Biological Journal of the Linnean Society, 2019, **127**, 742–761. With 9 figures.

Freshwater-to-marine transitions may explain the evolution of herbivory in the subgenus *Mollienesia* (genus *Poecilia*, mollies and guppies)

JESSICA L. SANCHEZ*, HEATHER D. BRACKEN-GRISSOM and JOEL C. TREXLER

Department of Biological Sciences, Florida International University, Miami, FL, USA

Received 25 October 2018; revised 7 March 2019; accepted for publication 9 March 2019

The ability of organisms to cross ecosystem boundaries is an important catalyst of evolutionary diversification. The genus *Poecilia* (mollies and guppies) is an excellent system for studying ecosystem transitions because species display a range of salinity and dietary preferences, with herbivory concentrated in the subgenus *Mollienesia*. We reconstructed ancestral habitats and diets across a phylogeny of the genus *Poecilia*, evaluated diversification rates and used phylogenetically independent contrasts to determine whether diet evolved in response to habitat transition in this group. The results suggest that ancestors of subgenus *Mollienesia* were exclusively herbivorous, whereas ancestral diets of other *Poecilia* included animals. We found that transitions across euryhaline boundaries occurred at least once in this group, probably after the divergence of the subgenus *Mollienesia*. Furthermore, increased salinity affiliation explained 24% of the decrease in animals in the gut, and jaw morphology was associated with the percentage of animals in the gut, but not with the percentage of species occupying saline habitats. These findings suggest that in the genus *Poecilia*, herbivory evolved in association with transitions from fresh to euryhaline habitats, and jaw morphology evolved in response to the appearance of herbivory. These results provide a rare example of increased diet diversification associated with the transition from freshwater to euryhaline habitats.

ADDITIONAL KEYWORDS: adaptive evolution – ancestral state reconstruction – diet evolution – freshwater habitat – habitat transition – herbivory – marine habitat – *Mollienesia* – phylogeny – *Poecilia*.

INTRODUCTION

The ability of organisms to cross habitat and ecosystem boundaries and invade new space is an important driver of evolutionary diversification. Habitat shifts by organisms may provide new foraging opportunities with little competition and decreased predation threats (Betancur-R *et al.*, 2012). In addition, invading a new habitat can have significant evolutionary consequences for the invading species by enhancing the possibility for novel phenotypes to evolve. These novel phenotypes can promote new ecological interactions between species, ultimately resulting in adaptive radiation (Lee, 1999; Betancur-R *et al.*, 2012; Davis *et al.*, 2012). However, the ability of an organism to transition across an ecological boundary requires a suite of adaptations suited for the new environment (Vermeij & Dudley, 2008; Betancur-R, 2009). Although such adaptations can be energetically costly to maintain, many metazoans are derived from ancestors that have crossed ecosystem boundaries (e.g. Vermeij & Dudley, 2008; Davis *et al.*, 2012; Mitterboeck et al., 2016), suggesting that the relative costs of transitioning can be outweighed by the ecological opportunities afforded to those with the ability to do so.

In aquatic systems, the interface between marine and freshwater habitats represents a boundary that creates a physiological challenge for potential invaders (Lee, 1999). As a result, colonization of marine habitats by freshwater organisms, or reinvasion of freshwater by secondary marine clades, is uncommon (McDowall, 1997; Vermeij, 2000; Betancur-R, 2009). Furthermore, approximately half of marine animal phyla have not colonized freshwater habitats (Betancur-R, 2009). However, several clades have successfully crossed aquatic ecosystem boundaries and have experienced rapid diversification in the freshwater clades relative to their marine counterparts (Davis *et al.*, 2012; Bloom *et al.*, 2013). For example, fish from the family

*Corresponding author. E-mail: jsanc318@fiu.edu

Terapontidae originated in marine habitats, but after a single marine-to-freshwater transition, 40 out of 54 extant species are restricted to freshwaters (Davis *et al.*, 2012). After their incursion from marine waters, freshwater terapontids diversified three times as quickly as the marine clade, accompanied by a shift from a carnivorous diet in marine habitats to an herbivorous diet in freshwater (Davis *et al.*, 2012). This diet diversification is likely to have occurred because freshwater systems have greater habitat complexity than marine systems (Strathmann, 1990; May, 1994). However, herbivory is thought to be a nutritionally inefficient feeding strategy relative to omnivory and carnivory (for a review, see Sanchez & Trexler, 2016); therefore, it is unclear why a habitat transition would prompt the evolution of a nutritionally 'inferior' diet.

It has been suggested that herbivory evolved as an adaptive strategy that allowed organisms to persist in habitats with decreased resource quality (i.e. 'suboptimal habitat hypothesis'; Sanchez & Trexler, 2016). Moving into 'suboptimal' habitats might allow organisms to minimize interspecific competition or to escape the negative effects of predation (Sanchez & Trexler, 2016). Given that marine systems are generally considered less productive per unit area than freshwater aquatic habitats (e.g. Colinvaux, 1980; May & Godfrey, 1994; Vermeij & Grosberg, 2010), they could be considered 'suboptimal' under the suboptimal habitat hypothesis. Therefore, the evolution of herbivory could also benefit organisms that make freshwater-to-marine transitions, not only those that transition in the opposite direction (e.g. terapontids).

The genus *Poecilia* is an excellent model system for studying transitions across ecosystem boundaries, because it consists of species with limited ranges and species with large, overlapping distributions with strong capacities for dispersal (Palacios *et al.*, 2016). Although all *Poecilia* species have some capacity to survive in both fresh and euryhaline waters, species with limited dispersal capacities tend to thrive in freshwater habitats (e.g. *Poecilia reticulata*), whereas others thrive in brackish and/or marine habitats (e.g. *Poecilia vivipara*, *Poecilia latipinna* and *Poecilia mexicana*; Meffe & Snelson, 1989). In addition, all *Poecilia* species exhibit some degree of herbivory; however, we hypothesize that obligate herbivory is concentrated in the subgenus *Mollienesia* (Sanchez, pers. obs.). As such, transitions from freshwater to less productive marine waters might have prompted the evolution of the herbivorous strategy in the genus *Poecilia*, particularly in the subgenus *Mollienesia*.

Our objective for this study is to reconstruct ancestral states of habitat and diet across a phylogeny of the genus *Poecilia* to identify patterns of diet evolution and habitat transition from freshwater to euryhaline (marine and/or brackish) systems (or vice versa) in the subgenus *Mollienesia*. This information will allow us to evaluate the suboptimal habitat hypothesis by determining whether habitat affiliations explain patterns of diet evolution throughout the phylogeny.

MATERIAL AND METHODS

Taxon sampling

There are 44 documented species in the genus *Poecilia*, spread across seven subgenera (Poeser, 2002; Ho *et al.*, 2016): *Acanthophacelus*, *Poecilia* (subgenus), *Micropoecilia*, *Curtipenis*, *Psychropoecilia*, *Allopoecilia* and *Mollienesia*. In this study, we assembled a dataset of 36 *Poecilia* species with at least one representative from all seven of the described subgenera, with two species from the sister genus *Limia*, to construct an updated topology. We chose *P. reticulata* as an outgroup taxon. Although this species is in the genus *Poecilia*, it has been shown to be a reliable outgroup taxon in previous studies focusing on the subgenus *Mollienesia* (e.g. Ptacek & Breden, 1998) and on the genus *Poecilia* (Alda *et al.*, 2013; Ho *et al.*, 2016). To date, our sampling represents the highest number of representative species collected across all *Poecilia* subgenera in a single study.

We collected diet and habitat data (see methodology below) from a subsample of our collection, represented by 15 *Poecilia* species spread across our six sampled subgenera (excluding *Curtipenis*). These were: *P. butleri*, *P. orri*, *P. mexicana*, *P. sphenops*, *P. gilli*, *P. caucana*, *P. hispaniolana*, *P. dominicensis*, *P. vivipara*, *P. latipinna*, *P. kyesis*, *P. velifera*, *P. picta*, *P. parae* and *P. reticulata* (Table 1). Of these, eight were representatives of the *Mollienesia* subgenus (*P. butleri*, *P. orri*, *P. mexicana*, *P. sphenops*, *P. gilli*, *P. latipinna*, *P. kyesis* and *P. velifera*) and represent individuals from the three recognized *Mollienesia* complexes (*P. mexicana*, *P. latipinna* and *P. sphenops*) listed by Ho *et al.*, (2016).

Phylogenetic analyses

Previous *Poecilia* phylogenies were constructed using several mitochondrial genes and one ribosomal gene (Alda *et al.*, 2013; Ho *et al.*, 2016; Palacios *et al.*, 2016): 5′ region of the cytochrome oxidase subunit I (*COI*; mtDNA), *ATPase 8/6* (mtDNA), NADH dehydrogenase subunit 2 (*ND2*; mtDNA) and the nuclear S7-like ribosomal protein (S7). The previous topologies did not include all available *Poecilia* species sequences and lacked a few of our subsampled species (*P. velifera*, *P. dominicensis*, *P. parae* and *P. picta*). To compare diet and habitat characteristics, it was necessary to create an updated tree that included the species represented in our entire dataset ($N = 36$). We retrieved sequences

Table 1. Complete list of sampled *Poecilia* specimens for gut and jaw morphology analyses

	Sample ID	Species	Locality description	State, country	Latitude	Longitude	Gut content sample size	Jaw measurement sample size
1.	UF 7333*	*P. sphenops*	Kilometre marker 583 between Lerdo de Tejada and Santiago Tuxtlas	Veracruz, Mexico	18.5869100°N	95.3650980°W	25	25
2.	UF 87585*	*P. sphenops*	Aguan River, on road CA 13, 44.6 miles W of Trujillo	Colon, Honduras	15.5281790°N	86.2305890°W	10	25
3.	UF 15249*	*P. butleri*	Rio Quelite, 22.6 miles NNW of Mazatlan	Sinaloa, Mexico	23.5226570°N	106.4978210°W	–	4
4.	UF 15253*	*P. butleri*	Mangrove swamp, 1.7 miles SE and 4.5 miles SW of Tecoman	Colima, Mexico	18.8703980°N	103.9322370°W	5	13
5.	UF 19554*	*P. gilli*	Quepos, stream near Los Junta de Alregados, at Pan American Highway bridge	Colima, Mexico	9.4515450°N	84.1680030°W	5	13
6.	UF 19567*	*P. gilli*	Rio Corobici and canal tributary, at La Pacifica Hotel, 5 km NW of Las Canas, near Pan American Highway	Guanacaste, Costa Rica	10.4721250°N	85.1226740°W	–	7
7.	UF 23988*	*P. dominicensis*	River 14 km NW of Sabina Grande de Boya	San Cristobal, Hispaniola	19.0092590°N	69.9094420°W	15	25
8.	UF 25044*	*P. dominicensis*	Rio Maimon, 7 km SW of Piedra Blanca, 250 m elevation	La Vega, Hispaniola	18.9022540°N	70.2830910°W	15	25
9.	UF 25049*	*P. hispaniolana*	Rio Yaque del Sur 9 km SW of Jarabacoa	La Vega, Hispaniola	19.0780560°N	–70.7186420°W	25	25
10.	UF 111695*	*P. hispaniolana*	Lago Enriquillo, 4 km from Descubierta	Hispaniola	18.5150000°N	–71.6608330°W	15	25
11.	UF 74903*	*P. picta*	Salybia River Bridge #3 1/2, E of 1.25 mile post between Salybia Bay and Galera Point	Trinidad, Trinidad and Tobagao	10.8339450°N	60.9206520°W	–	25
12.	UF 112133*	*P. vivipara*	Tenesopolis Municipality; Guarani farm	Rio de Janeiro, Brazil	19.9000000°S	55.8000000°W	15	25
13.	UF 188017*	*P. vivipara*	Itapicuru River off BA-381 between Filadélfia and Itiúba	Bahia, Brazil	–10.7041944°S	-39.8965278°W	15	25
14.	UMMZ 55052*	*P. caucana*	Small pools in course of small stream, Rio Camarones, at Arroyo de Arena	Columbia	11.2624590°N	72.9197800°W	–	25
15.	UMMZ 186930*	*P. caucana*	Rio Portillo, tributary called Rio Carache	Venezuela	9.61482222°N	70.54972222°W	15	25
16.	UMMZ 233640*	*P. parae*	Rio Maguari near Maguary, Belem	Para, Brazil	1.2818030°S	48.4274700°W	–	11

Table 1. Contined

	Sample ID	Species	Locality description	State, country	Latitude	Longitude	Gut content sample size	Jaw measurement sample size
17.	UMMZ 247482*	*P. parae*	Canals at Anna Regina on Essiquibo coast	Guyana, Brazil	7.2596680°N	58.4848630°W	–	19
18.	UF 24504*	*P. orri*	Below dam of reservoir on Salt Creek	Isla de Providencia, Columbia	13.3435810°N	81.3877640°W	25	24
19	ECOSUR donation 1	*P. orri*	Laguna Ubero	Quintana Roo, Mexico	19.0530250°N	–87.5739000°W	10	10
20.	ECOSUR donation 2	*P. mexicana*	Close to Carreterra El Cafetal-Mahahual	Quintana Roo, Mexico	18.96838333°N	–87.9472611°W	20	20
21.	POEMEX A	*P. mexicana*	Arroyo Escondido	Quintana Roo, Mexico	18.6111111°N	–88.8122222°W	4	4
22.	ECOSUR donation 3	*P. kykesis*	Champton	Campeche, Mexico	19.2652972°N	–87.5739583°W	15	15
23.	ECOSUR donation 4	*P. kykesis*	Arroyo Nuevo Loria	Quintana Roo, Mexico	19.3011111°N	88.5347222°W	10	10
24.	POEVEL A	*P. velifera*	Homochen	Yucatan, Mexico	21.2001510°N	–089.9484400°E	25	25
25.	POEVEL B	*P. velifera*	Ojo de Agua Ex Granja Pecis	Yucatan, Mexico	21.1834400°N	–089.9791300°E	21	25
26.	POELAT A	*P. latipinna*	Water Conservation Area 3B, boatramp near S-333 water structure	Florida, USA	25.7623722°N	–80.6731833°W	16	19
27.	POELAT B	*P. latipinna*	Mangrove area on the right of South-bound US 1, Everglades National Park	Florida, USA	25.2361583°N	80.4336722°W	20	20
28.	POERET A	*P. reticulata*	Tacarigua River via Caura Royal Road	Trinidad, Trinidad and Tobagao	10.6789333°N	–61.3194666°W	24	23
29.	POERET B	*P. reticulata*	Quare River	Trinidad, Trinidad and Tobagao	10.6000000°N	–61.1000000°W	22	20

*Museum samples obtained from the *Fishnet2* data base (http://www.fishnet2.net/).

(36 *Poecilia* species + two *Limia* species) for the same suite of genes used in previous works, because they provided strongly supported phylogenetic relationships at both the genus (e.g. Alda *et al.*, 2013; Ho *et al.*, 2016) and subgenus (e.g. Palacios *et al.*, 2016) level. These sequences were obtained from data deposited in Dryad by the previous authors (Alda *et al.*, 2013; Ho *et al.*, 2016) and were supplemented with additional sequences not included in these previous works using GenBank (for accession numbers and sample IDs, see Supporting Information, Table S1). We assembled the sequences using MEGA 7 (Kumar *et al.*, 2015). Pseudogenes were investigated by: (1) translating nucleotides to amino acids; (2) examining the sequences for stop codons; and (3) searching for insertions/deletions (mitochondrial and ribosomal genes). The sequences were aligned using the Muscle option in MEGA 7 and concatenated (*COI* + *ATPase 8/6* + *ND2* + S7) using Sequence Matrix (Vaidya *et al.*, 2011). We removed the first base of the *COI* sequences to set them in reading frame 1 (651 bp) and split the *ATPase 8/6* sequences into the partial *ATPase 8* (158 bp) segment and complete *ATPase 6* (684 bp) sequence. We used PartitionFinder v.2.1.1 (Lanfear *et al.*, 2012) to identify the best partitioning scheme and models of evolution that fitted the data. We used the Bayesian information criterion (BIC) to evaluate the best-fitting scheme and model with the greedy search algorithm, linked branch lengths and models restricted to those that can be used in MrBayes. We repeated these methods to obtain the best-fitting scheme for a second dataset composed of the subsampled sequences (15 *Poecilia* species). All replicate sequences were included in the pruned tree except *P. mexicana*, *P. sphenops* and *P. reticulata*. For these species, we included only individuals that were sampled in the same country as the specimens we used to collect dietary data.

We used MrBayes v.3.2 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003) to create a Bayesian inference (BI) phylogeny using the partitions and models specified in PartitionFinder for the concatenated datasets (all sequences and subsampled sequences). We constructed an analysis with uninformed priors, which ran for 1×10^6 generations, on four Markov chains. Trees were sampled every 100 generations. We performed three separate runs, each with two replicate runs. Following methods of Ho *et al.* (2016), we evaluated convergence of parameters using Tracer v.1.6 (Rambaut *et al.*, 2014) for each replicate and combined run and found that all values for effective sample size were > 200. Pairwise convergence of resulting tree topologies was evaluated using the RWTY package (Warren *et al.*, 2017) in R v.3.4.1 (R Core Team, 2017), using a 25% burn-in. In addition, we visually verified that the 50% majority-rule consensus trees for the three separate runs had matching topologies with minor deviations in branch lengths. We constructed a consensus tree for each posterior sample of trees using the *sumt* function in MrBayes and visualized the topologies using FigTree v.1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

Diversification analyses

We used the MEDUSA method ('Geiger' package in R; Alfaro *et al.*, 2009) to estimate diversification rate (r) shifts within the history of the genus *Poecilia*. This method applies a stepwise procedure to a time-calibrated phylogeny and species richness matrix (assigned to each tip based on taxonomic diversity) by fitting a birth–death model using a likelihood function (Rabosky *et al.*, 2007). We used the CHRONOS function ('ape' package in R; Paradis *et al.*, 2004) and *Poecilia* spp. divergence estimations taken from Palacios *et al.* (2016) to produce a time-calibrated tree (containing only subsampled species), and we created a species richness matrix using the complete list of described *Poecilia* species listed by Ho *et al.* (2016: table 1). Then, we used the MEDUSA function to fit a series of increasingly complex models to the tree to reveal the internal node rate shifts that give the highest likelihood. Models were compared using Akaike's information criterion (AIC) by calculating ΔAICc (ΔAICc = AIC_i – min AICc, where i = model I; Anderson & Burnham, 2002).

Habitat characterization

Poecilia species can survive in both fresh and salt waters and therefore show marked intraspecific variation in habitats they occupy. However, the rate of occurrence of individual species in fresh, brackish and marine habitats varies among species, revealing subtle differences in species-specific habitat preferences (Meffe & Snelson, 1989). We used the *Fishnet2* database to estimate interspecific habitat preferences. For each of our subsampled species, we performed a *Fishnet2* search using the species name. Given that these searches returned thousands of results (many of which were duplicates), we collected habitat information on the first 25 independent hits with logged latitude/longitude coordinates. This sample size was based on preliminary power analyses on species habitat types ($\alpha = 0.05$) for commonly studied species (e.g. *P. latipinna*, *P. reticulata*, *P. mexicana*), but for species that are not as heavily sampled, we compiled habitat data from the samples available (see Supporting Information, Table S2). Using the field collection notes provided by *Fishnet2*, Google Earth searches and accompanying geographical information, we determined whether each sample was collected from a freshwater, brackish or marine site. We then calculated

the proportion of samples collected from each habitat type for all species (Supporting Information, Table S2). We verified our predicted habitat associations with data reported in the literature for well-studied species (e.g. *P. reticulata*, *P. mexicana*, *P. latipinna*; Trexler & Travis, 1990; Nordlie *et al.*, 1992; Bussing, 1998; Miller, 2005), but these classifications are approximate and do not take into account seasonal or climatic changes in salinity, migration/dispersal events to or from different habitat types, or effective population sizes at each site. We assumed that if a species was able to be collected at a site, it had established there. We used an agglomerative hierarchical clustering procedure using the Sorensen (Bray–Curtis) distance measure with flexible beta linkage ($\beta = 0.25$) to classify habitat types into categorical variables for use in ancestral state reconstructions (CLUSTER package in R; Maechler *et al.*, 2017). We plotted the results using dendrograms, which were subjectively pruned, and the resulting groups were tested for validity using non-parametric multi-response permutation procedures (MRPPs; McCune & Grace, 2002; methods of Davis *et al.*, 2012).

Diet characterization

Subsampled species were obtained from Florida Museum of Natural History (retrieved from the *Fishnet2* database, http://www.fishnet2.net/), University of Michigan Museum of Zoology (*Fishnet2*), El Colegio de la Frontera Sur (ECOSUR) Ichthyology Collection (donations made to the authors) and collections made by the authors (Table 1). We used the most recent naming convention for *P. kykesis*; therefore, our *Fishnet2* search was performed using the former species name, *P. petenensis* (Poeser, 2002).

Adult individuals of each species were sampled from two distinct populations (i.e. no gene flow likely) within their native range using methods that do not interfere with diet characterization (e.g. by seining or cast nets, but not minnow traps), and were fixed in formalin and preserved in 70% ethanol after capture. We tried to capture intraspecific diet variation in by obtaining specimens collected from both the classified habitat type (euryhaline or freshwater) and a contrasting habitat type (e.g. sampling a freshwater population of a primarily euryhaline species), but our stringent sampling criteria limited our ability to do so for all species. Specifically, we were unable to collect dietary information on euryhaline populations of *P. vivipara*, *P. kykesis*, *P. sphenops*, *P. gilli*, *P. mexicana* and *P. velifera*; however, previous diet studies on these species corroborate our diet classifications (see Zaret & Rand, 1971; Winemiller, 1993; Bizerril & Primo, 2001; Plath *et al.*, 2005; Sa-nguansil, 2009) and suggest that interspecific variation is even greater than intraspecific variation in diet.

An analysis of several poeciliid species found that jaw morphologies varied among genera with different dietary habits, with more herbivorous species displaying a larger degree of intramandibular bending (IMB), larger gape angles (GAs) and a large degree of neurocranial rotation (NCR) (Gibb *et al.*, 2008; Hernandez *et al.*, 2008, 2009). We measured these jaw angles to the nearest 0.01 mm standard length and placed them under a dissecting scope with an attached digital camera. Using ImageJ software, we measured the vertex of a line along the ventral margin of the dentary bone that forms the lower jaw and a second line along the ventral margin of the angular–articular bone complex. We then subtracted the measured angle from 180° to obtain the degree of IMB. For GA, we measured the vertex of a line along the anterior-ventral margin of the upper jaw and a line along the anterior-dorsal margin of the lower jaw. Finally, we measured NCR by measuring the angle between a vertical line posterior to the eye and a line along the top of the skull above the eye (modified methods of Gibb *et al.*, 2008).

After jaw measurements, we assessed gut contents and morphology for each of the subsampled species. We were unable to dissect any specimens of *P. parae*, or specimens of *P. butleri* from a second locality owing to museum limitations; therefore, only jaw measurements were obtained for these individuals. We dissected all other fish to remove the gut tract. Once the tract was removed, we weighed it to the nearest 0.001 g, stretched it out onto a Petri dish lined with grid paper (6.35 mm grid) and recorded the length. To standardize the length for comparison among species, we divided the length of the gut (in millimetres) by standard body length (in millimetres). We removed a subsample from each gut (from the oesophagus to the first bend of the gut tract) and weighed it to the nearest 0.0001 g. We extracted the contents of the subsample onto a tared microscope slide using the blunt end of a razor blade. We then added a drop of deionized water to each slide, mounted them with a coverslip and sealed them using clear nail polish.

We examined slides using a light microscope at ×40 magnification and counted and identified all organisms (to genus) in ten random fields of view (counted area = 2.37 mm) to obtain the number of organisms per millilitre of gut material. We grouped the organisms found in the guts by trophic group (diatoms, green algae, cyanobacteria, metazoans) and calculated the relative abundance of each group for each fish species at both sampled localities to obtain the number of organisms per millilitre of gut material. Although we did not quantify detritus in the gut, we believe that detritus and detrital components (e.g. heterotrophic bacteria) are not the dietary target of these species, but instead supplement an algae-based diet in benthic ecosystems (Sanchez & Trexler, 2018). Therefore, we assumed that that detritus marginally

contributed to the diet of *Poecilia* fishes. We used an agglomerative hierarchical clustering procedure using the Sorensen (Bray–Curtis) distance measure with flexible beta linkage (β = 0.25) to classify gut contents into categorical variables. Similar to habitat clusters, diet groupings were tested using MRPP (McCune & Grace, 2002; methods of Davis *et al.*, 2012). All individuals of the same species (collected from different localities) clustered together, suggesting that intraspecific variation in gut contents was less than interspecific variation in gut contents. As such, we performed the clustering procedure again using the average gut content values for each species.

We used the morphological data (IMB, GA, NCR and standardized gut length) and gut content estimations to determine whether these diet characters are potential adaptations for the herbivorous diet in *Poecilia* species. For simplicity, we converted gut content data into the percentage of animal material in the gut. We then generated phylogenetically independent contrasts (PICs) between the percentage of animal material in the gut and each morphological character with the ape package in R (Paradis *et al.*, 2004) using branch lengths from our pruned topology (containing only subsampled species). Contrasts were used in linear-regression analyses, where the regression was forced through the origin (Felsenstein, 1985). Although our interspecific comparisons were relatively small ($K = 15$), Lajeunesse & Fox (2015) concluded that phylogenetic generalized least squares models are likely to conclude significant intercept and slope values irrespective of the number of species (K). Any characters that were significantly correlated ($P < 0.05$) with the percentage of animal material in the gut were assumed to have evolved in response to an herbivorous diet and were used as characters in ancestral state reconstruction.

Tracing the evolution of habitat and diet

We used ancestral state reconstruction to trace the dietary habits and habitat affiliations of ancestral *Poecilia* species. Given that we did not sample diet and habitat for every species belonging to each subgenus, we were restricted to interpretations of deep ancestral nodes of each species, rather than the most recent common ancestors (MRCAs) of entire subgenera. The exception was the subgenus *Mollienesia*, because we sampled multiple species across the three complexes, which is representative of the entire subgenus. Initially, we coded diet categories estimated from hierarchical cluster analysis as categorical traits (ranging from zero to five, and '?' for *P. parae*). Likewise, we coded the proportion of samples collected from each habitat type (estimated from *Fishnet2*) as categorical traits (ranging from zero to six). We created character matrices from these coded diet and habitat characters and from the morphological characters (IMB, GA, NCR and standardized gut length).

We uploaded our pruned consensus tree (subsampled species only) and character matrices into MESQUITE v.3.2 (Maddison & Maddison, 2017) and ran the 'trace character' analysis using maximum parsimony (MP) and maximum likelihood (ML) methods for habitat affiliation and diet category. We were able to run MP analyses only for jaw/gut morphology characters because these are continuous data and ML can analyse only categorical data. Parsimony ancestral state reconstruction minimizes the amount of character change over the tree topology based on the character state distribution and has thus been criticized for underestimating rates of evolutionary change (Cunningham *et al.*, 1998; Royer-Carenzi *et al.*, 2013). Maximum likelihood makes use of branch lengths and possible rates of character evolution to find the ancestral state that maximizes the probability that the observed character state (i.e. diet or habitat affiliation) would evolve under a stochastic model of evolution (Schluter *et al.*, 1997). In the present study, we used the symmetrical Mk1 model, which assumes equal forward and backward character transition rates (Lewis, 2001). Given that there has been some debate between using maximum parsimony (MP) and maximum likelihood (ML) methods, and because we were limited to more conservative MP methods for a subset of our data, we present the resulting reconstructions from both methods. The reconstructed states were plotted with the 'balls and sticks' model, with ancestral states marked at each node.

Identifying patterns of diet evolution in response to habitat transitions

We used phylogenetic independent contrasts (derived from our pruned tree) to compare diet and habitat affiliations across the genus *Poecilia*. Given that this method can be performed only on continuous data, we generated contrasts from the percentage of samples collected from euryhaline habitats (*Fishnet2* data) as a metric for habitat affiliation. We then used contrasts for habitat affiliation and all characters related to diet (percentage of animal material in the gut and our four measured morphological characters) in linear-regression analyses to identify the relationships between habitat affiliation, herbivory and the morphological adaptations related to herbivory.

RESULTS

Phylogenetic analyses

Full phylogeny (37 Poecilia *species)*

We partitioned the dataset by genes and by codons for the mtDNA (*COI*, *ATPase 8/6* and *ND2*) genes.

PartitionFinder identified the optimal partitioning scheme as four subsets of partitions (out of 13) for the complete *Poecilia* dataset (36 *Poecilia* species + two *Limia* species). Their estimated models of DNA substitution were as follows: (1) GTR+I+G for *COI* codon position 1, positions 2 and 3 of *ATPase 8*, *ATPase 6* and *ND2*; (2) K80+G for *COI* position 2 and complete S7; (3) F81 for *COI* codon position 3; and (4) HKY+G for position 1 of *ATPase 8/6* and *ND2*.

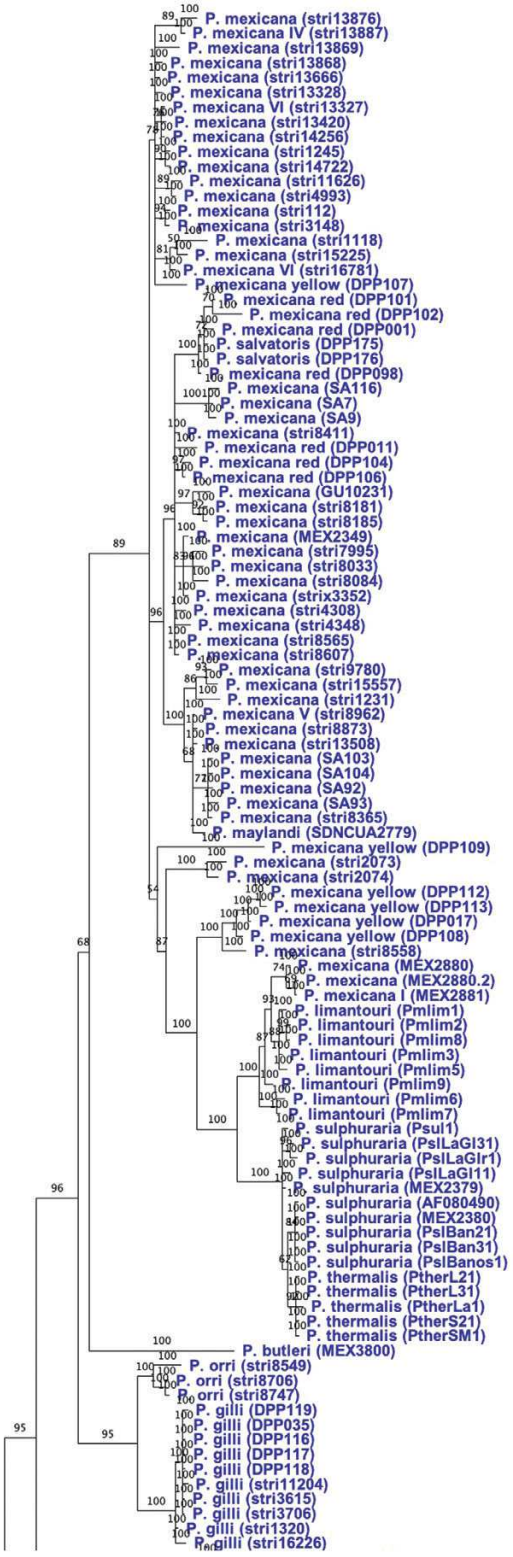
Our Bayesian phylogenetic analysis derived from the concatenated mitochondrial *COI*, *ATPase 8/6* and *ND2* and the ribosomal S7 genes from 36 *Poecilia* species (and two *Limia* species) resulted in a well-supported consensus tree, with the exception of the node linking the subgenera *Poecilia* and *Micropoecilia* [85% posterior probability (PP)]. Furthermore, these subgenera grouped together as an unresolved polytomy, which is not a supported pattern in previous studies (e.g. Palacios *et al.*, 2016). The low nodal support and polytomy are likely to have resulted from missing sequence data for individuals of the subgenus *Micropoecilia*, because only *ND2* sequences were available for these species (Fig. 1).

Although our analyses resulted in a tree with high support values, we found that *P. mexicana* species are not monophyletic as suggested by Ho *et al.*, (2016). Their topology placed *P. salvatoris* and several *P. mexicana* morphs (clades V–VI, yellow and red morphs) in a monophyletic group (Fig. 1). In our study, Bayesian analysis placed *P. salvatoris*, *P. maylandi*, *P. limantouri*, *P. sulphuraria* and *P. thermalis* with *P. mexicana* species, resulting in paraphyly.

Although monophyly was not supported, the position of these species within the *P. mexicana* complex is supported in our tree. The exception is *P. maylandi*, which is hypothesized to belong to the *P. sphenops* complex (Ho *et al.*, 2016). Given that no phylogenetic work has included *P. maylandi*, we are unable to conclude whether this species is in fact part of the *P. mexicana* complex instead of the *P. sphenops* complex, or if missing data and/or misidentification of the voucher specimen has resulted in the incorrect assignment of this species. Furthermore, *P. wandae* (sequences obtained from Ho *et al.*, 2016) was included in the subgenus *Mollienesia*, although this species has been classified as belonging to the subgenus *Allopoecilia*. Correspondence with Ho *et al.*, (2016) suggests that these vouchers were possibly misidentified and could be *P. koperi*, although this claim was never verified. All other deep nodes were highly supported (PP ≥ 90%) and congruent to those revealed in previous studies. Trees constructed from separate mitochondrial and ribosomal genes are available in the supplementary material (Figs S1, S2).

Subsampled phylogeny (15 Poecilia *species)*

Similar to the full phylogeny, we partitioned the dataset by genes and by codons for the mtDNA (*COI*,



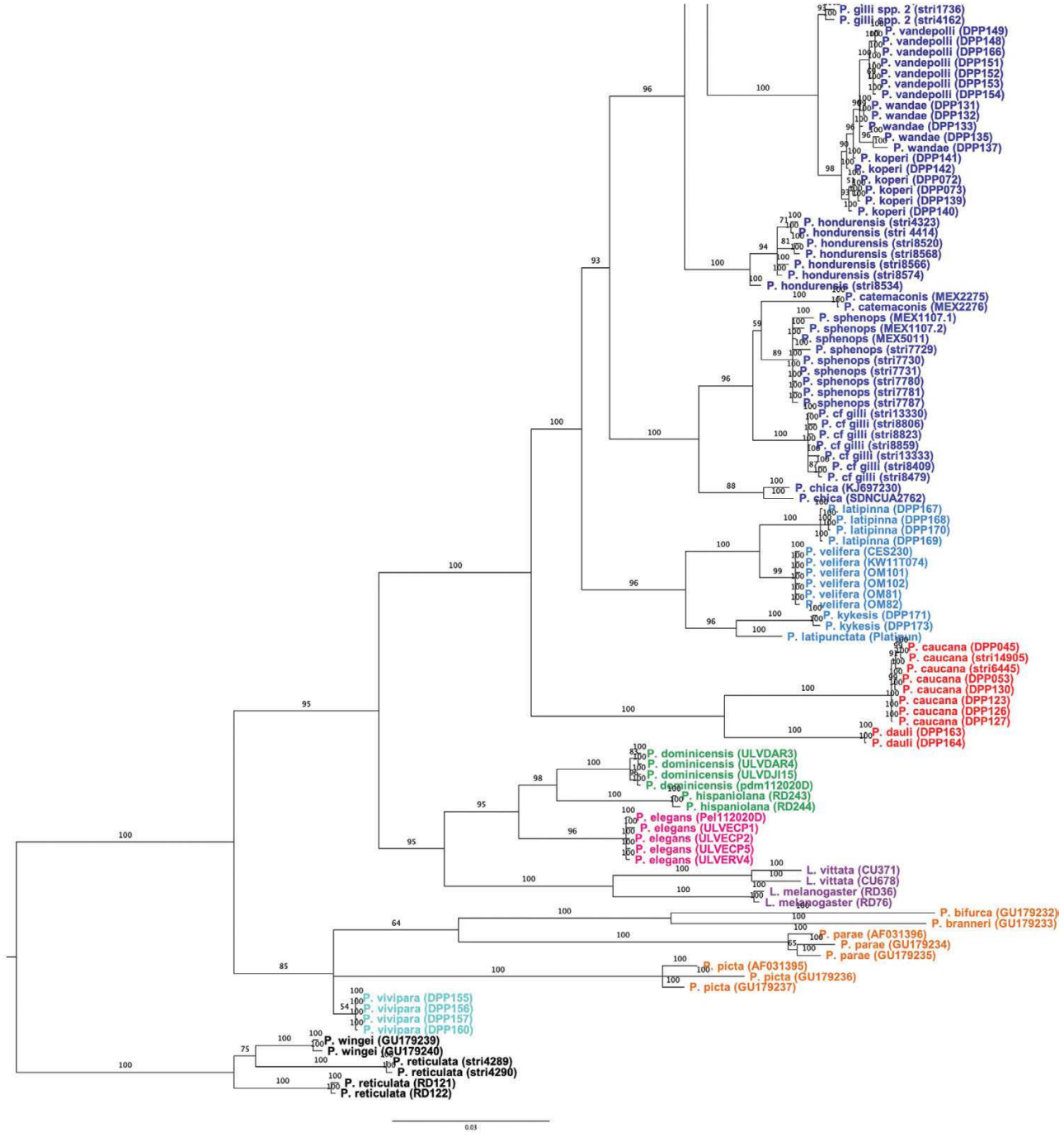


Figure 1. Bayesian phylogenetic tree (50% majority-rule) derived from concatenated mitochondrial cytochrome oxidase subunit I, *ATPase 8/6*, NADH dehydrogenase subunit 2 and ribosomal protein S7 genes for 36 *Poecilia* and two *Limia* species. Bullets at each node represent the posterior probability (PP). Nodes with posterior probabilities > 99% are considered highly supported, those with posterior probabilities > 95% are well supported, nodes with posterior probabilities > 75% are moderately supported, and those with posterior probabilities > 75% have no support. The GenBank ID for each species is listed in parentheses. Species are coloured by subgenus.

ATPase 8/6 and *ND2*) genes. PartitionFinder identified four subsets of partitions (out of 13) for the subsampled *Poecilia* dataset (15 species). Their corresponding models of evolution were as follows: (1) GTR+G for *COI* position 1 and position 3 of *ATPase 8/6* and *ND2*; (2) K80+G for position 2 of *COI* and *ATPase 8* and for

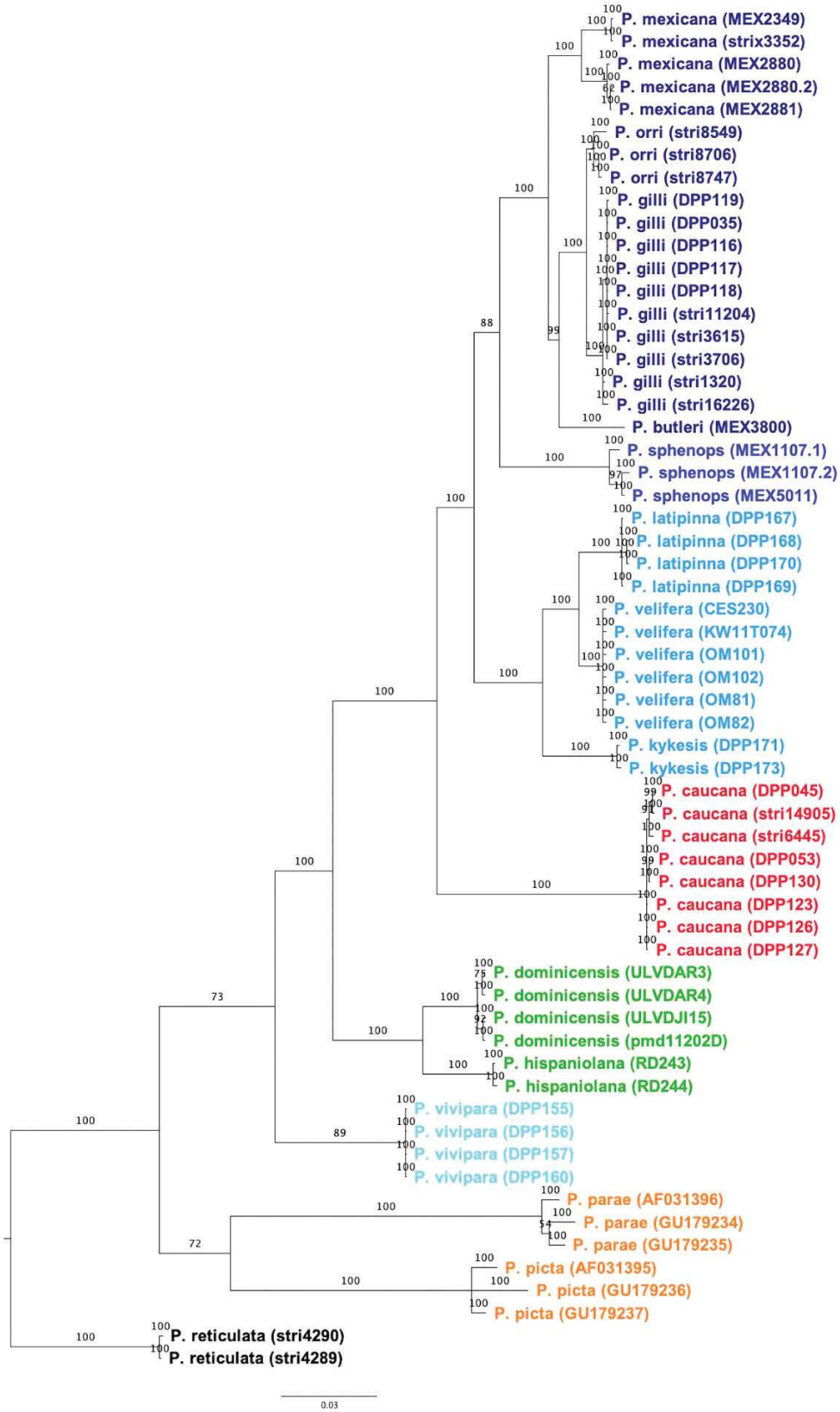


Figure 2. Pruned Bayesian phylogenetic tree (50% majority-rule) derived from concatenated mitochondrial cytochrome oxidase subunit I, *ATPase 8/6*, NADH dehydrogenase subunit 2 and ribosomal protein S7 genes for the 15 subsampled

complete S7; (3) HKY+I for *COI* codon position 3 and for position 2 of *ATPase 6* and *ND2*; and (4) HKY+G for codon position 1 of *ATPase 8/6* and of *ND2*.

The phylogenetic analysis of the subsampled *Poecilia* species resulted in a well-supported consensus tree, with few nodes of low support (Fig. 2). Specifically, the node linking species of the subgenus *Micropoecilia* (72% PP) and the node linking the subgenus *Poecilia* to the other subgenera (73% PP) had low support, probably as a result of missing sequence data (see previous subsection). However, unlike the full phylogeny, the pruned tree placed *P. vivipara* (subgenus *Poecilia*) in a different clade from *P. parae* and *P. picta* (subgenus *Micropoecilia*), a relationship that is congruent with previous studies (e.g. Palacios *et al.* 2016). Unlike the full phylogeny, we found that *P. mexicana* species formed a monophyletic clade with two subspecific groups (100% PP). The entire *P. mexicana* complex was composed of three subgroups: (1) *P. mexicana* species (including species listed above); (2) *P. orri* and *P. gilli*; and (3) *P. butleri*. This relationship and all others were highly supported (PP ≥ 90%) and congruent to those revealed in previous studies (Fig. 2). Pruned trees constructed from separate mitochondrial and ribosomal genes are available in the supplementary material (Figs S3, S4).

Diversification analyses

The net rate of diversification (r) of the genus *Poecilia* in a whole-tree birth model was 0.10, with a log-likelihood value of −86.71 (Table 2). The MEDUSA analysis found support for a pure-birth model with two shifts in diversification rate as the best-fitting model explaining the current diversity of the genus *Poecilia* (ΔAICc = 0.00; Table 2). The first shift occurred at the node containing the MRCA of subgenus *Mollienesia*, which showed a net diversification rate (r) that was 66% greater than the background rate of diversification occurring in other lineages (r = 0.29). The second shift occurred at the node linking *P. gilli* and *P. orri* (members of the *P. mexicana* complex, subgenus *Mollienesia*) and was 88% greater than the background diversification rate (r = 0.85; Fig. 3). Although the pure-birth model was the best fit, the birth–death model is also likely and also indicates a single shift in diversification rate (ΔAICc = 0.05; Table 2) at the node containing the MRCA of subgenus *Mollienesia*.

Habitat characterization

Our hierarchical cluster analysis produced six habitat categories (coded from zero to five in ancestral state reconstructions) that represented various salinity levels (Fig. 4A). All *Poecilia* species occupied freshwater habitats, but they occupied brackish and marine habitats at varying frequencies. Therefore, habitats were classified using the percentage of samples occupying euryhaline habitats. Of our subsampled species, *P. caucana*, *P. dominicensis*, *P. hispaniolana*, *P. reticulata* and *P. parae* were classified as having a low salinity affiliation (0% of samples collected from euryhaline habitats). Only 10–20% of *P. gilli*, *P. picta* and *P. vivipara* and 20–30% of *P. velifera* and *P. mexicana* were sampled in brackish or marine waters. Of the sampled *P. sphenops* and *P. butleri*, 30–35% were collected from euryhaline habitats. Approximately 35–40% of *P. latipinna* and *P. kykesis* and > 40% of *P. orri* samples were collected from euryhaline waters (Supporting Information, Table S2; Fig. 4A).

Diet characterization

We found differences in jaw and gut morphology among our subsampled species. Specifically, *P. reticulata* had the largest angles of neurocranial rotation, which were 75% more than the species with the smallest angles, *P. velifera* ($F_{15,587}$ = 23.314, P < 0.0001). Intramandibular bending was greatest in *P. mexicana*, where the degree of IMB was 13% greater than *P. reticulata*, the species with the smallest IMB angle ($F_{15,587}$ = 32.109, P < 0.0001). Gape angles showed a 53% difference between the species with the largest gape (*P. sphenops*) and the smallest gape (*P. picta*;

Table 2. MEDUSA models used to estimate diversification rates for clades in Figure 8

Model	No. of shifts	Clade	r	AICc	ΔAICc
Whole-tree birth–death	0	Whole tree	0.10	177.88	11.8
Birth–death	1	MRCA *Mollienesia*	0.29	166.13	0.05
Pure-birth	2	1. MRCA *Mollienesia* 2. Within *Poecilia mexicana* complex	0.84	166.08	0.00

Abbreviations: AICc, corrected Akaike information criterion; MRCA, most recent common ancestor; r, Net rate of diversification.

Poecilia species. Bullets at each node represent the posterior probability (PP). Nodes with posterior probabilities > 99% are considered highly supported, those with posterior probabilities > 95% are well supported, nodes with posterior probabilities > 75% are moderately supported, and those with posterior probabilities > 75% have no support. The GenBank ID for each species is listed in parentheses. Species are coloured by subgenus.

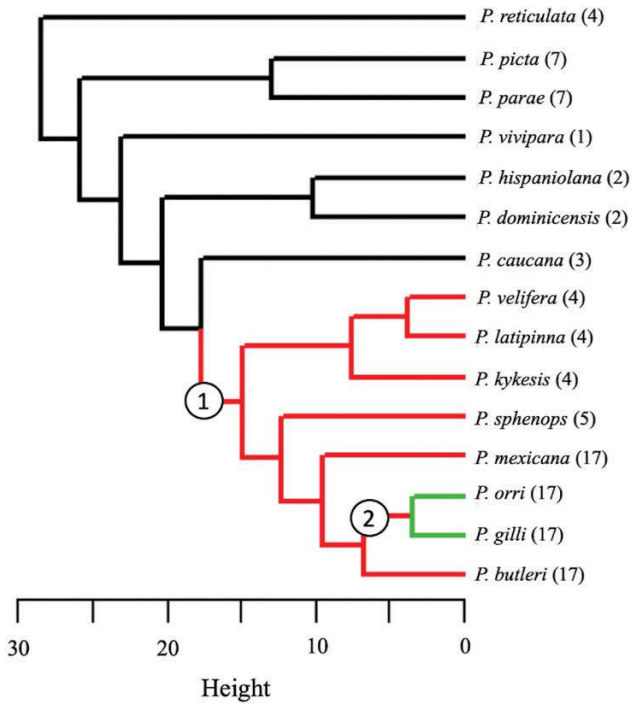


Figure 3. Diversification rate shifts across the *Poecilia* phylogeny. Tip richness values (for each subgenera) are listed in parentheses and were used to estimate lineage diversity. Node numbers represent changes in diversification rate (r) estimated using MEDUSA.

$F_{15,559}$ = 3.658, P < 0.0001). There were intraspecific differences in all three jaw measurements for *P. vivipara*, where the Rio de Janiero population had 38% greater neurocranial rotation and 24% greater gape angles (NCR, $F_{1,49}$ = 30.824, P < 0.0001; GA, $F_{1,49}$ = 13.325, P = 0.001), but the Bahia population had 9% greater IMB ($F_{1,49}$ = 6.105, P = 0.017). All other species did not differ in intraspecific jaw measurements. *Poecilia sphenops* had the longest standardized gut length, which was 43% longer than *P. reticulata*, our outgroup species ($F_{14,391}$ = 13.787, P < 0.0001; Supporting Information, Table S3).

Our hierarchical cluster analysis of gut content data produced six broad feeding categories (coded from zero to five in ancestral state reconstructions): carnivore (≥ 50% animals), three omnivore categories ('cyanobacteria + animals', 'diatoms + animals' and 'diatoms + cyanobacteria + animals') and two herbivore categories ('cyanobacteria' and 'diatoms + cyanobacteria'). Based on these groupings, *P. reticulata* (outgroup) were classified as carnivores, and *P. picta* ('cyanobacteria + animals'), *P. hispaniolana*, *P. caucana* ('diatoms + animals'), *P. dominicensis* and *P. vivpara* ('diatoms + cyanobacteria + animals') were classified as omnivores. All other *Poecilia* species were grouped as herbivores, where *P. sphenops*, *P. latipinna* and *P. gilli* guts contained diatoms and cyanobacteria, and all others contained cyanobacteria only (Fig. 4B). The relative abundance of each gut item can be found in the Supporting Information (Table S4).

Gape angles and the percentage of animal material in the gut were informative characters after correcting for phylogenetic relationships. Specifically, gape angles showed inverse relationships with the percentage of animal material in the gut, irrespective of phylogenetic relationship among species (y = −58.23x, r^2 = 0.27, P = 0.03). Intramandibular bending, neurocranial rotation angles and standardized gut lengths were not driven by the percentage of animal material in the diet once the phylogenetic relationships were accounted for (IMB, y = 4.36x, r^2 = −0.084, P = 0.796; NCR, y = 14.08x, r^2 = 0.089, P = 0.169; gut length, y = −94.35x, r^2 = 0.038, P = 0.250); therefore, these characters were not used in ancestral state reconstruction.

Tracing the evolution of habitat and diet

We used ancestral state reconstructions to estimate the habitat and diet of the MRCA of the subgenus *Mollienesia*, but we were limited to inferences on deep ancestral nodes of all other species. Habitat reconstructions varied between the methods used. Specifically, MP analyses suggest that the deeper ancestral nodes were represented by species that inhabited freshwater habitats, with the exception of subgenus *Mollienesia*, which inhabited euryhaline habitats (Fig. 5). However, the ML analyses revealed that the ancestral nodes of all species, including subgenus *Mollienesia*, represented inhabitants of fresh, brackish and marine waters (Fig. 6). Ancestral diet reconstructions using both methods suggested that the MRCA of the subgenus *Mollienesia* was exclusively herbivorous. The maximum parsimony analysis revealed that ancestral nodes of all other species were represented by either carnivorous or omnivorous species (Fig. 5), whereas ML analysis suggested that the ancestral node of *P. reticulata* (subgenus *Acanthophacelus*) was represented solely by a carnivore. The ancestral nodes of all other species (belonging to subgenera *Micropoecilia*, *Poecilia* and *Psychropoecilia*) were represented by omnivorous species (Fig. 6). Ancestral state reconstructions estimating jaw morphology revealed that GAs were increased in the MRCA of subgenus *Mollienesia* relative to the ancestral nodes of the other species (Fig. 7).

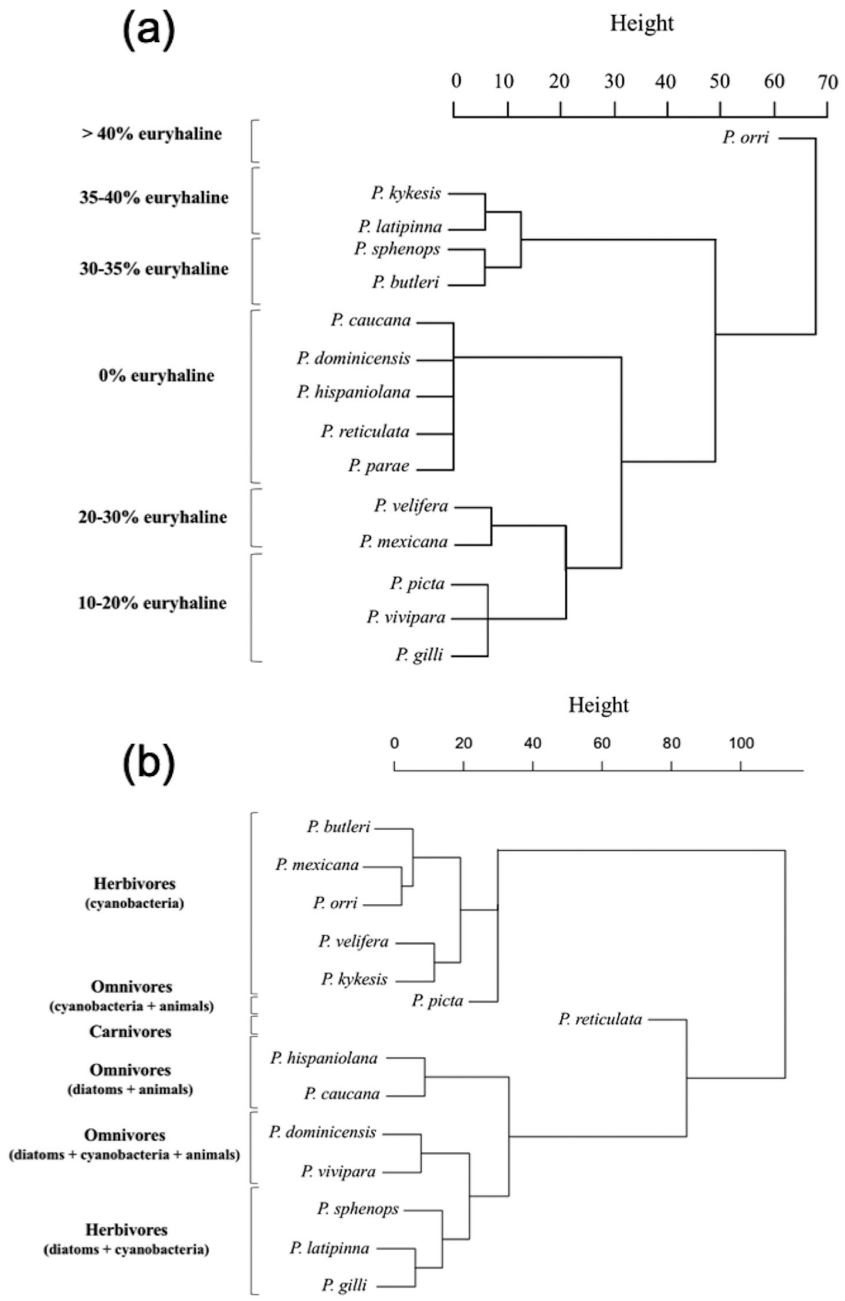


Figure 4. A, classification of *Poecilia* habitats using Sorensen (Bray–Curtis) distance measures with flexible beta linkage. Hierarchical cluster analysis identified seven habitat categories. B, classification of *Poecilia* diets using Sorensen (Bray–Curtis) distance measures with flexible beta linkage. Hierarchical cluster analysis identified six diet categories.

IDENTIFYING PATTERNS OF DIET EVOLUTION IN RESPONSE TO HABITAT TRANSITIONS

Phylogenetic independent contrasts on habitat affiliation (percentage of species occupying euryhaline habitats) and diet characters revealed contrasting patterns. Habitat affiliation did not predict GA ($y = 0.232x, r^2 = 0.033, P = 0.260$), despite the relationship between the percentage of animal material in the gut and GA. However, salinity affiliation explained 24% of the percentage of animal material in the gut ($y = -94.35x$, $r^2 = 0.24$, $P = 0.05$), suggesting that increased salinity affiliation might drive an increase in herbivory (decrease in animal material in the gut; Fig. 8).

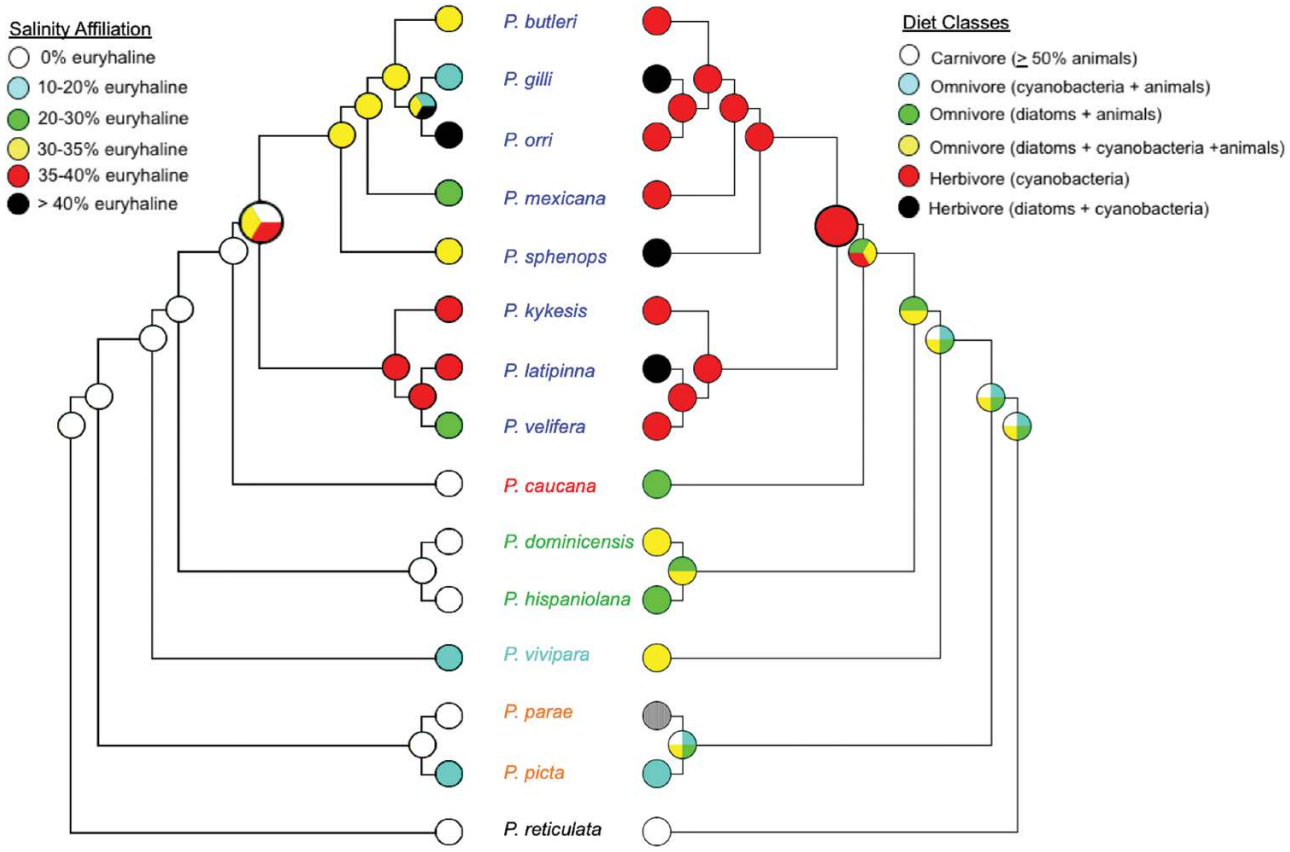


Figure 5. Maximum parsimony ancestral character reconstruction for the evolution of habitat (left cladogram) and diet (right cladogram) in the genus *Poecilia*. Circles at terminal nodes represent the observed character states for extant species, and pie charts for ancestral nodes show estimated probabilities for reconstructed character states. Species are coloured by subgenus, and the node representing the most recent common ancestor for subgenus *Mollienesia* is marked with a large circle. Black, *Acanthophacelus* (outgroup); blue, *Mollienesia*; green, *Psychropoecilia*; orange, *Micropoecilia*; red, *Allopoecilia*; teal, *Poecilia* (subgenus).

DISCUSSION

Our results revealed that herbivory might have evolved in response to invading less productive euryhaline habitats, thereby supporting the suboptimal habitat hypothesis (Sanchez & Trexler, 2016). We found that the MRCA of the exclusively herbivorous subgenera *Mollienesia* had euryhaline (both MP and ML analyses) roots and was also herbivorous. Furthermore, the ancestral node of our outgroup species, *P. reticulata*, probably contained a carnivorous (MP) or omnivorous (ML) species that inhabited fresh (MP) or euryhaline (ML) waters. All other *Poecilia* ancestors (deep nodes representing ancestral species of subgenera *Micropoecilia*, *Poecilia*, *Psychropoecilia* and *Allopoecilia*) inhabited fresh (MP) or euryhaline (ML) waters and were likely to be omnivorous (both MP and ML). Salinity affiliation (measured by the percentage of samples collected from brackish + marine habitats) explained 24% of the total variation in the diet of *Poecilia* species (measured by the percentage of animal material in the gut), and GAs were associated with the percentage of animal material in the gut, but not with the percentage of species occupying euryhaline habitats. These findings suggest that in this genus, herbivory evolved in response to habitat transitions between fresh and euryhaline habitats, and jaw morphology evolved in response to the appearance of herbivory.

Incorporating additional *Poecilia* species for phylogenetic analyses did not reveal any new relationships compared with previous studies, but instead verified the relationships among subgenera within the tree, allowing us to use these data for ancestral state reconstructions of diet and habitat. Dietary ancestral state reconstructions revealed that all species belonging to the subgenus *Mollienesia* displayed obligate herbivory (both MP and ML), whereas other *Poecilia* species were either carnivorous (MP) or omnivorous (both MP and ML). Two herbivorous strategies emerged ('cyanobacteria' and 'diatoms +

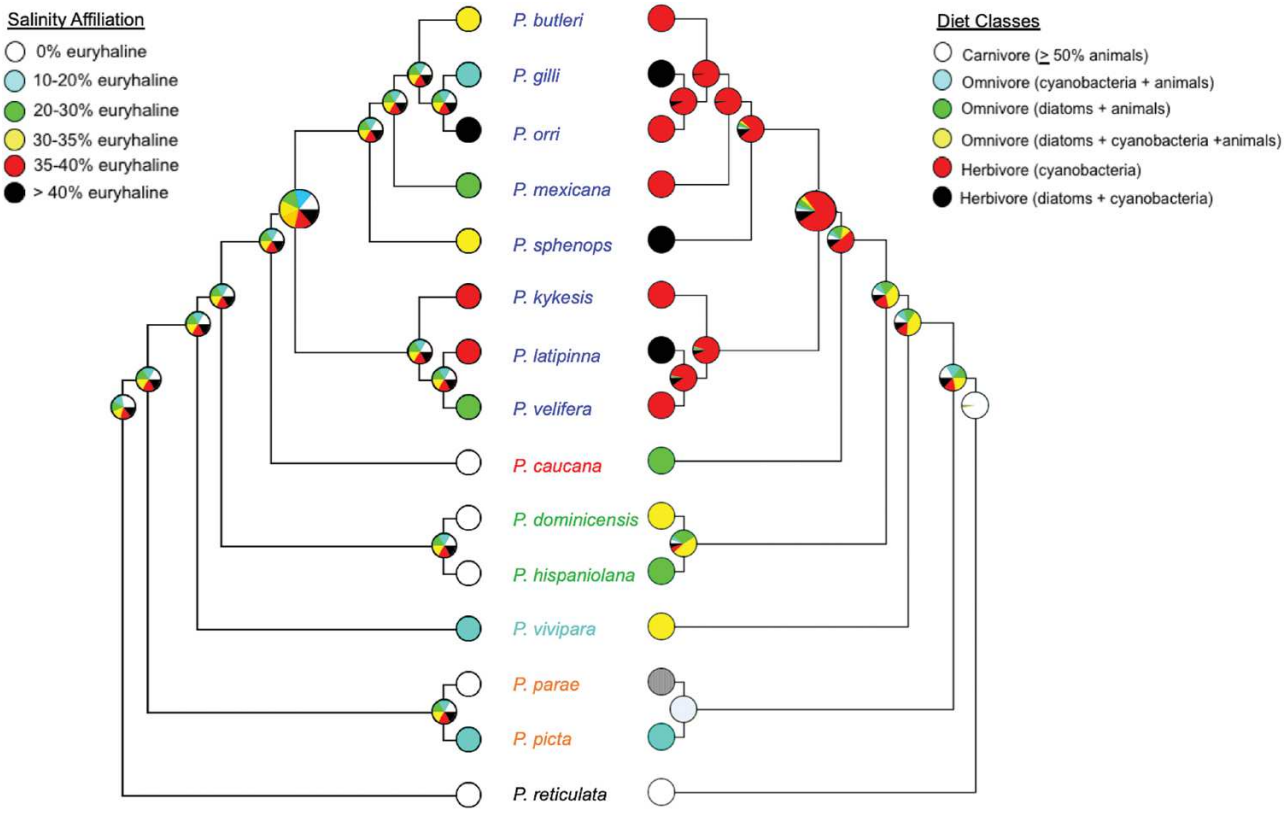


Figure 6. Maximum likelihood ancestral character reconstruction for the evolution of habitat (left cladogram) and diet (right cladogram) in the genus *Poecilia*. Circles at terminal nodes represent the observed character states for extant species, and pie charts for ancestral nodes show estimated probabilities for reconstructed character states. Species are coloured by subgenus, and the node representing the most recent common ancestor for subgenus *Mollienesia* is marked with a large circle. Black, *Acanthophacelus* (outgroup); blue, *Mollienesia*; green, *Psychropoecilia*; orange, *Micropoecilia*; red, *Allopoecilia*; teal, *Poecilia* (subgenus).

cyanobacteria') in the subgenus *Mollienesia*, and these correspond to the primary producer communities of tropical euryhaline habitats. Specifically, these primary producer communities are dominated by cyanobacteria (e.g. Flombaum *et al.*, 2013), which is reflected by the gut contents of the *Mollienesia* species sampled in the present study.

The results of our habitat ancestral reconstructions were dependent on the type of analysis performed (MP vs. ML), but phylogenetically independent contrasts allowed us to support these inferences better. Specifically, MP ancestral habitat reconstructions revealed a freshwater-to-euryhaline transition when the MRCA of subgenus *Mollienesia* diverged from the clade containing *P. caucana* (subgenus *Allopoecilia*). Alternatively, our ML model suggested that the entire genus probably originated in euryhaline habitats, with several euryhaline-to-freshwater transitions occurring before the divergence of the subgenus *Mollienesia*. Despite the uncertainty in our ancestral habitat estimations, we found that increased salinity affiliation explained 24% of the decrease in animal material in the gut. Our ancestral reconstructions suggested that the first appearance of obligate herbivory occurred in the MRCA of the subgenus *Mollienesia*, and our PICs indicated that increased salinity affiliation might have driven increased herbivory in this group. Taken together, these results might indicate that a freshwater-to-euryhaline transition occurred in the MRCA of this group (as predicted by the MP results).

Our diversification analyses support the hypothesis that salinity affiliation drove increased herbivory in the subgenus *Mollienesia*. More specifically, we found a 66% increase in diversification rate at the node containing the MRCA of the subgenus *Mollienesia*, which might suggest that a habitat transition prompted a shift to herbivory in this group. However, it is possible that salinity affiliation evolved before the divergence of the genus *Poecilia* (as predicted by ML). *Poecilia vivipara* and *P. picta* can also be

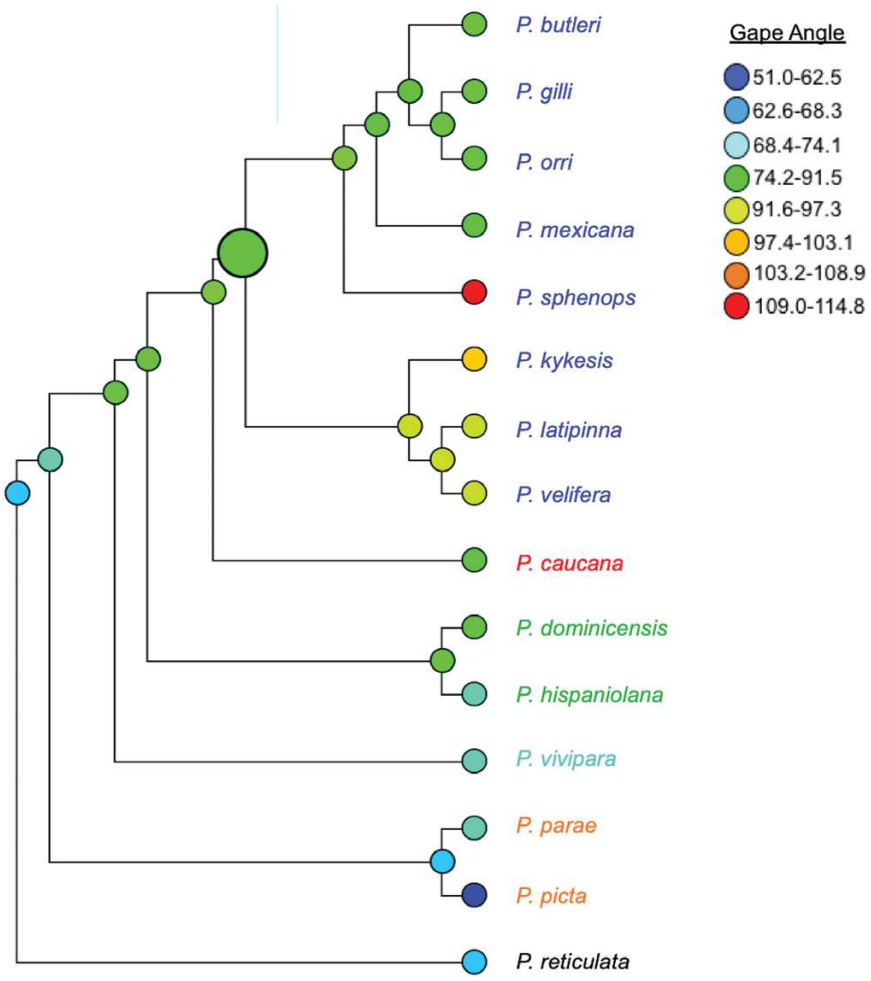


Figure 7. Maximum parsimony ancestral character reconstruction for the evolution of gape angles in the genus *Poecilia*. Circles at terminal nodes represent the observed character states for extant species, and pie charts for ancestral nodes show estimated probabilities for reconstructed character states. Species are coloured by subgenus, and the node representing the most recent common ancestor for subgenus *Mollienesia* is marked with a large circle. Black, *Acanthophacelus* (outgroup); blue, *Mollienesia*; green, *Psychropoecilia*; orange, *Micropoecilia*; red, *Allopoecilia*; teal, *Poecilia* (subgenus).

found in euryhaline habitats, and both these species diverged from the MRCA of the genus *Poecilia* ~3 Mya (Palacios *et al.*, 2016), suggesting that salinity affiliation evolved before the appearance of the subgenus *Mollienesia* 0.25 Mya (Palacios *et al.*, 2016). Furthermore, a salinity tolerance of up to 58 ppt has been documented for *P. reticulata* (Chervinski 1984), which diverged from the MRCA of the genus *Poecilia* 2.5 Mya (Palacios *et al.*, 2016). Our results suggest that salinity affiliation drove the evolution of herbivory in the subgenus *Mollienesia*; however, this finding does not explain why obligate herbivory failed to evolve in older lineages that also contain species with high salinity tolerances. In nature, carnivorous/omnivorous species, such as *P. reticulata*, do not typically occupy high-salinity habitats (Torres-Dowdall *et al.*, 2013), whereas *Mollienesia* species are found in habitats with a wide range of salinities (0–80 ppt; Nordlie *et al.*, 1992). Therefore, natural habitat preference might be a more informative metric than salinity tolerance when attempting to understand the mechanism of diet evolution in this group.

Freshwater-to-marine transitions are relatively rare in fishes (McDowall, 1997; Vermeij, 2000; Betancur-R, 2009), probably because of the decreased habitat complexity offered by marine habitats (Strathmann, 1990; May, 1994). In addition, herbivory is thought to be an energetically inferior diet compared with omnivory or carnivory; therefore, co-evolution of salinity affiliation and an herbivorous feeding strategy seems maladaptive when also experiencing a cost of osmoregulation. Our results support at least one transition across habitat boundaries in the genus *Poecilia*, but the direction (one fresh-to-euryhaline

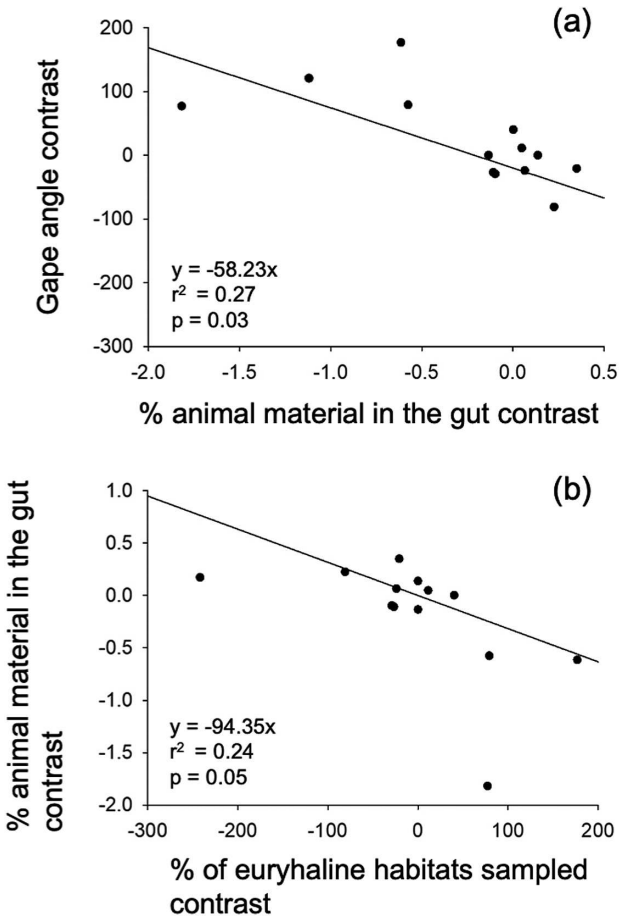


Figure 8. A, the relationship between the percentage of animal material in the gut and gape angle (plotted as phylogenetically independent contrasts) suggests that gape angle evolved as an adaptation to an increasingly herbivorous diet. B, the relationship between salinity affiliation and the percentage of animal material in the gut (plotted as phylogenetically independent contrasts) suggests that herbivory is an adaptation to euryhaline habitats.

transition vs. several euryhaline-to-fresh transitions) and timing (MRCA of subgenus *Mollienesia* vs. ancestral nodes of all other species) of the transition are unclear. However, we show that salinity affiliation might be related to rapid diversification favouring the evolution of herbivory in the subgenus *Mollienesia*, supporting a freshwater-to-euryhaline transition at the node containing the MRCA of the group.

The suboptimal habitat hypothesis posits that herbivory might be an adaptive strategy to allow organisms to persist in habitats with decreased resource quality, where animal prey are scarce and plant abundance is high (Sanchez & Trexler, 2016). Under this definition, a euryhaline habitat may be considered 'suboptimal' relative to a highly productive freshwater habitat. Therefore, our data partly support the suboptimal habitat hypothesis as an explanation for the appearance of herbivory in this group. It is important to note, however, that there might be other explanations supporting the evolution of herbivory in other metazoan groups (for alternative hypotheses, see Sanchez & Trexler, 2016) and that multiple mechanisms might be working simultaneously to explain the appearance and subsequent maintenance of herbivory in nature (see Sanchez & Trexler, 2018). Other studies have linked omnivore/herbivore richness to a decrease in latitude (proxy for temperature) in both freshwater (González-Bergonzoni *et al.*, 2012) and marine systems (Floeter *et al.*, 2005; González-Bergonzoni *et al.*, 2012). Furthermore, temperature, but not salinity, is positively correlated with the evolution of herbivory in fishes from the family Cleupeidae (Egan *et al.*, 2018). These results combined with the findings of the present study suggest that temperature might interact with salinity affiliation to promote the evolution of primary and secondary consumer diets in aquatic animals.

Our study suggests that obligate herbivory and, to some degree, brackish or marine affiliation are derived characters in the genus *Poecilia*. In addition, we show

that salinity affiliation partly drove the evolution of obligate herbivory. This result is surprising because there is ample evidence that freshwater-to-marine transitions generally result in decreased diversification relative to transitions in the opposite direction (e.g. McDowall, 1997; Vermeij, 2000; Betancur-R, 2009; Davis *et al.*, 2012). Although productive freshwater systems offer increased foraging opportunities compared with marine systems, we found that invading a 'suboptimal' habitat triggered diet diversification in the subgenus *Mollienesia*. The ability to cross ecosystem boundaries coupled with an adaptive diet strategy could allow *Poecilia* species to expand their range rapidly, thereby increasing opportunities for ecological diversification, ultimately resulting in species radiation.

ACKNOWLEDGEMENTS

We wish to thank David Reznick (University of California-Riverside), Juan Schmitter-Soto (ECOSUR) and Xavier Chiappa-Carrara National Autonomous University of Mexico-Sisal (UNAM-Sisal) for assistance with field collections, Donovan German (University of California-Irvine) for specimen collection/transportation advice, and Adeljean Ho (Chinese Academy of Sciences, South China Sea Institute of Oceanology) for sharing data that allowed us to construct our *Poecilia* phylogeny. In addition, we wish to thank the Florida Museum of Natural History and the University of Michigan Museum of Zoology for access to their ichthyology collections. This material was developed in collaboration with the Florida Coastal Everglades Long-Term Ecological Research program under National Science Foundation grant no. DEB-1237517. Animal use was approved by the Institutional Animal Care and Use Committee under protocol number IACUC-16-029. Funding for travel was provided by the Latin American and Caribbean Center at Florida International University (FIU) through two Tinker Field Research Grants awarded to J. Sanchez. This is contribution number 903 from the Southeast Environmental Research Center in the Institute of Water & Environment at Florida International University, and contribution number 125 from the Center for Coastal Oceans Research in the Institute for Water and Environment at Florida International University. Finally, we would like to thank the anonymous reviewers for providing valuable feedback towards the improvement of this manuscript.

REFERENCES

Alda F, **Reina RG**, **Doadrio I**, **Bermingham E. 2013.** Phylogeny and biogeography of *Poecilia sphenops* species complex (Actinopterygii, Poeciliidae) in Central America. *Molecular Phylogenetics and Evolution* **66:** 11011–1026.

Alfaro ME, **Santini F**, **Brock C**, **Alamillo H**, **Dornburg A**, **Rabosky DL**, **Carnevale G**, **Harmon LJ. 2009.** Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 13410–13414.

Anderson D, **Burnham K. 2002.** Avoiding pitfalls when using information-theoretic methods. *The Journal of Wildlife Management* **66:** 912–918.

Betancur-R R. 2009. Molecular phylogenetics support multiple evolutionary transitions from marine to freshwater habitats in ariid catfishes. *Molecular Phylogenetics and Evolution* **55:** 249–258.

Betancur-R R, **Orti G**, **Stein AM**, **Marceniuk AP**, **Pyron RA. 2012.** Apparent signal of competition limiting diversification after ecological transitions from marine to freshwater habitats. *Ecology Letters* **15:** 822–830.

Bizerril CRSF, **Primo PBS. 2001.** *Peixes de Águas Interiores do Estado do Rio de Janeiro*. Rio de Janeiro: FEMAR-SEMADS.

Bloom DD, **Weir JT**, **Piller KR**, **Lovejoy NR. 2013.** Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of new world silversides (*Atherinopsidae*). *Evolution* **67:** 2040–2057.

Bussing WA. 1998. *Peces de las aguas continentales de Costa Rica / Freshwater fishes of Costa Rica*. San José: Editorial de la Universidad de Costa Rica. 468 pp.

Chervinski J. 1984. Salinity tolerance of the guppy, *Poecilia reticulata* Peters. *Journal of Fish Biology* **24:** 449–452.

Colinvaux P. 1980. *Why big fierce animals are rare*. Princeton, NJ: Princeton University Press.

Cunningham CW, **Omland KE**, **Oakley TH. 1998.** Reconstructing ancestral character states: a critical reappraisal. *Trends in Ecology & Evolution* **13:** 361–366.

Davis AM, **Unmack PJ**, **Pusey BJ**, **Johnson JB**, **Pearson RG. 2012.** Marine freshwater transitions are associated with the evolution of dietary diversification in terapontid grunters (Teleostei: Terapontidae). *Journal of Evolutionary Biology* **25:** 1163–1179.

Egan JP, **Bloom DD**, **Kuo C**, **Hammer MP**, **Tongnunui P**, **Iglesias SP**, **Sheaves M**, **Grudpan C**, **Simons AM. 2018.** Phylogenetic analysis of trophic niche evolution reveals a latitudinal herbivory gradient in Clupeoidei (herrings, anchovies, and allies). *Molecular Phylogenetics and Evolution* **124:** 151–161.

Felsenstein J. 1985. Phylogenies and the comparative method. *The American Naturalist* **125:** 1–15.

Fishnet. **2013.** *A computerized database management system for the national fish collection at the J.L.B. Smith Institute of Ichthyology*. Available at: http://www.fishnet2.net/

Floeter S, **Behrens MD**, **Ferreira CEL**, **Paddack MJ** and **Horn M. 2005.** Geographical gradients of marine herbivorous fishes: Patterns and processes. *Marine Biology* **146(6):** 1435–1447.

Flombaum PJL, **Gallegos RA**, **Gordillo J**, **Rincon LL**, **Zabala N**, **Jiao DM**, **Karl WK**, **Li W**, **Lomas MW**, **Veneziano D**, **Vera CS**, **Vrugt JA**, **Martiny AC. 2013.** Present and future distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proceedings of the National Academy of Sciences of the United States of America* **110:** 9824–9829.

GenBank. *National Center for Biotechnology Information; [1982]–[2017]*. Bethesda: National Library of Medicine (US). Available at: https://www.ncbi.nlm.nih.gov/nucleotide/

Gibb AC, **Ferry-Graham LA**, **Hernandez LP**, **Romansco R**, **Blanton J. 2008.** Functional significance of intramandibular bending in Poeciliid fishes. *Environmental Biology of Fishes* **83:** 507–519.

González-Bergonzoni I, **Meerhoff M**, **Davidson TA**, **Teixeira-de Mello F**, **Baattrup-Pedersen A**, **Jeppesen E. 2012.** Meta-analysis shows a consistent and strong latitudinal pattern in fish omnivory across ecosystems. *Ecosystems* **15:** 492–503.

Hernandez LP, **Ferry-Graham LA**, **Gibb AC. 2008.** Morphology of a picky eater: a novel mechanism underlies premaxillary protrusion and retraction within cyprinodontiforms. *Zoology* **11:** 442–454.

Hernandez LP, **Gibb AC**, **Ferry-Graham LA. 2009.** Trophic apparatus in Cyprinodonitiform fishes: functional specializations for picking and scraping behaviors. *Journal of Morphology* **270:** 645–661.

Ho ALFC, **Pruett CL**, **Lin J. 2016.** Phylogeny and biogeography of *Poecilia* (Cyprinodontiformes: Poeciliinae) across Central and South America based on mitochondrial and nuclear DNA markers. *Molecular Phylogenetics and Evolution* **101:** 32–45.

Huelsenbeck JP, **Ronquist F. 2001.** MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17:** 754–755.

Kumar S, **Stecher G**, **Tamura K. 2015.** MEGA7: Molecular Evolutionary Genetics Analysis version 7.0. *Molecular Biology and Evolution* **33:** 1870–1874.

Lajeunesse MJ, **Fox GA. 2015.** Statistical approaches to the problem of phylogenetically correlated data. In: Fox GA, Negrete-Yankelevitch S, Sosa VJ, eds. *Ecological statistics: contemporary theory and application*. Oxford: Oxford University Press, 261–283.

Lanfear R, **Calcott B**, **Ho SYW**, **Guindon S. 2012.** PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* **29:** 1695–1701.

Lee CE. 1999. Rapid and repeated invasions of fresh water by the copepod *Eurytemora affinis*. *Evolution* **53:** 1423–1434.

Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* **50:** 913–925.

Maddison WP, **Maddison DR. 2017.** *Mesquite: a modular system for evolutionary analysis*, *Version 3.2*. Available at: http://mesquiteproject.org

Maechler M, **Rousseeuw P**, **Struyf A**, **Hubert M**, **Hornik K. 2017.** *cluster: cluster analysis basics and extensions. R package version 2.0.6.* Available at: https://cran.r-project.org/web/packages/cluster/cluster.pdf

May RM, **Godfrey J. 1994.** Biological diversity: differences between land and sea. *Philosophical Transactions of the Royal Society B: Biological Sciences* **343:** 105–111.

McCune B, **Grace JB. 2002.** *Analysis of ecological communities*. Glenden Beach: MjM Software Design.

McDowall RM. 1997. The evolution of diadromy in fishes (revisited) and its place in phylogenetic analysis. *Reviews in Fish Biology and Fisheries* **7:** 443–462.

Meffe GK, **Snelson EE. 1989.** An ecological overview of Poeciliid fishes. In: Meffe GK, Snelson EE, eds. *Ecology and evolution of livebearing fishes*. NJ: Prentice Hall, 13–32.

Miller RR. 2005. *Freshwater fishes of Mexico*. Chicago: University of Chicago Press.

Mitterboeak TF, **Chen AY**, **Zaheer OA**, **Ma YET**, **Adamowicz SJ. 2016.** Do saline taxa evolve faster? Comparing relative rates of molecular evolution between freshwater and marine eukaryotes. *Evolution; international journal of organic evolution* **70:** 1960–1978.

Nordlie FG, **Haney DC**, **Walsh SJ. 1992.** Salinity tolerances and osmotic regulatory capabilities in populations of Sailfin Molly (*Poecilia latipinna*) from brackish and fresh waters. *Copeia* **1992:** 741–746.

Palacios M, **Voelker G**, **Rodriguez LA**, **Mateos M**, **Tobler M. 2016.** Phylogenetic analyses of the subgenus *Mollienesia* (*Poecilia*, Poeciliidae, Teleostei) reveal taxonomic inconsistencies, cryptic biodiversity, and spatio-temporal aspects of diversification in Middle America. *Molecular Phylogenetics and Evolution* **103:** 230–244.

Paradis E, **Claude J**, **Strimmer K. 2004.** APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20:** 289–290.

Plath M, **Heubel KU**, **Garcia de León FJ**, **Schlupp I. 2005.** Cave molly females (*Poecilia mexicana*, Poeciliidae, Teleostei) like well-fed males. *Behavioral Ecology and Sociobiology* **58:** 144–151.

Poeser FN. 2002. *Poecilia kykesis* nom. nov., a new name for *Mollienesia petenensis* Gunther, 1866, and redescription, revalidation and the designation of a lectotype for *Poecilia petenensis* Gunther, 1866 (Teleostei: Poeciliidae). *Contributions to Zoology* **704:** 243–246.

Ptacek MB, **Breden F. 1998.** Phylogenetic relationships among the mollies (PoeciliidaeL PoeciliaL Mollienesia group) based on mitochondrial DNA sequences. *Journal of Fish Biology* **53:** 64–81.

R Core Team. **2017.** *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at: https://www.R-project.org/

Rabosky DL, **Donnellan SC**, **Talaba AL**, **Lovette IJ. 2007.** Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. *Proceedings of the Royal Society B: Biological Sciences* **274:** 2915–2923.

Rambaut A, **Suchard MA**, **Xie D**, **Drummond AJ. 2014.** *Tracer v.1.6*. Available at: http://tree.bio.ed.ac.uk/software/tracer/

Ronquist F, **Huelsenbeck JP. 2003.** MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:** 1572–1574.

Royer-Carenzi M, **Pontarotti P**, **Didier G. 2013.** Choosing the best ancestral character state reconstruction method. *Mathematical Biosciences* **242:** 95–109.

Sa-nguansil S. 2009. *Population characteristics and diet of Sailfin Molly Poecilia velifera (Regan, 1914) (Poeciliidae: Cyprinodonitformes): an alien fish invading Songkhla Lake Basin, South Thailand*. Masters Thesis, Prince of Songkla University.

Sanchez JL, **Trexler JC. 2016.** The adaptive evolution of herbivory in freshwater systems. *Ecosphere* **7:** e01414.

Sanchez JL, **Trexler JC. 2018.** When is an herbivore not an herbivore? Detritivory facilitates herbivory in a freshwater system. *Ecology and Evolution* **8:** 5977–5991.

Schluter D, **Price T**, **Mooers AO**, **Ludwig D. 1997.** Likelihood of ancestor states in adaptive radiation. *Evolution* **51:** 1699–1711.

Strathmann RR. 1990. Why life histories evolve differently in the sea. *American Zoologist* **30:** 197–207.

Torres-Dowdall J, **Dargent F**, **Handelsman CA**, **Ramnarine ID**, **Ghalambor CK. 2013.** Ecological correlates of the distribution limits of two poeciliid species along a salinity gradient. *Biological Journal of the Linnean Society* **108:** 790–805.

Trexler JC, **Travis J. 1990.** Phenotypic plasticity in the Sailfin Molly, *Poecilia latipinna* (Pisces: Poeciliidae). I. Field experiments. *Evolution; international journal of organic evolution* **44:** 143–156.

Vaidya G, **Lohman DJ**, **Meier R. 2011.** SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* **27**:171–180.

Vermeij GJ. 2000. Why are there so few evolutionary transitions between aquatic and terrestrial ecosystems? *Biological Journal of the Linnean Society* **70:** 541–554.

Vermeij GJ, **Dudley R. 2008.** Why are there so few evolutionary transitions between aquatic and terrestrial ecosystems? *Biological Journal of the Linnean Society* **70**(4): 541–554.

Vermeij GJ, **Grosberg RK. 2010.** The great divergence: when did diversity on land exceed that in the sea? *Integrative and Comparative Biology* **50:** 675–682.

Warren DL, **Geneva AJ**, **Lanfear R. 2017.** RWTY (R We There Yet): an R package for examining convergence of Bayesian phylogenetic analyses. *Molecular Biology and Evolution* **34:** 1016–1020.

Winemiller KO. 1993. Seasonality of reproduction by livebearing fishes in tropical rainforest streams. *Oecologia* **95:** 266–276.

Zaret TM, **Rand AS. 1971.** Competition in tropical stream fishes: support for the competitive exclusion principle. *Ecology* **52:** 336–342.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. GenBank accession numbers for genes used to reconstruct *Poecilia* phylogeny.
Table S2. Percentage of habitat types occupied by each species based on collections logged in the *Fishnet2* database (http://www.fishnet2.net/).
Table S3. Measured jaw angles of each sampled *Poecilia* species. Abbreviations: GA, gape angle; IMB, intramandibular bending (angle subtracted from 180°); NCR, neurocranial rotation.
Table S4. Relative abundance of diet items in the gut of each sampled *Poecilia* species.
Figure S1. Bayesian phylogenetic tree (50% majority-rule) derived from mitochondrial genes cytochrome oxidase subunit I, *ATPase 8/6* and NADH dehydrogenase subunit 2 from 36 *Poecilia* and two *Limia* species. The GenBank ID for each species is listed in parentheses. Species are coloured by subgenus.
Figure S2. Bayesian phylogenetic tree (50% majority-rule) derived from ribosomal gene S7 from 36 *Poecilia* and two *Limia* species. The GenBank ID for each species is listed in parentheses. Species are coloured by subgenus.
Figure S3. Pruned Bayesian phylogenetic tree (50% majority-rule) derived from mitochondrial genes cytochrome oxidase subunit I, *ATPase 8/6* and NADH dehydrogenase subunit 2 from 15 *Poecilia* species. The GenBank ID for each species is listed in parentheses. Species are coloured by subgenus.
Figure S4. Pruned Bayesian phylogenetic tree (50% majority-rule) derived from ribosomal gene S7 from 15 *Poecilia* species. The GenBank ID for each species is listed in parentheses. Species are coloured by subgenus.