2) Ojo de HaciaíDa Dolores, 11.2 km S Jimenez, Chihuahua [18228]; C. pecosensis (11:2/2) Ojo de San Diego, 57 km SE Chihuahua, Chihuahua [18230]; C. fontinalis (12:2/2) Ojo de Carbonera, 35 km W Villa Ahumada, Chihuahua [18238]; C. fisteri (13:2/2) tributary of Rio Casas Grandes, 14.4 km W Janos, Chihuahua [18237]; C. albivelis (14:2/2) springs ENE Ojo de Arrey, 4.8 km SSE Guela, Chihuahua [18242]; C. nazas (15:2/2) outflow of Ojo de la Concha, 9 km W Peñon Blanco, Durango [18235], (16:2/2) tributary of Laguna Santiaguito, Rio Guatimape at Guatimape, Durango [18244]; C. meeki (17:2/2) tributary of Rio del Tunal, 40 km NE Durango, Durango [18224]; C. sp. AguanaVal (18:2/2) Río AguanaVal at Ranco Grande, Zacatecas [18232]; C. alvarezi (19:2/2) irrigation canal from Ojo del Potosí at Ejido Catarino Rodriguez, Nuevo Leon [18243]; C. veronicae (20:2/2) Charco Azul, 1 km S San Juan de Avilés, Nuevo Leon [18223]; C. atrius (21:1/1) Laguna Grande at inlet from Río Churince, 15.5 km SE Cuatro Ciénegas, Coahuila [18231]; C. bifasciatus (22a:1/1) Pozo Churince [18227] and (22b:1/1) Pozo Escobeda, respectively, 16 km SE, 13 km S Cuatro Ciénegas, Coahuila; C. variegatus variegatus (23:12/2) Edinburg Water Supply Canal, 1 km N Edinburg, Hidalgo Co., Texas [18317], (24:12/1) isolated pool near Highway 90 bridge over Biloxi River,
Biloxi, Harrison County, Mississippi, (25:12/1) roadside pool near Yanketown, Levy Co., Florida, (26:12/2) Vero Beach at Riverside Park, Indian River Co., Florida, (27:12/1) Flagler Beach at Highway 201, Flagler Co., Florida; *C. variegatus ovinus* (28:12/1) S end of Chincoteaque Island, Accomac Co., Virginia; *C. bondi* (29:1/1) Lago del Fondo, Dominican Republic; *C. sp. “Enriquillo pupfish”* (30:1/1) Lake Enriquillo, Dominican Republic; *C. dearborni* (31:2/2) Carlisle Salt Works at Canas, Island of Bonaire; *C. artifrons* (32:1/1) tidepool at Río Lagartos, Yucatan, (33:2/2) tidepool at Cancun, Quintana Roo; *C. radiatus* (34:1/1) Warm Spring Refugium, S of Bishop, Inyo Co., California [18309]; *C. neodensii* (35:1/1) tailwaters of Shoshone Spring in Amargosa River at State Highway 178 SE of Shoshone, Inyo Co., California; *C. maculatus* (36:1/1) captive population at Boyce Thompson Arboretum, Pinal Co., Arizona, originally from Santa Clara Slough, Sonora; *Megapielos aporus* (37:1/1) captive stock from outdoor tanks at Universidad de Autonoma de Nuevo Leon, Monterrey, Nuevo Leon; *Culiceps tesselatus* (38:1/1) La Media Luna, 11 km S Rio Verde, San Luis Potosí [18225]; *Jordanella floriae* (39:1/1) 40 km NW West Palm Beach, Palm Beach Co., Florida [18314].

**Acknowledgments**

We dedicate this paper to the memories of R. Miller and W. Minckley. We thank S. Contreras and L. Lozano for help with fieldwork and many fruitful discussions; S. Cather, D. Love, R. Myers, and especially J. Hawley, for consultation on geology; R. Hershler for comments on drafts of the manuscript; T. Bunt, R. Edwards, L. Echelle, T. Echelle, H. Fitch, L. Fuselier, M. Smith, B. Turner, and A. Valdez for help in obtaining specimens; M. Childs, S. Hofer, L. Knowles, T. Malloy, M. Hainey, G. Wilson, and R. Pfau for help in the laboratory; T. Hrbek for providing his dataset on *Aphanopus*, S. Norris for unpublished manuscripts, S. Fox for help with the Spanish abstract, D. Swofford and M. Vences for analytical advice, and the Government of Mexico (SEDUE and Departamento de Pesca) for permission to collect in 1989. The research protocol satisfied requirements of the Oklahoma State University Animal Care and Use Committee. Collecting in Cuatro Ciénegas was done with permit DAH19990626-1-5 to D. Hendrickson from the government of Mexico. Funding provided by the National Science Foundation (AAE and AFE, RVDB, TED, and AM), the U.S. Fish and Wildlife Service (AAE and AFE), Fond der Chemischen Industrie and Deutsche Forschungsgemeinschaft (to AM), and the NASA Astrobiology Program (to TED).

**Literature Cited**


---

**REFERENCES**


---


---


**CONANT, R. 1963.** Semiaquatic snakes of the genus *Thamnophis* from the isolated drainage system of


Villwock, W. 1982. Aphanius (Nardo, 1827) and...


