Title

Introgressive hybridization among Gulf of California endemic species of the genus Colpichthys in relation to environmental change in the Colorado River Delta

Author

Lau, Long Fung

Publication Date

2017

UNIVERSITY OF CALIFORNIA

Los Angeles

Introgressive hybridization among Gulf of California endemic species of the genus *Colpichthys* in relation to environmental change in the Colorado River Delta

A thesis submitted in partial satisfaction

of the requirements for the degree Master of Science

in Biology

by

Long Fung Lau

2017

ABSTRACT OF THE THESIS

Introgressive hybridization among Gulf of California endemic species of the genus *Colpichthys* in relation to environmental change in the Colorado River Delta

by

Long Fung Lau

Master of Science in Biology
University of California, Los Angeles, 2017
Professor David K. Jacobs, Chair

The unique habitats of the Northern Gulf of California are home to many endemic vertebrate and invertebrate taxa. Owing to their limited geographic range, these taxa are often vulnerable to extinction. Located at the tip of the Northern Gulf is the Colorado River Delta. Upstream damming and water diversions in the past century have turned the Delta into an inverse estuary. The resulting rise in salinity and environmental impacts in the region have raised concerns for the local endangered fauna. Among these is the endemic silverside fish, *Colpichthys hubbsi*, which is narrowly restricted to the tidal channels within the Delta. Its sister species, *Colpichthys regis*, inhabits lagoons and estuarine channels on both sides of the Gulf. The two species share a region of syntopy at the southwestern edge of the Delta where hybridization is suspected to occur. Combining phenotypic and molecular data, this study shows evidence for introgressive hybridization between *C. regis* and *C. hubbsi*. While we find no evidence for recent population decline in *C. hubbsi*, the reduction of river outflow may lead to diminished ecological separation between the sister species, putting *C. hubbsi* at risk of extinction by introgression. Future work on this system will need to focus on identifying genes under divergent selection and assessing the full extent of genomic introgression using high-throughput sequencing approaches. Molecular analyses reveal a surprising lack of genetic diversity in the less geographically restricted species (*C. regis*). Results from neutrality tests based on mitochondrial DNA suggest that both species of *Colpichthys* may have undergone recent demographic expansion. Recent studies have demonstrated profound effects of Pleistocene climate fluctuations on population structure of estuarine fish, prompting the need for additional investigations to elucidate the complex demographic histories of these taxa.

The thesis of Long Fung Lau is approved.

Kirk Edward Lohmueller

Blaire Van Valkenburgh

David K. Jacobs, Committee Chair

University of California, Los Angeles

2017

Table of Contents

List of Tables

Supplementary Tables

List of Figures

Supplementary Figures

Acknowledgements

I am grateful for the years of guidance my advisor, David Jacobs, has given me. I would like to thank him for serving as my inspiration and for providing me with all the opportunities to learn from him. His continuous support, patience, and encouragements are deeply appreciated.

Besides my advisor, I would like to thank members of my thesis committee, Kirk Lohmueller and Blaire Van Valkenburgh, for their comments and help on revising this thesis.

I’ve been lucky to have Greer Dolby, Ryan Ellingson, and David Gold as my mentors. It was under Greer’s guidance that I started my path in research; her mentorship has played an essential role in my training as a scientist.

A big thanks to all my fellow labmates for all the intellectually stimulating discussions. I thank Paige Zhang and Hayden Speck for offering their insightful advice. I also thank Bruno Passerelli for assistance in data collection.

Additionally, I thank Rick Feeney of the Los Angeles County Natural History Museum for granting me access to the use of the x-ray machine and the specimens in the ichthyology collection. I also thank the Scripps Institute of Oceanography for providing the historical specimens.

Finally, I would like to express my gratitude to my parents, on whom I can always depend for support.

INTRODUCTION AND BACKGROUND

The Gulf of California has long been known for its productivity and biodiversity. The northern portion of the Gulf is a unique home to many endemic vertebrate and invertebrate taxa such as *Gillichthys detrusus*, *Totoaba macdonaldi*, *Uca monolifera*, and *Phocoena sinus* (Brusca, 1980; Swift et al., 2011; Palacios-Salgado et al., 2012; Silber, 1990). Owing to their limited geographic range, endemic species are especially vulnerable to extinction and therefore are of special interest in conservation biology. At the tip of the Northern Gulf of California is the mouth of the Colorado River; significant changes to the hydrology of the region have taken place since the completion of Hoover Dam in 1935. Upstream damming and water diversion for irrigation and municipal purposes have drastically reduced the flow of freshwater into the Northern Gulf; the ensuing desiccation of the Colorado River Delta transformed the former brackish wetlands into dry salt flats intersected by hypersaline tidal channels (Glenn et al., 1996). While the cessation of river output has raised much concern for the endangered species in the local ecosystem, the full extent of its environmental impacts has been a contentious issue. There have been debates on the extent to which the reduced Colorado River flow has impacted the local fauna. Recently, Brusca et al. (2017) has argued that throughout the Holocene, periodic natural water impoundment at the ancient Lake Cahuilla limited freshwater from entering the Gulf, and as a result, the Northern Gulf ecosystem has long been accustomed to the seasonal droughts of the Colorado River. The same study has argued that Northern Gulf marine productivity has not suffered from reduced Colorado River flow, and that overfishing and accidental bycatch should be the critical environmental concern. Nevertheless, the issue is far from settled as productivity is only one of many measures of ecosystem health, and local endangered species, especially those protected within the Biosphere Reserve (Reserva de la Biosfera del Alto Golfo de California y

Delta del Rio Colorado), could still be affected by the freshwater shutoff in other less predictable ways. In this study, we examine two Gulf of California endemic species of silverside fish in the genus *Colpichthys* that have been listed on the IUCN Red List of Threatened Species as Near Threatened and Endangered (Findley et al., 2010a; Findley et al., 2010b).

Colpichthys is a relatively obscure silverside belonging to the family Atherinopsidae that inhabits the near-shore habitats of the northern part of the Gulf of California. The two recognized species in this genus are *Colpichthys regis* (Jenkins and Evermann) and *Colpichthys hubbsi* (Crabtree). Often found in hypersaline coastal lagoons north of the latitude 26°N, historically and currently *C. regis* has a range that is disjointed across the tip of the Northern Gulf at the Colorado River Delta (hereinafter: Delta). Its sister species, *C. hubbsi*, is a geographically restricted species within the Delta region, and has been found to co-occur with *C. regis* only at the southwestern border of the delta where it meets the Baja coast (Fig. 1A). Despite significant genetic divergence (~6% cytochrome *b* sequence divergence; see results below) between the sister clades, we hypothesize that the region of syntopy could represent a hybrid zone, where introgressive hybridization may occur.

Introgressive hybridization is commonly defined as the transfer of alleles from one species to the gene pool of another species through repeated backcrossing of hybrids (Heiser, 1973; Harrison and Larson, 2014). In this study, we compare both the external and internal morphology of specimens recovered from *C. regis* and *C. hubbsi* populations to show that putative hybrid specimens at the area of syntopy have intermediate morphology in several key characters. We present clear instances of introgression of *C. regis* mitochondrial and nuclear haplotypes in various populations of *C. hubbsi*, some of which reside deep within the Delta where *C. regis* should be absent. We also discover evidence for admixture based on six microsatellite loci. All the available evidence points towards the presence of a hybrid zone at the southwestern border of the Delta on the Baja Coast. This hybridization may prove to have evolutionary significance, as introgression could allow for the spread of adaptive as well as deleterious alleles from one species to the other across the hybrid zone. Alternatively, introgression may promote the homogenization of the two species, ultimately dissolving the ecological boundaries necessary for maintaining species distinction.

In addition to investigating the hybridization, with the available genetic data we performed a number of population genetics analyses to infer genetic diversity and demographic histories of the two species. Our results show that populations of *C. regis* lack the genetic diversity one would expect for a species with its distribution within the Gulf of California. Taken together, our reported results indicate a complicated history in the Gulf, and additional investigations are clearly warranted.

MATERIALS AND METHODS

Sample Collection

Specimens of the genus *Colpichthys* were collected by seine from estuarine habitats along the coast of the Northern Gulf of California and within the Colorado River Delta in the years 2005, 2006, 2007, and 2011 (Table 1). Collected specimens were preserved in 95% ethanol in the field and stored at -20°C upon returning to the laboratory for use in both molecular and morphological analyses. For this study, we divided the coastline of the Northern Gulf of California into 4 regions: the Baja Coast, the Delta Edge, the Colorado River Delta, and the Sonora Coast. A full list of collection sites along with the abbreviated designation for each site are shown in Table 1. In addition to the field-collected specimens, formalin-fixed specimens from the Natural History Museum of Los Angeles County (LACM) and Scripps Institute of Oceanography were examined for their morphology to confirm species identification. A list of these “historical” specimens is included in the supplementary materials (Table S1).

Phenotypic Analyses

The dorsal-lateral scale counts, vertebral counts, hemal spine morphology, and the position of the first dorsal fin were used as species-distinguishing features for this study; these features were chosen based on the species description of *Colpichthys hubbsi* (Crabtree, 1989). The first dorsal fin position is assessed based on the linear length from the snout to the first dorsal fin origin, normalized by the total length. Length measurements were taken with dial calipers. Morphometric and meristic counts were performed according to Hubbs and Lagler (2004). Radiographs of specimens were taken for examination of their vertebral counts and hemal spine morphology. Six specimens from location NWC and 3 specimens from location

YAV were too small to accurately assess their hemal spine morphology; they were thus excluded from the hemal spine data set.

Molecular Procedures

Genomic DNA was isolated from caudal peduncle muscle tissues using the DNeasy Blood and Tissue Kit (Qiagen, Inc., Valencia, CA). The mitochondrial cytochrome *b* gene was amplified by PCR using primers AJG15 and H5 (Akihito et al., 2000) under the following cycling conditions: 95°C for 2 min; 40 cycles of 95°C for 30 sec, 51°C for 30 sec, followed by 72°C for 90 sec; then finally 72°C for 10 min. PCR products were size-checked on 1.5% agarose gel and cleaned with ExoSap before Sanger sequencing with BigDye Terminator v3.1. Samples were then submitted to DNA Analysis Facility on Science Hill at Yale University for capillary electrophorasis.

Amplification of the nuclear RAG1 gene was carried out using primers RAG1-2533F and RAG1-4090R (López et al., 2004) under the following cycling conditions: 95°C for 2 min; 40 cycles of 95°C for 1 min, 51°C for 1 min with an incremental increase of 0.5°C per cycle until 55°C, followed by 72°C for 90 sec; then finally 72°C for 10 min. Post-PCR processing and sequencing of the RAG1 gene were performed as described above. From the sequences we were able to identify four polymorphic sites; several individuals display ambiguity in base calling at such sites, these were tagged as putative heterozygotes. A total of four RAG1 variants were inferred from the sequence data. Homozygous individuals were used to confirm the existence of three RAG1 variants; the remaining RAG1 variant was inferred from a heterozygous individual. To confirm the gametic phase of these RAG1 sequences, we employed allele-specific polymerase chain reaction (ASPCR) as detailed in Wu et al. (1989). Three allele-specific primers

(T717F: 5'-CTACAAAATCTTCCAGGAT-3', T770R: 5'-TTTATCTAAGGCTGCCCTCCAGA-3', and A909F: 5'-AACTGGTGCCCTCAGAAGAA-3') were designed to uniquely amplify one of the two diploid copies of RAG1 in the putative heterozygotes. The 3' terminating nucleotide of these primers are designed to specifically anneal at the predetermined polymorphic site, thus enabling unique amplification of alleles and determination of the RAG1 genotypes of the heterozygotes.

Six microsatellite loci (Odont08, Odont09, Odont11, B18, B19, B39) were amplified for genotyping. Primer sequences for loci Odont08, Odont09, and Odont11 were obtained from Beheregaray and Sunnucks (2000). Primer sequences for loci B18, B19, B39 were obtained from Byrne and Avise (2009). Loci which can be easily differentiated by lengths were multiplexed two at a time using QIAGEN Multiplex PCR kits with 6-FAM labeled M13 primer (Boutin-Ganache et al. 2001). Only the forward primers of each locus are 5' tagged with the M13 sequence, and reactions were performed under the following cycling conditions: 95°C for 15 min; 25 cycles of 94ºC for 30 sec, 55ºC for 90 sec, 72ºC for 1 min; then another 25 cycles of 94ºC for 30 sec, 50ºC for 90 sec, 72ºC for 1 min; and finally 60ºC for 30 min. Diluted PCR products were submitted to the UCLA GenoSeq core facility for genotyping on an ABI3730 (Applied Biosystems).

Molecular Analyses

All sequences were aligned, trimmed, and analyzed using Geneious Pro ver. 5.5.6 (Drummond et al., 2011). A Maximum Likelihood tree of Cyt *b* sequence was made with MEGA7 (Kumar et al., 2016) and rooted using *Atherinops affinis* as the outgroup. A time-calibrated tree was inferred using the Reltime method (Tamura et al. 2012) and the Tamura-Nei model (Tamura and Nei, 1993). The timetree was computed using one calibration constraint at the node of the split between *C. regis* and *C. hubbsi*; the calibration point of 4.83Ma was obtained from the age of the ash bed within the Bouse Formation, the rock unit from which two fossilized *C. regis* were documented (Todd, 1976; Spencer et al., 2013).

Microsatellite data were obtained from 101 out of 103 specimens and imported into Geneious for allele scoring and analyzed using STRUCTURE ver. 2.3.4 (Pritchard et al. 2000). STRUCTURE is a program that uses a Bayesian clustering approach to infer population structure from genotypic data. Under an admixture model of K populations or genetic groupings, STRUCTURE computes the q-value, a quantity between 0 and 1, that reflects the proportions of an individual's genome originating from a certain population. Hybridization can then be inferred based on the computed q-value; for example, first-generation (F1) hybrids of two populations (K=2) are expected to have a q-value close to 0.5 (Vähä and Primmer, 2006). We ran STRUCTURE analyses using an admixture model with 1,000,000 MCMC generations after a burn-in period of 100,000. Sampling locations were incorporated as prior parameters in the *LOCPRIOR* model to improve clustering for our data set; the *LOCPRIOR* model is preferred when there are weak signals of structure due to low number of markers analyzed (Hubisz et al., 2009). Both correlated and independent allele frequency models were used with number of groups set from K=1 to K=4; for each K value, analysis was repeated ten times. Of the two allele frequency models, the correlated model provides greater power in differentiating closely related populations (Falush et al., 2003). The optimal K value was informed by ΔK (Evanno et al., 2005) as computed using Structure Harvester (Earl et al., 2012). Output of the STRUCTURE analyses were summarized in CLUMPP ver. 1.1.2 (Jakobsson and Rosenberg, 2007) and visualized in DISTRUCT ver. 1.1 (Rosenberg, 2004).

Multivariate Analysis

We further analyzed our phenotypic and microsatellite data by discriminant analysis of principal components (DAPC) using the R-package adegenet (Jombart et al., 2011). This analysis is used to help describe the phenotypic and genetic clustering of our specimens. Cluster priors for DAPC were identified by *k*-means clustering; based on BIC scores we selected the 2-cluster model to use for the DAPC. To perform DAPC, the program transforms the data into principal components, which is then analyzed through discriminant analysis. Repeated cross-validation was used to evaluate the optimal number of principal components to retain for the discriminant analyses to avoid overfitting of data. For the phenotypic analysis, we used measurements of the 4 species-distinguishing features as our data and retained one PC (91.28% of the cumulative variance) for generating the morphology discriminant function. For the microsatellite analysis, we used the allele frequencies of the 6 loci as our data and retained 20 PCs (66.35% of the cumulative variance) for generating the microsatellite discriminant function.

Demographic Histories

To explore the demographic histories of *Colpichthys*, Cyt *b* data sets were separated by species according to the Cyt *b* gene tree (Fig. 1B). 44 sequences from *C. hubbsi* and 52 sequences from *C. regis* were analyzed for their haplotype diversity. Median-joining haplotype networks were drawn in POPART ver. 1.7 (Leigh and Bryant, 2015). Cyt *b* haplotypes were examined by various neutrality tests – Tajima's *D* (Tajima, 1989), Fu and Li's D^* and F^* (Fu and Li, 1993), Fu's *Fs* (Fu, 1997), and R_2 (Ramos-onsins and Rozas, 2006) – in DnaSP v. 5.10.01(Librado and Rozas, 2009). P-values for these test statistics were obtained by coalescent simulations of 1000 replicates, conditional on the number of segregating sites. Pairwise differences were analyzed by mismatch distributions in Arlequin ver. 3.1 (Excoffier and

Schneider, 2005); coalescent simulations were run with 10,000 bootstrap replicates. Past demographic expansions can be inferred from haplotypes based on the mismatch distribution of the pairwise differences; an uneven, multi-modal distribution is expected for populations under demographic equilibrium, whereas a “smooth” uni-modal distribution is expected for populations that had undergone demographic expansion (Rogers and Harpending, 1992). The observed mismatch distributions of the two species were compared to the expected mismatch distribution under a sudden expansion model using the sum of square deviation (SSD) and the raggedness index.

RESULTS

Pure-bred Specimens from Both Species Can Be Distinguished by Morphology

Although specimens of *C. hubbsi* superficially resemble those of *C. regis* in morphology, pure-bred *C. regis* and *C. hubbsi* can be readily distinguished based on phenotypic characters documented by Crabtree (1989). Consistent with the findings thereof, our morphological analysis based on both modern and historical specimens, summarized in Fig. 2, shows that *C. regis* (green and gold) differs from *C. hubbsi* (pink) in having higher dorsal-lateral scale counts, higher vertebral counts, and shorter snout to first dorsal fin origin distance. In addition, the hemal spines of the most anterior caudal vertebrae of *C. regis* exhibit expanded, flattened processes (Fig. 2D, Fig. 3B). These differences are significant enough to classify most of our specimens; however, there appears to be enough variation within each species that some hybrids may be difficult to ascertain based solely on the visual inspection of morphology.

Introgression of Cytochrome b *and RAG1 Haplotypes*

Based on the Maximum-Likelihood phylogeny of Cyt *b*, our samples grouped into two clades separated by ~6% sequence divergence, corresponding to the two species of *Colpichthys* (Fig. 1b). Consistent with the geographic range of the species, all samples with *C. hubbsi* haplotypes are exclusively found within the Delta region. On the other hand, Baja Coast and Sonora Coast populations were all composed of *C. regis* haplotypes. Both species were expected to cohabit the PRI, SGU, and TRC localities, but all sampled specimens apart from those of TRC show a mitochondrial affinity to the *C. regis* clade. Within the Delta, five specimens collected from *C. hubbsi* native populations—ELZ, CHY, and SFG—were found to possess the *C. regis* haplotype; these specimens are morphologically indistinguishable from other *C. hubbsi* specimens.

We identified four RAG1 variants (labeled as H_1, H_2, H_3, and *R*; Table 2) based on four polymorphic sites in the RAG1 sequence data (975bp). We find that H_1, H_2, and H_3 are common in populations of *C. hubbsi* in the Delta, while *R* is fixed in all populations of *C. regis* in the Baja Coast and Sonora Coast regions. Thus, the two species are well differentiated by their RAG1 genotypes, with the diagnostic SNPs at positions 717 and 909. Based on this we identified six specimens that are heterozygous at both of these positions; these RAG1 heterozygotes are flagged as putative hybrids. Four of the six putative hybrids occur at the Delta Edge, while the remaining two are found within the Delta proper (Table 3). No RAG1 heterozygotes are detected in any of the Baja and Sonora collection sites.

As noted above, a discrete phenotypic difference between the two species of *Colpichthys* is the morphology of the anterior-most caudal hemal spines (Fig. 2D, Fig. 3). This trait difference associates with the morphology of the gas bladder. In all Atherinopsinae species except *C. hubbsi*, the gas bladder extends posteriorly past the visceral cavity (Crabtree, 1989). Compared to other Atherinopsinae species, the hemal spines of *C. hubbsi* appears to show no signs of expansion. This phenotype is observed in all specimens possessing any of the three *C. hubbsi* RAG1 alleles (H_1, H_2, or H_3), suggesting that there may be a correlation between this phenotype and the genotype at the RAG1 locus.

Bayesian Analyses of Microsatellites Show Evidence of Admixture

Based on the evaluation of ΔK, the STRUCTURE analyses recovered two genetically distinct clusters corresponding to the two species of *Colpichthys* (Fig. 4; for K=3 and K=4, see

Fig. S1). Under the admixture model, the q-value represents admixture proportion of each individual. A q-value threshold could be used to distinguish putative purebred individuals from putative hybrids; for optimal efficiency, this threshold is set to 0.1 (Vähä and Primmer, 2006). Under this q-value threshold, populations in the Delta proper—ELZ, SFG, and CHY—were identified as purebred *C. hubbsi*. On the other hand, several populations away from the Delta such as ANI, NWC, and YAV were identified as purebred *C. regis*. All individuals within or near the Delta Edge—PRI, SGU, and TRC—show substantial degrees of genetic admixture, having q-values between 0.1 and 0.9. Moreover, among these populations there appears to be a gradient in q-values as a function of geographic distance from the mouth of the Colorado River. TRC has an average q-value of 0.17 while PRI, which is farther from the river mouth, has an average q-value of 0.72 (Fig. 4A). In the analysis with the correlated allele frequency model, a relatively weak signal for admixture was detected in the ADR population; this signal is absent in the analysis with the independent allele frequency model (Fig. 4B).

Consilience of Phenotypic and Nuclear Data

There are only two clusters, consequently only one discriminant function was generated for each of the DAPC analyses. Fig. 5 shows the plot of the discriminant function obtained from microsatellite data against the discriminant function obtained from phenotypic data. The results show that the two species are well separated into two distinct clusters by both data sets, and that the cluster assignment from both data sets are mostly congruent. Furthermore, both data sets show stratification across the four geographic regions, with the Delta Edge specimens situating mostly in between clusters. RAG1 heterozygotes, plotted as triangles on Fig. 5, are shown to be more phenotypically *C. hubbsi* than *C. regis*.

Genetic Diversity and Demographic History

The Cyt *b* median-joining haplotype networks of the two species are displayed in Fig. 6. Twenty-Eight *C. hubbsi* haplotypes are recovered from the four delta populations; in contrast, only ten haplotypes are found in *C. regis*. Among *C. regis*, haplotype diversity is markedly different between the Baja Coast and the Sonora Coast; with the exception of one individual from PRI, all non-major haplotypes are found exclusively in Sonora Coast populations. In fact, the only two locales that are excluded from the major haplotype are the two southernmost Sonora Coast sites, SLD and YAV. The five mitochondrial-introgressed individuals are also displayed in Fig. 6; all five possess the major *C. regis* haplotype. Neutrality tests were conducted with the exclusion of the mitochondrial-introgressed individuals; all neutrality tests report significant departure from neutrality. Significantly negative values for Tajima's *D* and Fu's Fs in both species are consistent with either population growth or selective sweep (Table 4). The sum of square deviations and raggedness indices for the mismatch distributions are close to zero and not significant (Table 4), thus we are unable to reject the sudden expansion model. Fig. 7 shows the nucleotide mismatch distributions for both species; *C. hubbsi* has a bimodal distribution, while *C. regis* has a unimodal distribution that matches closely with the expected curve under sudden expansion.

DISCUSSION

In the species description of *Colpichthys hubbsi*, Crabtree (1989) recognized the region just north of San Felipe (a region designated as Delta Edge in this study) as an area of syntopy of the two *Colpichthys* species. While external morphology of specimens collected in 1968 confirms the co-occurrence of the two species at this region (Table S2), there is no clear indication of historical hybridization. This study presents the first evidence for the modern hybridization of the sister species based on morphological and molecular data. Results from both STRUCTURE and DAPC analyses strongly support our hypothesis that the Delta Edge region currently represents a hybrid zone. Our STRUCTURE results based on the correlated allele frequency model further raise the possibility that hybridization may also be occurring on the eastern border of the Delta at Bahía Adair (ADR). The fact that weak admixture signals at ADR are not detected using our other approaches suggests that hybridization at this region, if present, is likely infrequent.

Directional Introgression

Our study detected the presence of *C. regis* mitochondrial and RAG1 haplotypes in Delta populations; this could be explained by either 1) the dispersal of first generation hybrids from the hybrid zone to the Delta, or 2) gradual backcrossing of hybrids with parental *C. hubbsi* from the hybrid zone to the Delta (i.e. introgressive hybridization). Our data lend support to the latter, as unlike the RAG1 putative hybrids at the Delta Edge, specimens that possess *C. regis* mitochondrial and/or RAG1 haplotypes within the Delta do not exhibit intermediate morphology. Without genetic data from these loci, these specimens would indisputably have been classified as *C. hubbsi* and thus they likely represent backcrosses. We conclude that *C. regis* haplotypes have been introduced into Delta *C. hubbsi* populations via introgressive hybridization. Thus, one important conclusion based on the samples well within the delta is that introgression of alleles is not confined strictly to the hybrid zone.

We failed to detect any *C. hubbsi* haplotypes outside of the Delta and Delta Edge regions. Thus, introgression does not appear to be reciprocal. There are only two estuaries between Estero Primero (PRI) and Las Animas (ANI), so introgression to the south may be limited by the scarcity of estuarine habitats. Nevertheless, more sampling effort in these sites is needed to rule out the possibility of reciprocal introgression.

Parental Sex Asymmetry in Hybridization

The hybrid zone at the southwestern Delta border is comprised of three neighboring tidal channels—Estero Primero (PRI), Estero Segundo (SGU), and Estero Tercero (TRC); these locales span roughly 10km along the coastline. Phenotypic and microsatellite data reveal a narrow geographic cline across this region (Fig. 4, Fig. S2). Based on geography we can infer that first generational hybrids occur most frequently in PRI and SGU, yet all thirteen specimens from PRI and SGU possess the *C. regis* mitochondrial haplotype. As the mitochondria is matrilineally inherited, this pattern may indicate that hybridization is occurring between female parental *C. regis* and male parental *C. hubbsi*. Assuming that this mitochondrial pattern is not due to random factors or drift, sex asymmetry in hybridization may be the result of pre-zygotic or post-zygotic factors (Wirtz, 1999). Unlike their better studied relatives, the grunions (genus *Leuresthes*), which practice beach spawning, relatively little is known about the life history and spawning behavior of *Colpichthys*. However, pre-zygotic bias in hybridization may be imposed by a morphological difference in reproductive organs. Crabtree (1989) noted that both species of *Colpichthys* possess ovipositors which are used to deposit mucous-covered eggs on aquatic vegetation such as mangrove roots. The ovipositors of the two species differ greatly in size, with *C. hubbsi* having nearly double the ovipositor size compared to that of *C. regis*. Other pre-zygotic biases such as ecological or behavioral differences may also play a role, but these hypotheses remain to be tested by future studies.

Ecological Divergence

Due to the lack of available genetic data from historical samples, the onset of hybridization cannot be easily constrained. However, the ~6% mitochondrial sequence divergence between *C. regis* and *C. hubbsi* suggests that the two species had been in separation for a substantial duration and supports the notion that secondary contact may have been relatively recent. This amount of mitochondrial divergence is comparable to that of another Gulf fish species pair, *Gillichthys mirabilis* and *Gillichthys detrusus* (Swift et al., 2011). Within the Gulf of California, *Gillichthys* share similar phylogeographic distributions with *C. regis* and *C. hubbsi*. Like *C. hubbsi*, *G. detrusus* is narrowly restricted to the silty tidal channels of the Colorado River Delta; its congener, *G. mirabilis*, can be found in the estuaries and lagoons on both coasts of the Gulf. As sister species, *G. detrusus* and *G. mirabilis* have been estimated to have diverged at ~5Ma (Swift et al, 2011). Assuming that the mitochondrial rate in *Colpichthys* is comparable to that of *Gillichthys*, the ages of these two clades are likely similar. Given that both *C. hubbsi* and *G. detrusus* are Delta endemics, their divergences from their respective sister clades may have been linked to the formation of the Delta estuaries after the opening of the Gulf of California. The distribution of *C. regis* is discontinuous across the Delta, but fossils of *C. regis* have been discovered in the Bouse Formation at the location of Cibola Lake, Arizona (Todd, 1976). The marine fossils found in this rock unit have been used to argue for marine incursions in the late Miocene / early Pliocene (Poulson and John, 2003; McDougall and Miranda Martinez, 2014). However, Spencer and Patchett (1997) and Roskowski et al. (2010) favored a strictly lacustrine interpretation of the Bouse Formation based on isotopic analyses. Details in the interpretation of this rock unit may vary, but stratigraphic and tectonic reconstructions have estimated that the Colorado River had reached the Gulf of California by ~4-5Ma (Winker and Kidwell, 1986; Dorsey et al., 2007; Spencer et al., 2013; Crossey et al. 2015; Howard et al., 2015). Speciation leading to the Delta endemics presumably followed the establishment of the Delta habitats around this time.

The parapatric distribution of *Colpichthys* in the Northern Gulf suggests that *C. regis* may have diverged from *C. hubbsi* due to ecological differentiation reinforced by divergent selection on local adaptations. A possible driver for divergent selection could be the gradient in salinity in the Northern Gulf. While virtually no physiological data exist for *C. hubbsi*, *C. regis* is known to tolerate hypersaline conditions as it is frequently observed in lagoons with salinity values reaching 50 psu (Castro-Aguirre and Espinosa-Pérez, 2006). Prior to the completion of Hoover Dam in 1935, the water output at the Delta was estimated to be 16 to 18 billion cubic meters per year (Stockton and Jacoby 1976). According to Levín and Sánchez (1999), pre-dam salinity values of 35.2-35.7 could be found in the Delta. Thus, it is reasonable to infer that salinity gradients in the delta may have played a role in reproductive isolation between the sister species. Other potential selective factors could involve the unusual silty conditions and tides in the Delta and the limited depth which might be associated with the reduced gas bladder of *C. hubbsi* as noted above.

Evolutionary Implications

The introgressive hybridization revealed in this study may have potential evolutionary consequences. For example, introgression could lead to the transfer of adaptive variations from one species to another; examples of such adaptive introgression are abundant in both plant and animal species (Martin et al., 2006; Whitney et al., 2010; Pardo-Diaz et al., 2012). Conversely, introgressive hybridization may bring about negative effects on endemic species. An example of this could be found in the case of the introduction of smallmouth bass (*M. dolomieu*) to the native population of Guadalupe Bass (*Micropterus treculii*) in Central Texas. Despite supplemental stocking of *M. treculii*, Littrell et al. (2007) found that introgressive hybridization with *M. dolomieu* had driven *M. treculii* to local extirpation in the Blanco River. For sympatric or parapatric species, homogenization by introgression may be exacerbated by the breakdown of ecological species barriers resulting from habitat changes (Olden et al., 2004). Such "speciation reversal" has been documented in whitefish in Swiss lakes where unique spawning habitats were lost as a result of eutrophication (Vonlanthen et al., 2012). If salinity was indeed one of the ecological drivers for *Colpichthys* speciation, as the current Delta continues to desiccate, the rise in salinity within the Delta could create habitats more favorable to *C. regis* genotypes. This decrease in habitat heterogeneity may lead to an increase in permeability to gene flow at the hybrid zone and, subsequently, the degradation of genetic integrity in *C. hubbsi*. In the long run, the risk of extinction by "reverse speciation" for *C. hubbsi* will depend on the extent to which habitat heterogeneity is lost in the Northern Gulf. While this study only identified a few backcrossed hybrids based on the limited genetic loci sampled, it is possible that many of the morphological "pure bred" *C. hubbsi* are in fact "cryptic" hybrids. Depending on the frequency and relative abundance of such "cryptic" hybrids, interspecific gene flow may be more pervasive than suggested by our data. To assess whether "speciation reversal" could pose a threat to *C.*

hubbsi, future whole-genome analyses will be required to quantify the amount of genomic introgression as well as identify loci that may remain under divergent selection in the face of gene flow.

Given the geographic restriction of *C. hubbsi,* one immediate concern is whether the reduced Colorado River outflow has directly contributed to a population decline. The rise in salinity at the Delta has been implicated to be linked to the decline of the endemic Delta clams, *Mulinia coloradoensis* (Rodriguez *et al*., 2001). Based on our results from neutrality tests and mismatch analyses, we find significant departure from neutrality but no evidence to support a recent bottleneck in *C. hubbsi*. On the other hand, contrary to expectation, *C. regis* exhibits surprisingly low genetic diversity compared to its geographically restricted congener. This pattern is seen in all three types of molecular markers (Cyt *b*: Fig. 6; RAG1: Table 2; microsatellites: Table S3). The results from this study are consistent with demographic expansions in the recent past, although other factors, such as selective sweep or fine-scaled population structure (Ptak and Przeworski, 2002), could produce similar mtDNA patterns. These confounding factors cannot easily be ruled out without more complete sampling. Nevertheless, we offer the following scenario as a possible explanation for the observed data. In the mismatch analysis, the model parameter τ estimates the age of demographic expansion. Our results indicate that *C. hubbsi* may have experienced a more ancient expansion relative to *C. regis* (Table 4). One possible explanation for this could come from climate fluctuations in the Pleistocene. Recent work by Dolby et al. (2016) demonstrated that formation of glacial refugia due to interactions between coastal geomorphology and Pleistocene sea-level fluctuations could have lasting impacts on modern population structures of estuarine fish. During the Pleistocene, wetter climates may have increased Colorado River outflow into the Gulf; the gentle sloping bathymetry of the Northern Gulf together with the lowering of sea-level could have enabled habitats for *C. hubbsi* to extend much further south, allowing for population growth. Meanwhile, the steeper coastal shelves in the lower portion of the Gulf could have restricted suitable refugial habitats for *C. regis* during sea-level lowstand. The lack of genetic diversity in *C. regis* populations may represent rapid range expansion from these refugia following sea-level rise at the end of the Pleistocene. Crucial to this hypothesis is the precise timing of the demographic and range expansion. More complete sampling of *C. regis* within each of the modern localities is needed in future studies to further explore this issue.

CONCLUSION

While it has long been recognized that the two species of *Colpichthys* overlap in range, this study is the first to demonstrate clear evidence for their hybridization at the southwestern edge of the Colorado River Delta. Combining multiple lines of evidence, both phenotypic and genetic, we are able to detect signals of admixture consistent with directional introgressive hybridization from *C. regis* to *C. hubbsi*. With the limited genetic loci examined in this study, we expect a certain degree of underestimation of the genomic introgression. Our observations of the mitochondrial haplotypes of specimens from the hybrid zone revealed a possible sex bias in hybridization, with hybrids conceived from maternal *C. regis* and paternal *C. hubbsi*. While we find no evidence for population decline in *C. hubbsi,* continued introgression may jeopardize the genetic integrity of this endemic species. Though we cannot definitively link the hybridization to the anthropogenic reduction of freshwater discharge in the Colorado River, we caution that persistent desiccation at the Delta may diminish the habitat required for *C. hubbsi* to remain as a distinct species. Genetic diversity of *C. regis* is revealed to be lower than expected for a healthy, widely distributed species. Neutrality tests and mismatch analyses suggest possible recent demographic expansions in *C. regis,* which could be linked to glacial sea-level fluctuations. We recognize that our demographic analyses based solely on mitochondrial DNA are limited in power. Our work highlights the need for closer monitoring and further demographic investigations of *Colpichthys* populations not just within the Delta, but also throughout the Northern Gulf.

TABLES

Locality	Code	Region	Sampling Date	Coordinates	N =
Las Animas	ANI	Baja Coast	4 May 2007	28° 47.855N - 113° 20.894W	9
Estero Santa María	ESM	Baja Coast	10 Nov 2005	30° 44.73' N - 114° 42.01W	1
Estero Primero	PRI	Delta Edge	22 Sep 2006	31° 11.903N - 114° 53.437W	10
Estero Segundo	SGU	Delta Edge	24 Sep 2006	31° 15.355N - 114° 53.011W	3
Estero Tercero	TRC	Colorado River Delta	16 Jun 2011	31° 17.354N - 114° 54.831W	3
Estero Chayo	CHY	Colorado River Delta	3 Dec 2005	31° 40.119N - 114° 41.529W	15
"Port Elizabeth"	ELZ	Colorado River Delta	3 Dec 2005	31° 49.405N - 114° 49.566W	18
"Shrimp Farm" El Golfo	SFG/ SFGb	Colorado River Delta	2 Dec 2005/ 15 Jun 2011	31° 46.480N - 114° 34.931W	7 12
Bahía Adair	ADR	Sonora Coast	14 Jun 2011	31° 32.244N - 113° 58.910W	9
Northwest of Cholla	NWC	Sonora Coast	14 Jun 2011	31° 27.822N - 113° 37.898W	8
Estero del Soldado	SLD	Sonora Coast	6 Oct 2006	27°57'26"N - 110°58'48"W	1
Yavaros	YAV	Sonora Coast	5 Mar 2006	26° 40' 42"N - 109° 29' 36"W	7
				Total=	103

Table 1 List of collection localities.

RAG1 Variants	Position 495	**Position 717**	Position 770	**Position 909**
H_1	G	**C**	G	**G**
H_2	G	**C**	T	**G**
H_3	A	**C**	T	**G**
R	G	**T**	G	**A**

Table 2 Variants of RAG 1. H_1, H_2, H_3 are common in populations of *C. hubbsi* in the Delta; *R* is fixed in all populations of *C. regis*. SNPs at positions 717 and 909 (bolded) are species-specific and heterozygotes at these positions are used to identify putative hybrids.

		heterozygotes	non-heterozygotes	het./non-het. ratio
Delta	Port Elizabeth (ELZ)	0	18	0
	Shrimp Farm (SFG)	1	15	0.067
	Estero Chayo (CHY)	1	12	0.083
	Estero Tercero (TRC)	0	3	0
Delta Edge	Estero Segundo (SGU)	1	2	0.500
	Estero Primero (PRI)	3	7	0.429

Table 3 Counts of RAG1 heterozygotes and non-heterozygotes at positions 717 and 909 at each of the Delta and Delta Edge collection sites.

		Mismatch Analysis (sudden expansion)			
	SSD	raggedness	θ_0 (99% CI)	θ_1 (99% CI)	τ (99% CI)
C. regis	0.0015	0.0779	0.018 (0 – 0.965)	66842 (0 – 99999)	1.110 (0 - 4.605)
C. hubbsi	0.0092	0.0158	1.332 (0 – 13.558)	17339 (6 – 99999)	5.827 (0.105 – 27.619)

			Neutrality Tests				
	N	S	Tajima's *D*	Fu & Li's *D**	Fu & Li's *F**	Fu's *Fs*	R_2
C. regis	47	11	**-2.1410****	**-3.3113***	**-3.4502****	**-7.1724***	**0.0514**
C. hubbsi	44	40	**-1.5645**	**-2.2374**	**-2.3783***	**-16.1789****	**0.0596**

Table 4 Mismatch analysis and Neutrality test results. N = number of samples; S = number of segregating sites; SSD = sum of square deviation. (Bold: $p < 0.05$; *: $p \leq 0.02$; **: $p \leq 0.005$).

FIGURES

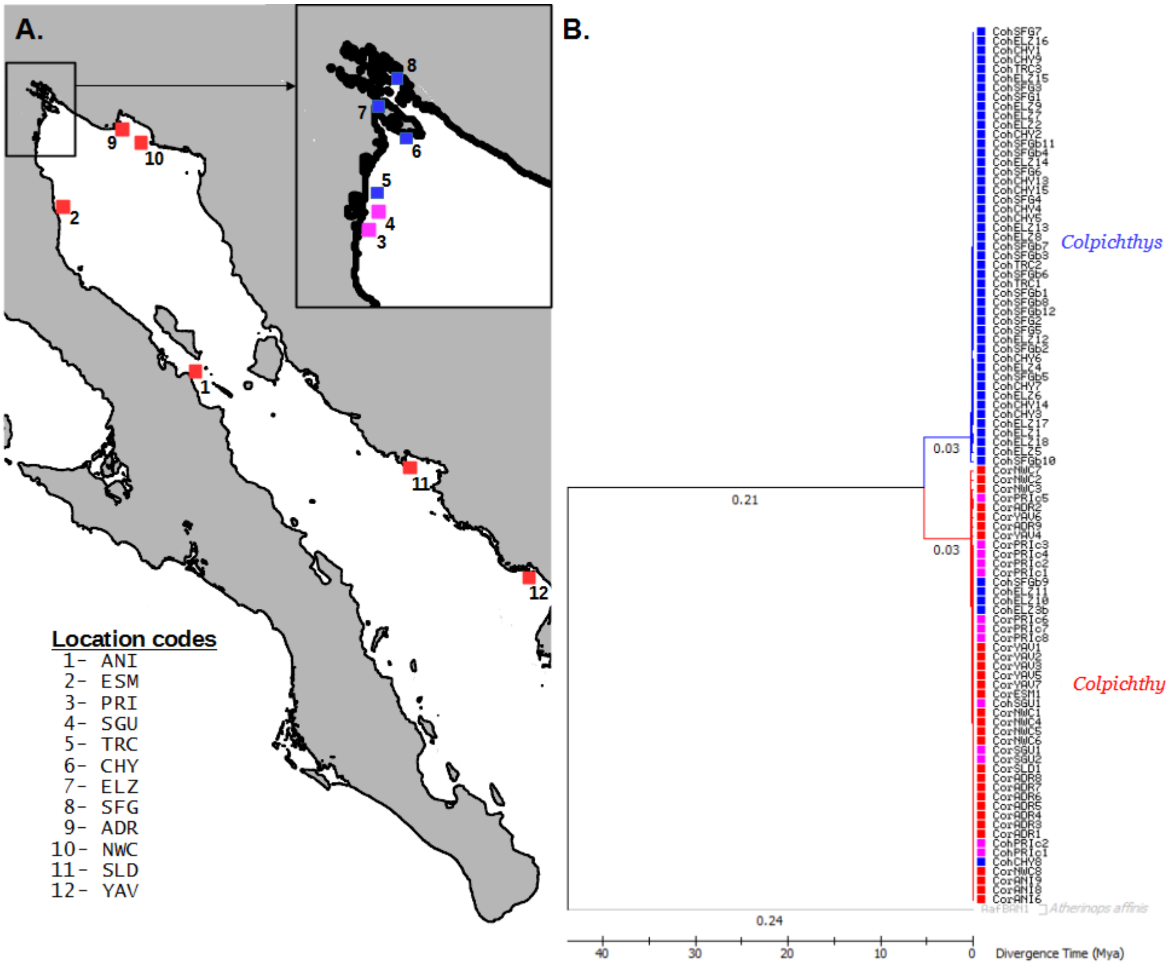

Fig. 1 (A) Map of the coastline of the Gulf of California with sampling localities, color coded by species distribution. *C. hubbsi* native populations are colored blue; *C. regis* native populations are colored red; and populations where both species co-occur are colored magenta. (B) Time calibrated Maximum-Likelihood (ML) tree of cytochrome *b* (872bp). Branch lengths (substitutions per site) are indicated below the branches. One specimen of *Atherinops affinis* was used as outgroup to root the tree. Branch labels show branch lengths (substitutions/site).

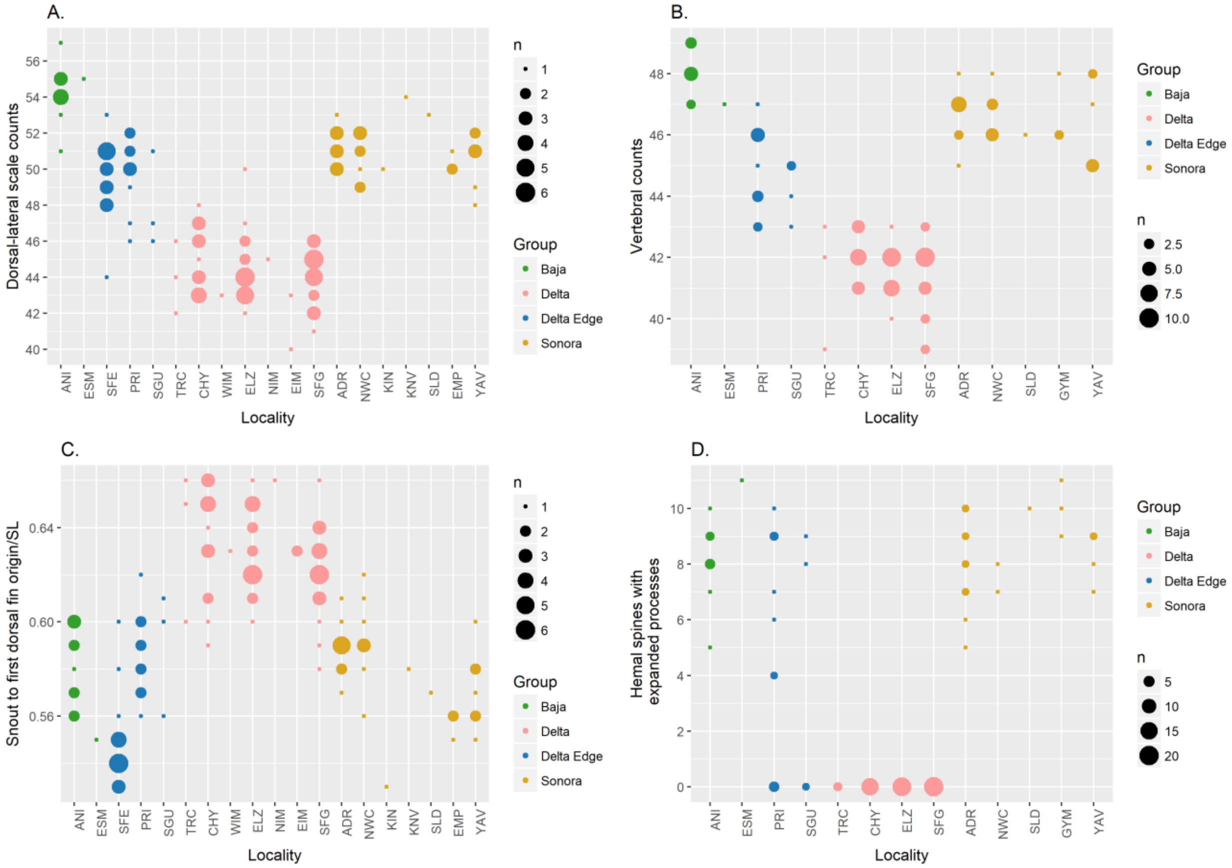

Fig. 2 Dot plots showing external (A and C) and vertebral (B and D) morphological data; size of dots is proportional to the number of specimens (n). Six specimens from NWC and three specimens from YAV were excluded from the hemal spine data set as these specimens were too small to reliably identify the presence/absence of the expanded process on the hemal spines (See Fig. 3).

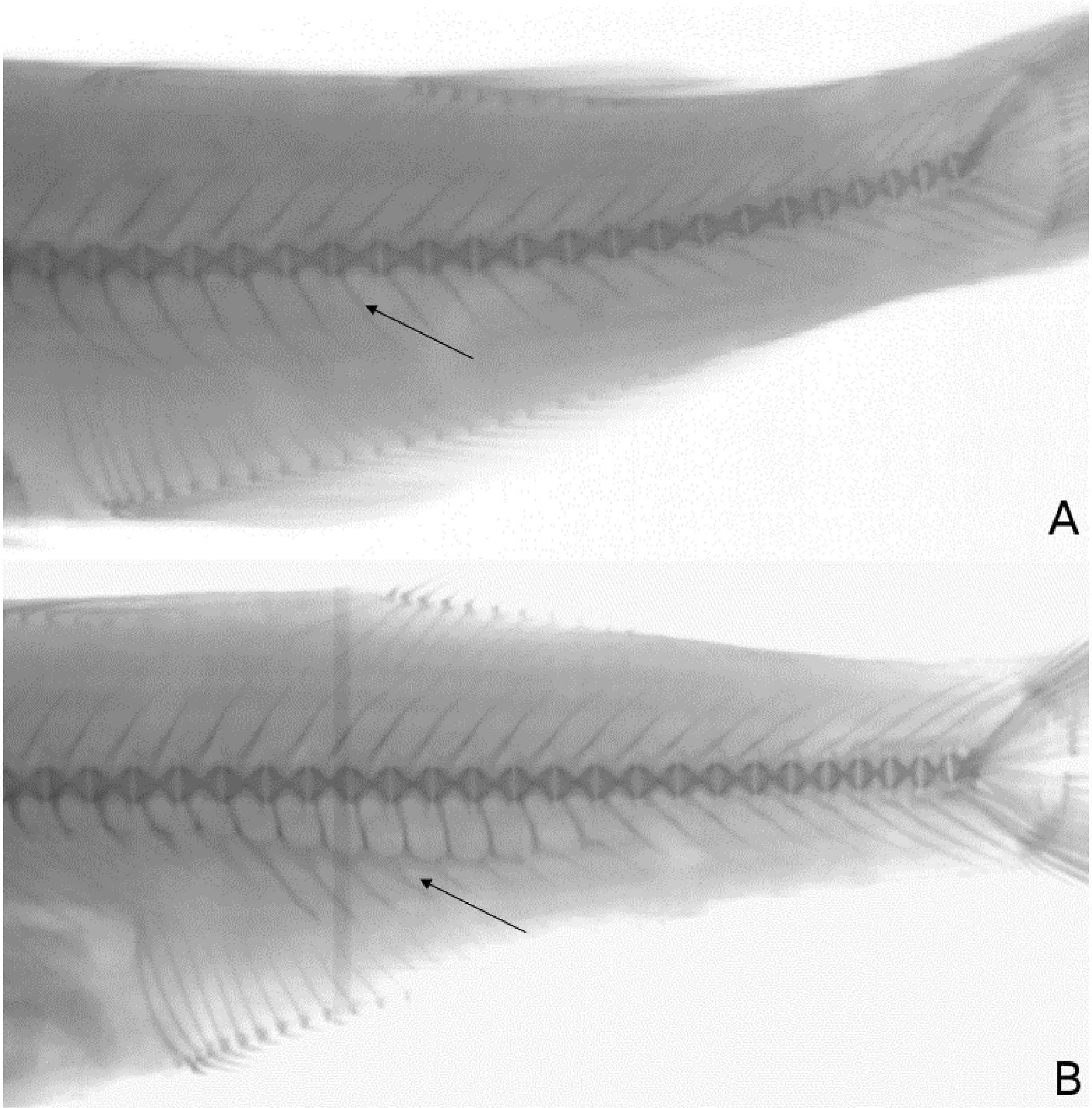

Fig. 3 Radiographs of a *C. hubbsi* specimen (A) and a *C. regis* specimen (B); arrow in (B) shows the expanded process in the hemal spines of *C. regis*.

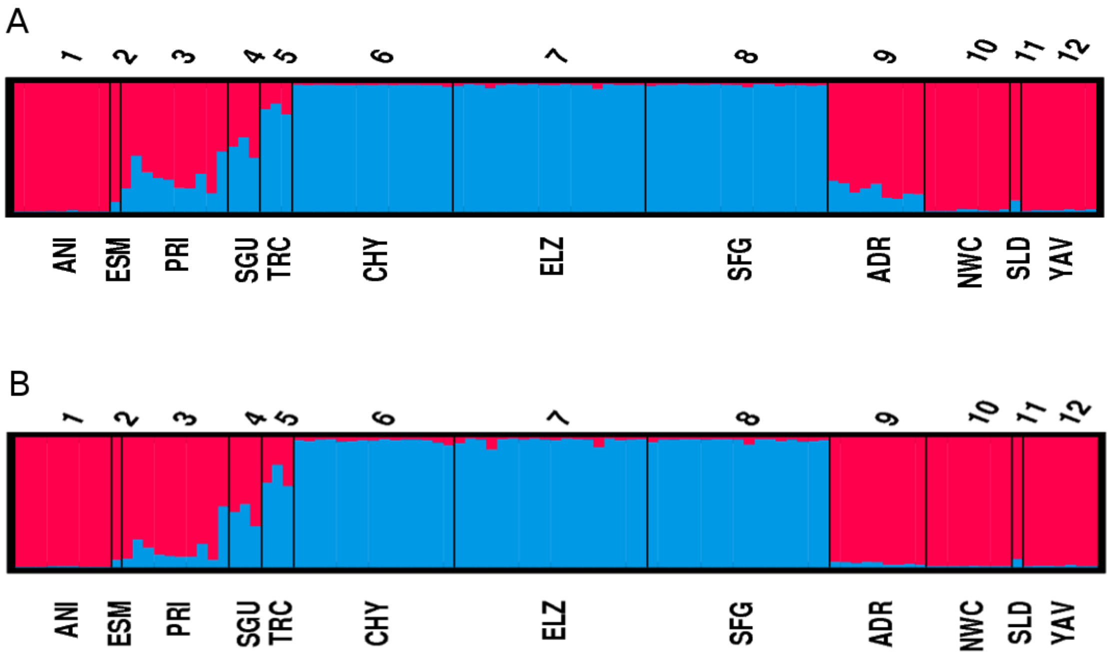

Fig. 4 STRUCTURE bar plots showing q-values of each individual in each sampling locale for K=2. Color for each bar represents admixture proportions of *C.regis* (red) and *C. hubbsi* (blue). Correlated allele frequency model (A). Independent allele frequency model (B).

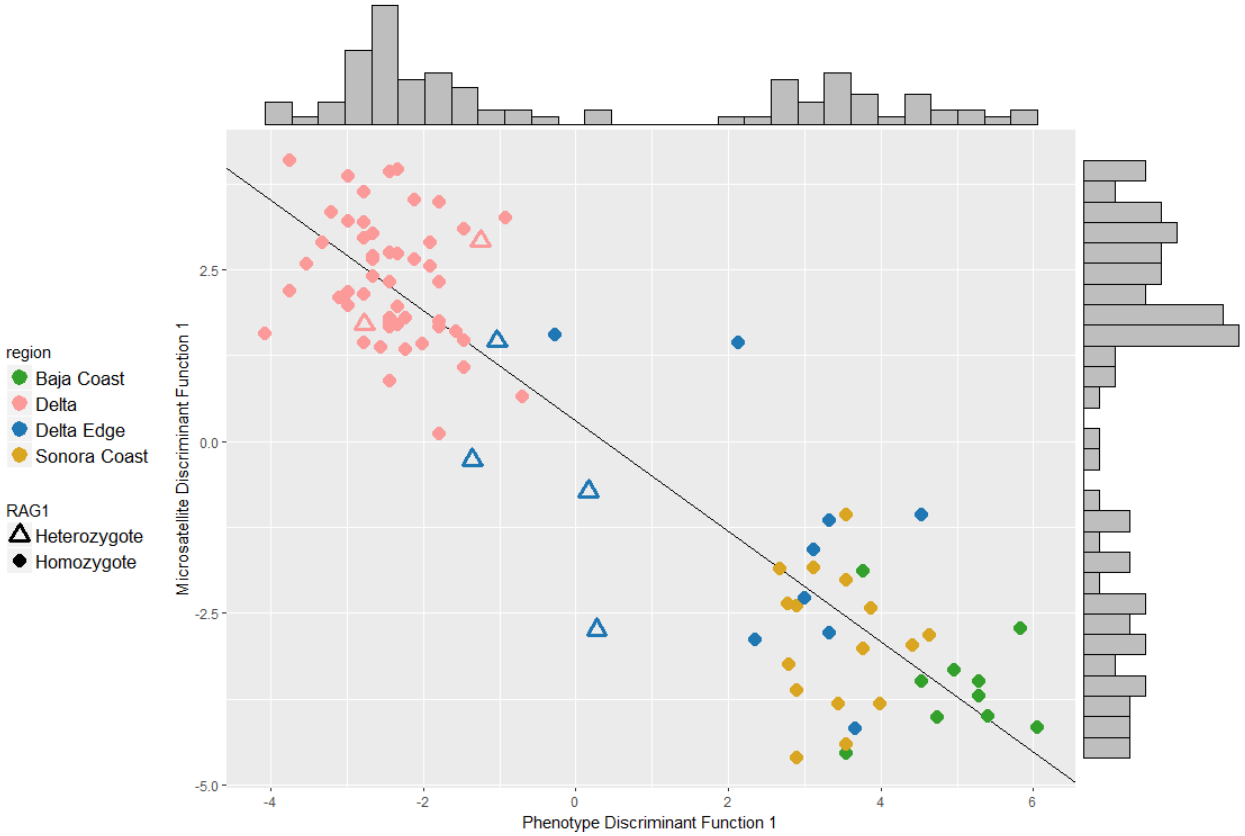

Fig. 5 Scatter plot of the first discriminant functions in the DAPC analyses based on phenotypic and microsatellite data. Histograms on the top and the right side show the distributions for each axis.

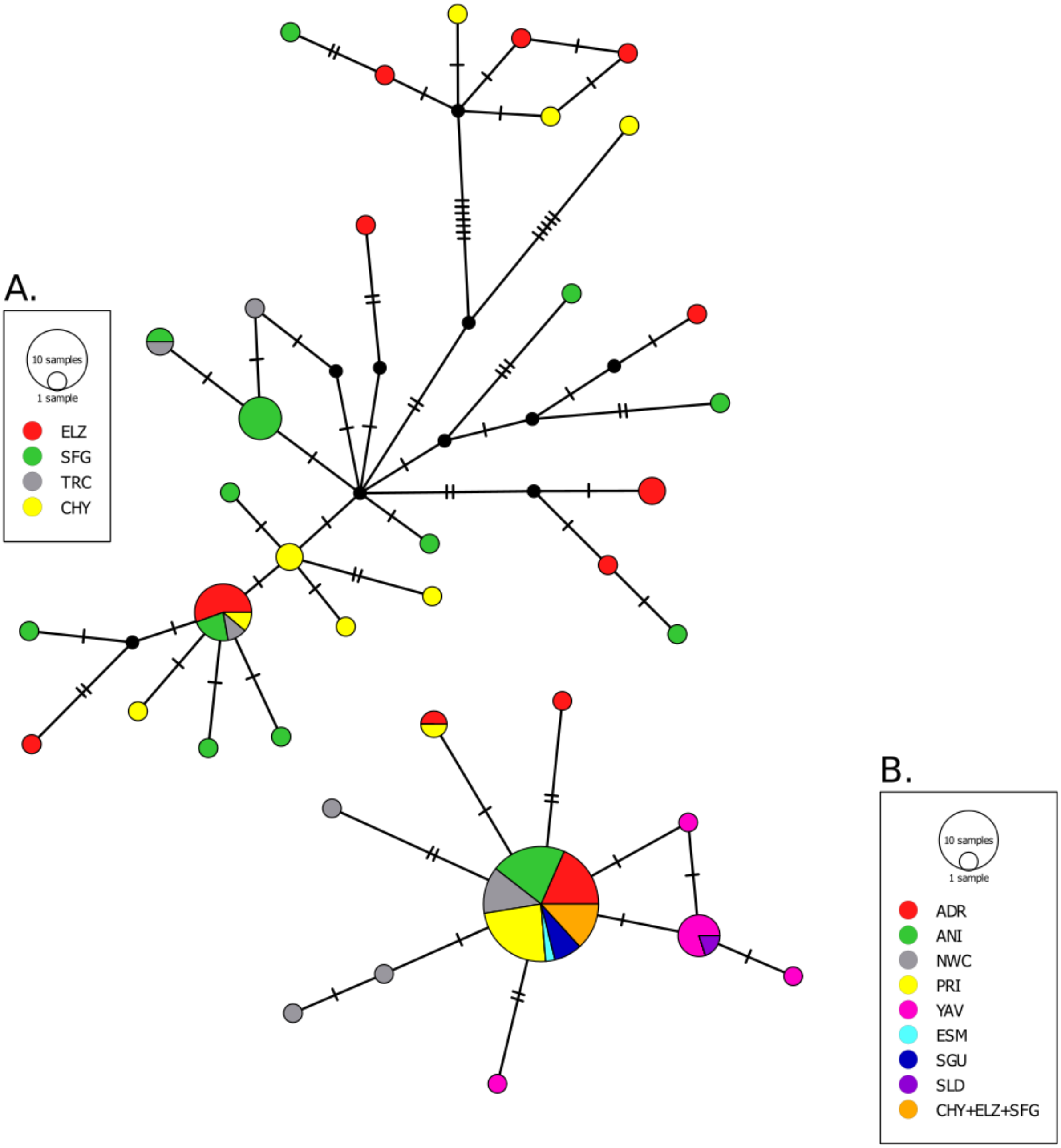

Fig. 6 Cyt *b* median-joining haplotype networks of *C. hubbsi* (A) and *C. regis* (B). Each circle represents a unique haplotype; size of circles indicates number of specimens with that haplotype. Each tick mark represents a nucleotide difference; black circles represent inferred missing haplotypes.

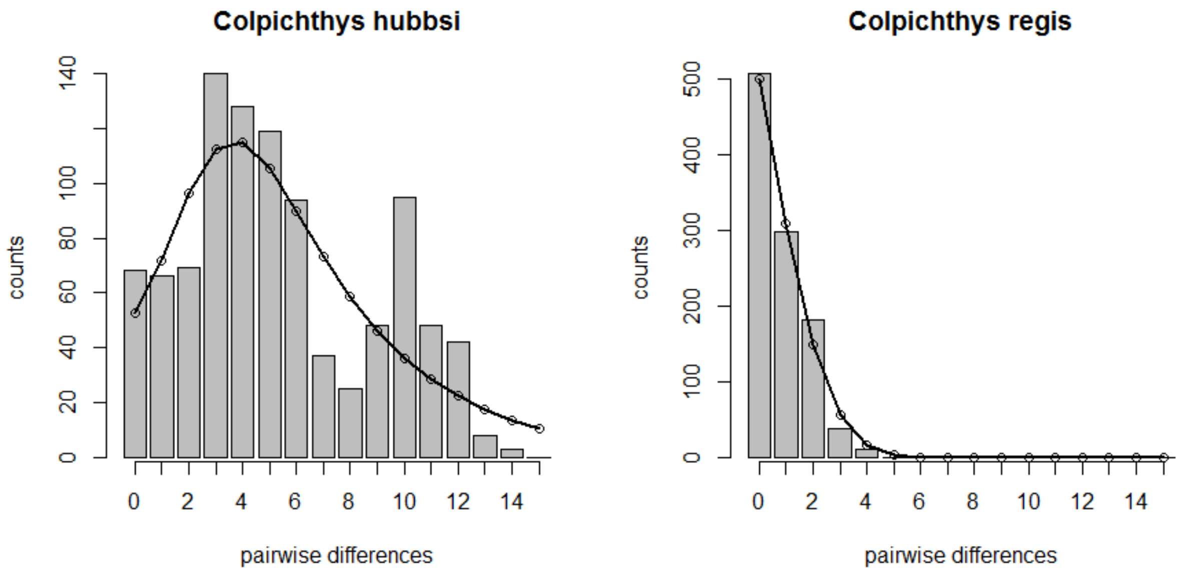

Fig. 7 Mismatch distributions for *C. hubbsi* (Left) and *C. regis* (Right). Histograms represent the observed pairwise differences; black lines represent the expected pairwise differences under the sudden expansion model.

SUPPLEMENTARY MATERIALS (TABLES)

Collection ID	Collection Date	Locality	Locality Code	Coordinates	Species	Feature Examined	N =
SIO 81-158	19 Jan 1973	East of Isla Montague	EIM	31°46.0’ N - 114°42.5’ W	*C. hubbsi*	External morphology	2
SIO 81-161	15 Apr 1973	North of Isla Montague	NIM	31°49.0’ N - 114°48.5’ W	*C. hubbsi*	External morphology	1
SIO 81-156	17 Dec 1972	West of Isla Montague	WIM	31°44.0’ N - 114°48.0’ W	*C. hubbsi*	External morphology	1
SIO 68-46	20 Jan 1968	North of San Felipe	SFE	31°0.0’ N - 114°52.0’ W	*C. hubbsi* *	External morphology	16
SIO 60-484	5 Dec 1960	West from Empalme	EMP	27°55.0’ N - 110°55.0’ W	*C. regis*	External morphology	3
SIO 63-532	17 Aug 1963	Kino Bay	KIN	28°51.5’ N - 112°1.5’ W	*C. regis*	External morphology	1
SIO 63-531	17 Aug 1963	Laguna La Cruz	KNV	28°48.0’ N - 111°54.5’ W	*C. regis*	External morphology	1
LACM 35730	23 May 1974	Guaymas	GYM	27° 55’ N - 110° 57’ W	*C. regis*	Vertebral morphology	3

* Specimens were classified as *C. hubbsi* on label, but all except one specimen from this locality display the external morphology of *C. regis*. (For details, see Table S2.)

Table S1 List of “historical” specimens from the Natural History Museum of Los Angeles County collection (LACM) and Scripps Institute of Oceanography collection (SIO).

Specimen	SL	Snout-D1 to SL ratio	DLS
SIO 68-46 1	117.6	0.56	49
SIO 68-46 2	113.5	0.55	49
SIO 68-46 3	124.3	0.53	48
SIO 68-46 4	138.1	0.55	51
SIO 68-46 5	144.4	0.53	51
SIO 68-46 6	123.8	0.54	51
SIO 68-46 7	126.5	0.54	50
SIO 68-46 8	120.4	0.54	48
SIO 68-46 9	119.8	0.58	48
SIO 68-46 10	115.65	0.54	50
SIO 68-46 11	121.45	0.54	53
SIO 68-46 12	125.9	0.54	51
SIO 68-46 13	116.6	0.55	50
SIO 68-46 14	119.1	0.53	51
SIO 68-46 15	116.4	0.55	49
SIO 68-46 16	**126.7**	**0.60**	**44**

Table S2 Measurements of the external morphology of *C. regis* specimens collected in 1968 at the San Felipe site. **SL** = standard length; **Snout-D1** = length from snout to first dorsal fin origin; **DLS** = dorsolateral scale count. Measurements for specimen 16 (in bold) identify it as *C. hubbsi* sensu Crabtree (1989). Specimens borrowed from the Scripps Institute of Oceanography collection (SIO).

Allele	**Baja Coast**		**Delta Edge**		**Colorado River Delta**				**Sonora Coast**			
Odont08	**ANI (9)**	**ESM (1)**	**PRI (10)**	**SGU (3)**	**TRC (3)**	**CHY (15)**	**ELZ (18)**	**SFG (19)**	**ADR (9)**	**NWC (8)**	**SLD (1)**	**YAV (7)**
146	-	-	-	-	-	0.033	-	-	-	-	-	-
148	-	-	0.1	0.333	0.5	-	-	-	-	0.063	-	-
150	-	0.5	-	-	-	0.233	0.5	0.324	0.333	-	-	0.143
152	-	-	-	-	-	0.133	0.056	0.029	-	-	-	-
154	0.111	-	0.05	-	0.333	0.367	0.028	0.029	-	0.063	-	0.214
156	-	-	0.05	-	-	-	-	0.029	-	-	-	-
158	0.167	0.5	0.25	-	-	0.1	0.056	0.147	0.333	0.438	0.5	0.429
160	0.167	-	0.2	-	-	0.133	0.111	0.206	-	-	-	0.071
162	0.111	-	0.1	0.333	0.167	-	0.111	-	0.111	0.125	-	-
164	0.056	-	0.15	0.333	-	-	0.083	0.059	0.056	0.188	0.5	0.143
166	-	-	0.1	-	-	-	0.028	0.089	0.056	0.125	-	-
168	-	-	-	-	-	-	0.028	0.029	-	-	-	-
170	0.056	-	-	-	-	-	-	-	-	-	-	-
Odont09	**ANI (9)**	**ESM (1)**	**PRI (10)**	**SGU (3)**	**TRC (3)**	**CHY (15)**	**ELZ (18)**	**SFG (19)**	**ADR (9)**	**NWC (8)**	**SLD (1)**	**YAV (7)**
160	-	-	-	-	-	-	-	0.029	-	-	-	-
161	-	-	-	-	-	-	0.028	0.029	-	-	-	-
165	-	-	-	-	-	-	0.056	-	-	-	-	-
171	-	0.5	-	0.167	0.167	-	0.168	0.029	-	0.125	-	0.071
173	1	0.5	0.95	0.833	0.833	0.867	0.667	0.824	1	0.875	1	0.929
175	-	-	0.05	-	-	0.133	0.056	0.059	-	-	-	-
179	-	-	-	-	-	-	0.028	0.029	-	-	-	-
Odont11	**ANI (9)**	**ESM (1)**	**PRI (10)**	**SGU (3)**	**TRC (3)**	**CHY (15)**	**ELZ (18)**	**SFG (19)**	**ADR (9)**	**NWC (8)**	**SLD (1)**	**YAV (7)**
154	-	-	0.05	-	0.167	0.033	0.028	0.059	0.056	-	-	-
156	1	1	0.9	0.833	0.667	0.767	0.806	0.853	0.889	1	1	1
158	-	-	0.05	0.167	0.167	0.2	0.167	0.089	0.056	-	-	-
B18	**ANI (9)**	**ESM (1)**	**PRI (10)**	**SGU (3)**	**TRC (3)**	**CHY (15)**	**ELZ (18)**	**SFG (19)**	**ADR (9)**	**NWC (8)**	**SLD (1)**	**YAV (7)**
275	-	-	-	-	-	0.033	-	-	0.056	0.063	-	-
277	-	-	-	-	-	-	-	0.088	0.056	-	-	-
279	-	-	-	-	-	-	0.083	0.059	-	-	-	-
281	0.222	-	0.25	0.167	-	0.1	0.056	0.059	0.222	-	-	0.357
283	0.278	0.5	0.2	-	0.333	0.033	0.028	0.059	0.056	0.063	-	0.071
285	-	-	0.05	-	0.167	0.1	-	0.059	0.167	0.25	-	0.071
287	-	-	0.05	0.167	-	0.033	0.139	0.118	-	0.125	0.5	0.286
289	0.5	0.5	0.1	0.167	-	0.067	0.194	0.059	0.167	0.25	-	0.143
291	-	-	0.05	-	0.167	0.067	0.083	0.176	-	-	0.5	-
293	-	-	0.1	-	-	0.1	0.028	0.176	0.056	0.125	-	-
295	-	-	0.1	0.167	-	0.1	0.028	0.029	0.167	0.125	-	-
297	-	-	-	-	-	0.033	0.056	-	-	-	-	0.071
299	-	-	0.1	0.167	-	0.067	0.028	-	-	-	-	-
302	-	-	-	-	-	-	0.028	0.029	-	-	-	-
304	-	-	-	-	-	0.1	0.056	-	-	-	-	-
306	-	-	-	0.167	0.167	0.033	0.056	0.059	-	-	-	-
310	-	-	-	-	-	0.033	-	-	-	-	-	-
313	-	-	-	-	-	-	-	-	0.056	-	-	-
314	-	-	-	-	-	-	-	0.029	-	-	-	-
316	-	-	-	-	0.167	0.067	-	-	-	-	-	-
318	-	-	-	-	-	0.033	0.056	-	-	-	-	-
320	-	-	-	-	-	-	0.028	-	-	-	-	-
331	-	-	-	-	-	-	0.028	-	-	-	-	-
333	-	-	-	-	-	-	0.028	-	-	-	-	-
B19	**ANI (9)**	**ESM (1)**	**PRI (10)**	**SGU (3)**	**TRC (3)**	**CHY (15)**	**ELZ (18)**	**SFG (19)**	**ADR (9)**	**NWC (8)**	**SLD (1)**	**YAV (7)**
224	-	-	-	0.333	-	0.2	0.222	0.235	--	-	-	-
226	-	-	-	-	-	0.033	0.028	0.029	0.056	-	-	-
228	-	-	0.3	0.167	0.667	0.267	0.222	0.324	0.167	0.063	-	-
230	-	-	-	0.167	0.167	0.133	0.25	0.147	-	-	-	-
232	-	0.5	0.3	-	0.167	0.3	0.222	0.235	0.333	0.125	-	0.143
234	0.889	0.5	0.4	0.167	-	0.067	0.056	0.029	0.278	0.688	1	0.571
236	-	-	-	0.167	-	-	-	-	-	-	-	0.143
238	-	-	-	-	-	-	-	-	0.056	0.063	-	-
240	-	-	-	-	-	-	-	-	-	0.063	-	0.143

Table continued on next page.

Table continued from previous page.

B39	ANI (9)	ESM (1)	PRI (10)	SGU (3)	TRC (3)	CHY (15)	ELZ (18)	SFG (19)	ADR (9)	NWC (8)	SLD (1)	YAV (7)
252	1	1	0.75	0.667	0.333	0.033	0.028	0.029	1	1	1	1
254	-	-	0.1	0.333	0.667	0.733	0.528	0.618	-	-	-	-
256	-	-	-	-	-	0.133	0.139	0.176	-	-	-	-
258	-	-	0.05	-	-	-	0.139	0.29	-	-	-	-
262	-	-	-	-	-	0.033	-	-	-	-	-	-
264	-	-	0.05	-	-	-	0.0833	0.059	-	-	-	-
266	-	-	-	-	-	-	0.056	0.059	-	-	-	-
268	-	-	0.05	-	-	0.033	0.028	-	-	-	-	-
271	-	-	-	-	-	0.033	-	0.029	-	-	-	-

Table S3 Frequencies of each microsatellite alleles at each sampling locale. Microsatellite alleles are arranged by size. Number in parentheses under each locale denotes number of individuals analyzed.

SUPPLEMENTARY MATERIALS (FIGURES)

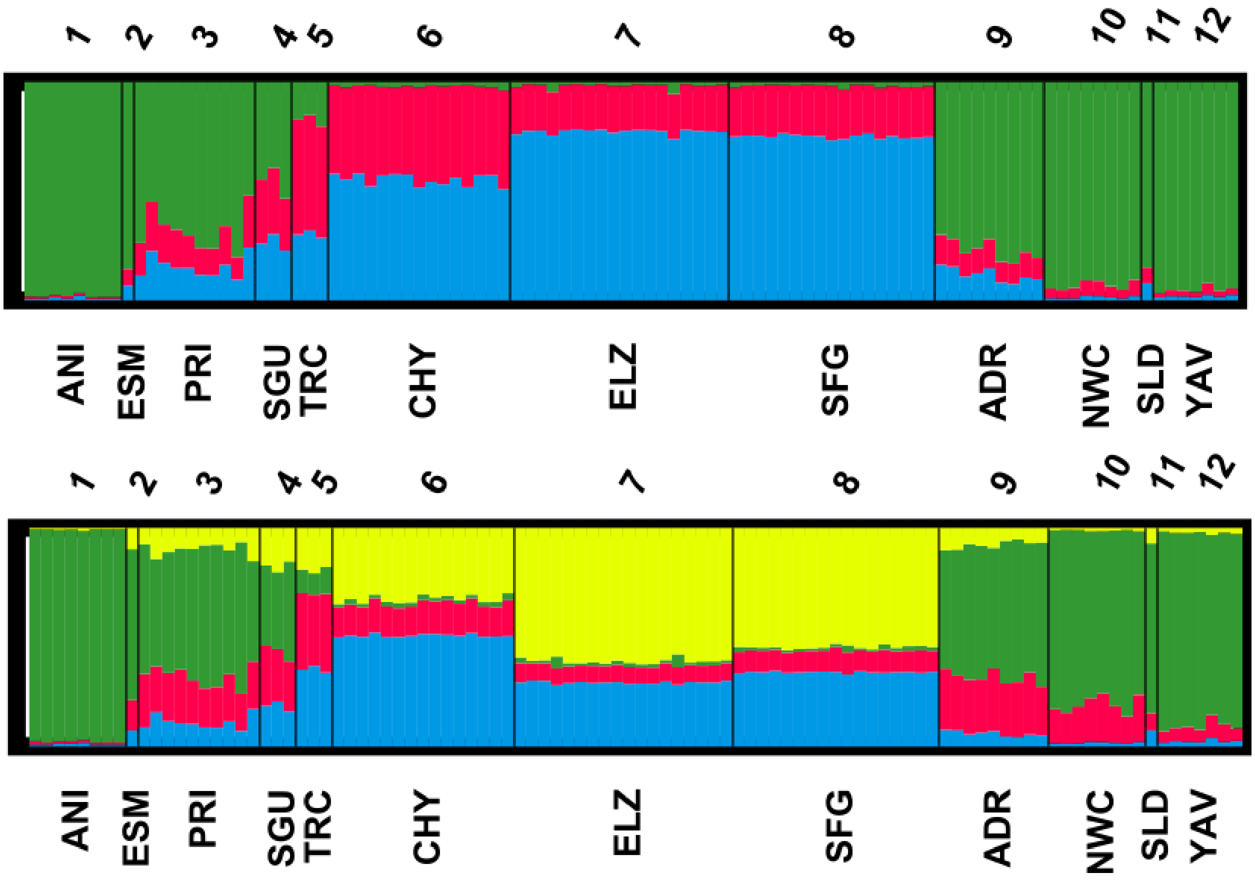

Fig. S1 Structure plots for the correlated allele frequency model for K=3 (top) and K=4 (bottom).

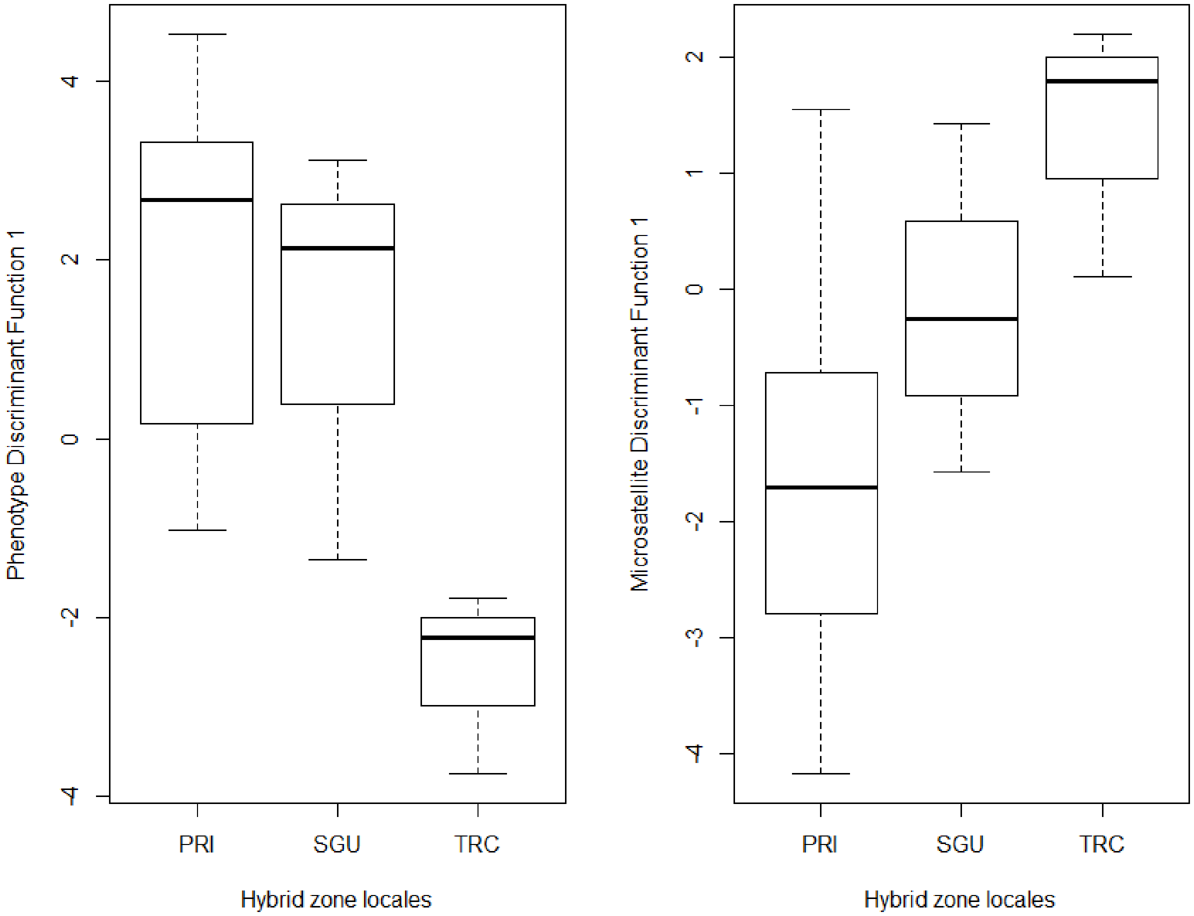

Fig. S2 Boxplots of the DAPC coordinates for specimens in the hybrid zone.

REFERENCES

Akihito, Iwata, A., Kobayashi, T., Ikeo, K., Imanishi, T., Ono, H., Umehara, Y., Hamamatsu, C. Sugiyama, K., Ikeda, Y., Sakamoto, K., Fumihito, A., Ohno, S. & Gojobori, T. (2000). Evolutionary aspects of gobioid fishes based upon a phylogenetic analysis of mitochondrial cytochrome b genes. *Gene, 259*(1), 5-15.

Beheregaray, L.B. & Sunnucks, P. (2000). Microsatellite loci isolated from *Odontesthes argentinensis* and the *O. perugiae* species group and their use in other South American silverside fish. *Molecular Ecology*, **9**, 629–631.

Boutin-Ganache, I., Raposo, M., Raymond, M. & Deschepper, C.F. (2001). M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. *Biotechniques*, **31**, 24–6.

Brusca, R.C. (1980). Common intertidal invertebrates of the Gulf of California. Univ of Arizona Pr.

Brusca, R. C., Álvarez-Borrego, S., Hastings, P. A., & Findley, L. T. (2017). Colorado River flow and biological productivity in the Northern Gulf of California, Mexico. *Earth-Science Reviews*, **164**, 1-30.

Byrne, R.J. & Avise, J.C. (2009). Multiple paternity and extra-group fertilizations in a natural population of California grunion (*Leuresthes tenuis*), a beach-spawning marine fish. *Marine Biology*, **156**, 1681–1690.

Castro-Aguirre, J.L. & Espinosa Pérez, H. (2006). Los peces de la familia Atherinopsidae (teleostei: Atheriniformes) de las lagunas costeras neutras e hipersalinas de México. *Hidrobiológica*, **16**, 89–101.

Crabtree, C.B. (1989). A New Silverside of the Genus *Colpichthys* (Atheriniformes: Atherinidae) from the Gulf of California, Mexico. *Copeia*, **1989**, 558–568.

Crossey, L.C., Karlstrom, K.E., Dorsey, R., Pearce, J., Wan, E., Beard, L.S., Asmerom, Y., Polyak, V., Crow, R.S., Cohen, A., Bright, J. & Pecha, M.E. (2015). Importance of groundwater in propagating downward integration of the 6-5 Ma Colorado River system: Geochemistry of springs, travertines, and lacustrine carbonates of the Grand Canyon region over the past 12 Ma. *Geosphere*, **11**, 660–682.

Dolby, G.A., Hechinger, R., Ellingson, R.A., Findley, L.T., Lorda, J. & Jacobs, D.K. (2016). Sea-level driven glacial-age refugia and post-glacial mixing on subtropical coasts, a palaeohabitat and genetic study. *Proceedings of the Royal Society of London B: Biological Sciences*, **283**.

Dorsey, R.J., Fluette, A., McDougall, K., Housen, B.A., Janecke, S.U., Axen, G.J. & Shirvell, C.R. (2007). Chronology of Miocene–Pliocene deposits at Split Mountain Gorge, Southern California: A record of regional tectonics and Colorado River evolution. *Geology*, **35**, 57.

Drummond, A., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M., Markowitz, S. & others. (2011). Geneious pro v5. 5.6. *available from http://www.geneious.com/*.

Earl, D.A. & others. (2012). STRUCTURE harvester: A website and program for visualizing structure output and implementing the evanno method. *Conservation genetics resources*, **4**, 359–361.

Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular ecology*, **14**, 2611–2620.

Excoffier, L., Laval, G. & Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary bioinformatics*, **1**.

Falush, D., Stephens, M. & Pritchard, J.K. (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics*, **164**.

Findley, L., Collette, B. & Espinosa, H. (2010a). *Colpichthys hubbsi*. *The IUCN Red List of Threatened Species 2010: e.T183457A8116697*.

Findley, L., Collette, B. & Espinosa, H. (2010b). *Colpichthys regis*. *The IUCN Red List of Threatened Species 2010: e.T183357A8099025*.

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Fu, Y.-X. & Li, W.-H. (1993). Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Glenn, E.P., Lee, C., Felger, R. & Zengel, S. (1996). Effects of Water Management on the Wetlands of the Colorado River Delta , Mexico. **10**, 1175–1186.

Harrison, R.G. & Larson, E.L. (2014). Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, **105**, 795–809.

Heiser, C.B. (1973). Introgression re-examined. *The Botanical Review*, **39**, 347–366.

Howard, K.A., House, P.K., Dorsey, R.J. & Pearthree, P.A. (2015). River-evolution and tectonic implications of a major Pliocene aggradation on the lower Colorado river: The bullhead alluvium. *Geosphere*, **11**, 1–30.

Hubisz, M.J., Falush, D., Stephens, M. & Pritchard, J.K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular ecology resources*, **9**, 1322–32.

Jacoby Jr, G.C. (1976). Long-term surface-water supply and streamflow trends in the upper colorado river basin based on tree-ring analyses. *Natl. Sci. Found. Lake Powell Res. Proj. Bull*, **18**, 70.

Jakobsson, M. & Rosenberg, N.A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jombart, T. & Ahmed, I. (2011). Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, **27**, 3070–3071.

Kumar, S., Stecher, G. & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*, **33-7**, 1870-1874. doi: 10.1093/molbev/msw054

Lavin, M. & Sánchez, S. (1999). On how the Colorado river affected the hydrography of the upper Gulf of California. *Continental Shelf Research*, **19**, 1545–1560.

Leigh, J.W. & Bryant, D. (2015). Popart: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, **6**, 1110–1116.

Librado, P. & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Littrell, B.M., Lutz-Carrillo, D.J., Bonner, T.H. & Fries, L.T. (2007). Status of an introgressed Guadalupe bass population in a central Texas stream. *North American Journal of Fisheries Management*, **27**, 785–791.

López, J.A., Chen, W.-J. & Ortí, G. (2004). Esociform phylogeny. *Copeia*, **3**, 449–464.

Martin, N.H., Bouck, A.C. & Arnold, M.L. (2006). Detecting Adaptive Trait Introgression Between *Iris fulva* and *I. brevicaulis* in Highly Selective Field Conditions. *Genetics,* **172**, 2481–2489.

McDougall, K. & Martínez, A.Y.M. (2014). Evidence for a marine incursion along the lower Colorado River corridor. *Geosphere*, **10**, 842–869.

Olden, J.D., LeRoy Poff, N., Douglas, M.R., Douglas, M.E. & Fausch, K.D. (2004). Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology & Evolution*, **19**, 18–24.

Palacios-Salgado, D.S., Burnes-Romo, L.A., Tavera, J.J. & Ramírez-Valdez, A. (2012). Endemic fishes of the Cortez biogeographic province (Eastern Pacific Ocean). *Acta Ichthyologica et Piscatoria*, **42**, 153–164.

Pardo-Diaz, C., Salazar, C., Baxter, S.W., Merot, C., Figueiredo-Ready, W., Joron, M., McMillan, W.O. & Jiggins, C.D. (2012). Adaptive Introgression across Species Boundaries in *Heliconius* Butterflies (M. R. Kronforst, Ed.). *PLoS Genetics*, **8**, e1002752.

Poulson, S.R. & John, B.E. (2003). Stable isotope and trace element geochemistry of the basal Bouse Formation carbonate, southwestern United States: Implications for the Pliocene uplift history of the Colorado Plateau. *Bulletin of the Geological Society of America*, **115**, 434–444.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Ptak, S.E. & Przeworski, M. (2002). Evidence for population growth in humans is confounded by fine-scale population structure. *Trends in Genetics*, **18**, 559–563.

Ramos-onsins, S.E. & Rozas, J. (2006). Statistical Properties of New Neutrality Tests Against Population Growth. *Molecular Biology and Evolution*, **23**, 1642–1642.

Rodriguez, C., Flessa, K. & Dettman, D. (2001). Effects of upstream diversion of Colorado River water on the estuarine bivalve mollusc *Mulinia coloradoensis*. *Conservation Biology,* **15**, 249–258.

Rogers, A.R. & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Rosenberg, N.A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes,* **4**, 137–138.

Roskowski, J.A., Patchett, P.J., Spencer, J.E., Pearthree, P.A., Dettman, D.L., Faulds, J.E. & Reynolds, A.C. (2010). A late Miocene-early Pliocene chain of lakes fed by the Colorado River: Evidence from Sr, C, and O isotopes of the Bouse Formation and related units between Grand Canyon and the Gulf of California. *Geological Society of America Bulletin,* **122**, 1625–1636.

Silber, G.K. (1990). Occurrence and distribution of the vaquita *Phocoena sinus* in the northern Gulf of California. *Fishery Bulletin*, **88**, 339–346.

Spencer, J.E. & Patchett, P.J. (1997). Sr isotope evidence for a lacustrine origin for the upper Miocene to Pliocene Bouse Formation, lower Colorado River trough, and implications for timing of Colorado Plateau uplift. *Geological Society of America Bulletin,* **109**, 767–778.

Spencer, J.E., Patchett, P.J., Pearthree, P.A., House, P.K., Sarna-Wojcicki, A.M., Wan, E., Roskowski, J.A. & Faulds, J.E. (2013). Review and analysis of the age and origin of the Pliocene Bouse Formation, lower Colorado River Valley, southwestern USA. *Geosphere*, **9**, 444–459.

Swift, C.C., Findley, L.T., Ellingson, R.A., Flessa, K.W. & Jacobs, D.K. (2011). The Delta Mudsucker, *Gillichthys detrusus*, a valid species (teleostei: Gobiidae) endemic to the Colorado River Delta, northernmost Gulf of California, México. *Copeia*, **2011**, 93–102.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tamura, K. & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution*, **10**, 512–526.

Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A. & Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proceedings of the National Academy of Sciences*, **109**, 19333–19338.

Todd, T.N. (1976). Pliocene occurrence of the recent Atherinid fish *Colpichthys regis* in Arizona. *Journal of Paleontology*, **50(3)**, 462–466.

Vähä, J.P. & Primmer, C.R. (2006). Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, **15**, 63–72.

Vonlanthen, P., Bittner, D., Hudson, A.G., Young, K.A., Müller, R., Lundsgaard-Hansen, B., Roy, D., Di Piazza, S., Largiader, C.R. & Seehausen, O. (2012). Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature*, **482**, 357–362.

Whitney, K.D., Randell, R.A. & Rieseberg, L.H. (2010). Adaptive introgression of abiotic tolerance traits in the sunflower *Helianthus annuus*. *New Phytologist*, **187**, 230–239.

Winker, C.D. & Kidwell, S.M. (1986). Paleocurrent evidence for lateral displacement of the Pliocene Colorado River delta by the San Andreas fault system, southeastern California. *Geology*, **14**, 788.

Wirtz, P. (1999). Mother species–father species: unidirectional hybridization in animals with female choice. *Animal Behaviour*, **58**, 1–12.

Wu, D.Y., Ugozzoli, L., Pal, B.K. & Wallace, R.B. (1989). Allele-specific enzymatic amplification of beta-globin genomic DNA for diagnosis of sickle cell anemia. *Proceedings of the National Academy of Sciences of the United States of America*, **86**, 2757–2760.