Lack of genetic differentiation between two sympatric species of Astyanax (Characidae:Teleostei) in Lake Catemaco, Mexico

Patricia Ornelas
Universidad Nacional Autonoma de Mexico https://orcid.org/0000-0003-3163-448X
Elena G. Gonzalez
Museo Nacional de Ciencias Naturales
Diethard Tautz
Max-Planck-Gesellschaft
Ignacio Doadrio (doadrio@mncn.csic.es)
https://orcid.org/0000-0003-4863-9711

Research article

Keywords: speciation, morphological divergence, lacustrine diversification, Astyanax, genetic differentiation

Posted Date: February 7th, 2020

DOI: https://doi.org/10.21203/rs.2.22839/v1

Abstract

Background

Fish of the genus Astyanax are known to be able to adapt to a wide range of ecological conditions and are especially known for repeated colorizations of cave systems. In lakes they often occur in species pairs. In the case we study here, they show major morphological differences, such that they were originally classified in different genera. Previous studies have shown that these morphological differences correlate with occupation of different trophic niches. Hence, this could be an example of ecological speciation under sympatric conditions, which predicts that differential ecological adaptation becomes coupled to assortative mating and to the formation of genetically distinct groups that may be called species. We have tested this prediction by typing a set of microsatellites for the two morphs in the lake in comparison to an allopatric population.

Results

While we find the expected differentiation with respect to the allopatric population, there is a complete lack of genetic differentiation between the two morphs within the lake. Hence, the two morphs in the lake are either in an extremely early phase of speciation or represent two extreme morphotypes derived from a single gene pool.

Conclusions

Even when we failed to recover the two morphs as reproductively isolated, this model provides a unique opportunity to characterize those factors that would promote the ecological divergence, thus, our lacustrine morphs system gives a unique opportunity to understand the genetic basis of how morphological divergence in the presence of gene flow.

Background

The origin and maintenance of biodiversity remains as one of the most challenging questions in biology. Ecological divergence has been recognized as a force promoting diversity and speciation (Schluter 1996). Under this scenario, adaptive evolution related to different ecological conditions promotes the emergence of alternative phenotypes that could differ in fitness, and those differences could promote genetic changes (West-Eberhard 2003; Dieckmann *et al.* 2004; Nosil 2012).

Ecological speciation occurs when adaptation to different environments or resources causes the evolution of reproductive isolation (Doebeli & Dieckmann 2003), but many cases of divergent selection and adaptive divergence are associated with only weak to modest levels of reproductive isolation (Thibert-Plante & Hendry 2009). Ecology-driven speciation is thought as a gradual process in which different measures of divergence can be used to assess the 'stage' of speciation (Nosil *et al.* 2009). Without the need to be ordered, such stages generally include the emergence of phenotypic, ecological, behavioral, and genetic divergence, however at the later stages of the speciation process we would expect a bimodal genotypic clustering (see Nosil, 2012; Doebeli & Dieckmann 2003).

In the continuum of the speciation process, the reproductive isolation could be from absent through complete, as are the examples of *Timema* walking-stick insects (absent) to the *Citrinellus* cichlids (complete) (Barluenga *et al.* 2006; Nosil *et al.* 2009). Ecological speciation applies independently of whether reproductive isolation evolves in sympatry, parapatry, or allopatry but it is particularly relevant under sympatric conditions, such as in small lakes (Hatfield & Schluter 1999; Schliewen *et al.* 2001).

The study of populations with discontinuous phenotypic variation (*i.e.,*, polymorphic populations) experiencing different degrees of reproductive isolation can provide valuable insights to determine which factors drive ecological divergence and ultimately lead to speciation (Magalhaes *et al.* 2015). Speciation of morphs occurs when fixation of a single morph is coupled with reproductive isolation. On the other hand, coexistence of polymorphic species with gene flow provide evidence of mechanisms of phenotypic divergence as either, alternative to, or part of the sequence on a speciation process (Schliewen *et al.* 2001; Magalhaes *et al.* 2015). A few studies have shown a stable coexistence of morphs within a species, suggesting alternative mechanisms to speciation, and explaining their coexistence (Hori 1993).

The *Astyanax* genus is characterized by its ability to adapt to a wide range of ecological conditions (Ornelas-García *et al.* 2008). The genus includes some striking examples of morphological divergence associated to extreme habitats, such as those reported in the cavefish *Astyanax mexicanus* (e.g. (Protas *et al.* 2007; Gross *et al.* 2009; Jeffery 2009; Elipot *et al.* 2014; McGaugh *et al.* 2014)). Another example of divergence related to ecological conditions has been described from several species pairs inhabiting lakes (Ornelas-García *et al.* 2008; Ornelas-García *et al.* 2014; Ornelas-García *et al.* 2018). In the present study, we focus on two closely related species: *Astyanax aeneus* and

A. caballeroi, that coexist in the tropical Lake Catemaco in Mexico. They show conspicuous morphological differences, which are mainly related to trophic habits: tooth shape, dental formula, eye size, snout length, body depth, head profile and mouth orientation (Contreras-Balderas & Rivera-Teillery 1983; Ornelas-García *et al.* 2008; Ornelas-García *et al.* 2014). The morphological divergence between the two species was so conspicuous that originally they were considered as different genera, namely *Bramocharax caballeroi* as a species described to be endemic to Lake Catemaco, and *Astyanax aeneus* described to show a wider distribution range (Ornelas-García *et al.* 2008). Recently, we could corroborate that morphological differentiation between the two morpho-species present in Lake Catemaco is associated with an ecological segregation. Isotope ratios and diet showed that *A. aeneus* and *A. caballeroi,* represent different trophic habits, suggesting an ecological niche partitioning between the two currently recognized species (Ornelas-García *et al.* 2018).

Little is known about the genetic basis of these ecologically divergent species in the Lake Catemaco, however previous phylogenetic and mitochondrial analyses have shown no differentiation between them, including the sharing of mitochondrial haplotypes (Ornelas-García *et al.* 2008; Ornelas-García *et al.* 2014). Here we evaluate the genetic structure as well as the gene flow between the two morpho-species based on a microsatellite analysis. Surprisingly, we find no signs of any genetic differentiation within the lake, even with the most sensitive analysis procedures. We suggest that this system is at its earliest possible type of ecological speciation, where the phenotypic differences are driven by only a small number of loci and where the phenotypic differentiation has not yet led to genetic differentiation.

Methods

Sample Collection and DNA Extraction

A total of 318 individuals assigned to both nominal species were collected during 2006 at 13 locations in the Lake Catemaco, Mexico (fig. 1, table S1). In the following, we retain the original genus names ('Astyanax' and 'Bramocharax') to assign the morphs, since they were collected according to these original diagnostic characters (Contreras-Balderas & Rivera-Teillery 1983). One allopatric population of *Astyanax aeneus* from Maquinas river basin, near to Lake Catemaco, was included in order to test if the microsatellite loci used were able to differentiate between allopatric populations. Fish were euthanized putting them in ice water (approximately 4ºC), and were collected under permits SGPA/DGVS/05464/ SAGARPA (México). A small fin clip sample was taken from all individuals, the remainder was preserved in ethanol as vouchers for future morphological analyses. The fin clips were also preserved in 90% ethanol and subsequently stored at −20 °C. DNA was extracted using a standard Proteinase-K in SDS/EDTA digestion and NaCl (4.5 M) and Chloroform, as described in (Sonnenberg *et al.* 2007). Voucher samples were stored at the Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain.

Microsatellite loci were taken from a previous study in a closely-related species, *Astyanax mexicanus* (Protas *et al.* 2006). Among the loci described therein, 12 were chosen considering their location on different chromosomes and absence of linkage to reported traits. The loci were multiplexed in two reactions using Multiplex PCR kit (QIAGEN), in 5µL reaction final volumes following kit instructions. PCR amplifications consisted of 1 cycle of denaturing at 95 °C for 5 min; 35 cycles of 94 °C for 30 s, annealing for 30 s at 52–60 °C and extension at 72 °C for 45 s; followed by 1 cycle of 7-min extension at 72 °C (table S2). Forward primers were labeled with fluorescent dyes (Invitrogen), and amplified PCR products were run on an ABI Prism 3730 DNA Analyzer (GS500 ROX size standard). All loci tested were successfully amplified after PCR optimization. Of those, all were polymorphic and were used in full screening of the rest of all samples (total N = 348; table S2) following the protocol described previously. Allele scoring was performed using Geneious v 10.2.3.

Genetic Diversity Analyses

Microsatellite genetic diversity was quantified for locus and sampling site based on the average number of alleles per locus (NA)„ number of alleles standardized to those of the population sample with the smallest size (NS (Nei & Chesser 1983)), and the observed (Ho) and expected (He) heterozygosities (Nei 1978), using Genetix 4.05 (Belkhir 2004) and Fstat 2.9.3 (Goudet 2001). Deviations from Hardy–Weinberg (HW) proportions were tested using the exact probability test for multiple alleles (Guo & Thompson 1992) available in Genepop 4.0.1 (Raymond 1995) at each locus for each population and over all loci for each population. Genotypic linkage disequilibrium between each pair of loci was estimated by Fisher's Exact tests with Genepop 4.0.1 software.

Both tests for deviations from HW proportions and for linkage disequilibrium used a Markov chain (10,000 dememorization steps, 1000 batches, 2000 iterations per batch) (Guo & Thompson 1992). Correction for multiple testing (type I error rates) was performed using the sequential Bonferroni (Rice, 1989). Additionally, MICRO- CHECKER v2.23 (Van Oosterhout *et al.* 2004) was used to explore the existence of null alleles and to evaluate their impact on the estimation of genetic differentiation.

Population Structure and Gene Flow Analyses

To visualize the relationship between population samples and morphs, a correspondence analysis (Guinand 1996) was implemented in GENETIX v.4.05.2 (Belkhir *et al.* 2000). A Bayesian clustering method was used to assess the possibility of genetic structure between both morphs. The number of populations (*K)* with the highest posterior probability (mean lnProb [D]) was calculated with the program STRUCTURE 2.0 (Pritchard *et al.* 2000), assuming an admixed model and a uniform prior probability of the number of populations, *K.* MCMC simulations consisted of 5×10^6 burn-in iterations followed by 5×10^5 sampled iterations. Furthermore, the modal value of λ, *ΔK (*Evanno *et al.* 2005) was also calculated to infer the best value of *K.* Ten runs for each value of *K* were conducted to check for consistency in the results.

The F_{ST} pairwise comparisons were carried out between sympatric morphs (Astyanax *vs.* Bramocharax), and each of them *vs.* Maquinas river population, the F_{ST} P-values were estimated based on 120 permutations in Arlequin, ver 3.5. (Excoffier & Lischer 2010). To determine the amount of genetic structuring among grouping levels, an analysis of molecular variance (AMOVA) was performed with Arlequin, ver 3.5. (Excoffier & Lischer 2010). AMOVA was used to assess the partitioning of microsatellite variation in allelic frequencies ($F_{ST)}$ between and within different grouping schemes according to morpho-species (Astyanax vs Bramocharax) and the allopatric *Astyanax aeneus* from Maquinas river.

Migrate 3.2.6 (Beerli & Felsenstein 1999, 2001) was used to infer the population size parameter Θ (where $\Theta = 4N_e\mu$, *Ne* is the effective population size and μ is the mutation rate per site) and the migration rate *M* (where $M = m/\mu$ and *m* is the immigration rate per generation) for both Astyanax and Bramocharax morphs, separated by lineages 1 and 2. Based on the estimates of historical migration rates (*M)* obtained, we tested different models of gene flow between species and lineages, using a Bayes factor approach (Beerli & Palczewski 2010). For the two species, we evaluated the following models: 1) Panmictic model with one population size (Astyanax and Bramocharax), 2) Two population sizes and one migration rate from Astyanax morpho-species to the Bramocharax morpho-species, and 3) Two population sizes and one migration rate from Bramocharax morph to the Astyanax morph. The models were compared using Bézier thermodynamic integration (Beerli & Palczewski 2010), and their marginal likelihoods were then used to calculate Bayes factors and model probabilities (Kass & Raftery 1995). We assumed a Brownian-motion mutation model with constant mutation across all loci, and F_{ST} and a UPGMA tree were the starting parameters for the estimation of Θ and *M.* MCMC consisted of ten short chain samplings (of 50,000 trees) and three long chain samplings (of 500,000 trees) using an adaptive heating scheme.

Results

All microsatellite loci were polymorphic in all locations and showed different levels of polymorphism across sites. MICRO-CHECKER inferred the presence of null alleles at two loci across sampling localities. Corrected estimates of F_{ST} by null allele presence were very similar to noncorrected values, consequently, all loci were used in the analyses. No linkage disequilibrium was observed between any loci pairs, which suggest independence of all loci (table S3). A total of 273 alleles were obtained, ranging from 13 (NYU14) to 32 (NYU29) per locus (mean *Na* = 22.21) (table S4), the number of alleles between the two morphs was very similar, with a mean for the Astyanax morph of 20.3 and for the Bramocharax morph of 18.58. The allopatric *Astyanax aeneus* from Maquinas river showed a considerably lower number of alleles (*Na* = 3.97). The amount of genetic variability was relatively high, across loci and morph clusters (H_O ranging from 0.847 to 0.855) (table 1). Similarly, the Maquinas river population showed the lowest *Ho* and *He* (0.60 and 0.57 respectively).

Lack of genetic structure among sympatric morpho-species

A Factorial Correspondence Analysis (FCA) clearly separated the population from Maquinas river from those of Lake Catemaco, but the Astyanax and Bramocharax morphs were not separated (fig. 2A). The Bayesian clustering analysis with STRUCTURE, including the Maquinas river population, revealed the highest likelihood for the model with K = 2, in congruence with the value for the highest ΔK for K = 2 (fig. 2B). Under this scenario, the two groups recovered correspond to the Maquinas river *vs.* Lake Catemaco. The latter was recovered as one genetic cluster, thus, showing the two morphs from Lake Catemaco as a panmictic unit. No additional peaks in the ΔK were evident (Figure 2C, D).

The F_{ST} pairwise comparisons were congruent with the previous results, showing no genetic differentiation between sympatric morphs (F_{ST} = 0.0003, P>0.05). In contrast, allopatric populations showed highly significant F_{ST} values (P<0.005), divergences between allopatric populations were higher, 0.202 Astyanax Lake Catemaco *vs.* Maquinas river, and 0.200 Bramocharax from Lake Catemaco and Maquinas

river. Similarly, in the hierarchical AMOVA our results support a lack of differentiation between sympatric morphs from Lake Catemaco (Astyanax *vs.* Bramocharax), only being significant the differences between geographic groups (i.e. Lake Catemaco vs. Maquinas river), and populations among groups (table S5).

Migration patterns between the two morpho-species

Migration rates and effective population size were estimated between differentAstyanax and Bramocharax morphs. Both analyses, ML and BI, were concordant in the best model which correspond to the panmictic unit (Ln ML –383825.44 and Ln BF = –247769.74 respectively, table 2).

Discussion

In the present study we evaluate the degree of genetic differentiation (*i.e.* effective reproductive isolation) between two morphologically well differentiated morphs of lacustrine fish, which were previously even classified as different genera, *Astyanax vs. Bramocharax* (Contreras-Balderas & Rivera-Teillery 1983). Their morphological divergence was shown to be related to different trophic habitats (Ornelas-García *et al.* 2014; Ornelas-García *et al.* 2018), indicating an adaptive divergence possibly in the context of ecological speciation. But contrary to the expectations that this should be coupled with at least some degree of reproductive isolation, our results show a complete lack of measurable genetic differentiation. At the same time, we find clear genetic differences with respect to an allopatric population (i.e. Maquinas river *vs* Lake Catemaco), supporting the notion that the marker system employed (microsatellites) would have captured genetic differences between the morphs, if they existed.

Our results show higher genetic diversity in the Lake Catemaco for both morphs than in the Maquinas river population (table 1). Heterozygosities in both *Ho* and *He,* were similar to those reported to other characid species (Beheregaray *et al.* 2005), and to the previous studies for the genus in a sister species (Bradic *et al.* 2012). This suggests that the morphs have been stably established since some time without any recent bottlenecks. The complete lack of any overall genetic differentiation between the morphs could be due to either an incipient divergence process or a polymorphic nature of a single species.

Some studies have highlighted that incipient divergence during a speciation process would be difficult to discriminate by using a restricted number of markers by genotypic cluster criterion (Coyne & Orr 1998). Moreover, under an ecological divergence framework, reproductive isolation could be time-dependent, that is, recently diverged populations would show weaker signs of reproductive isolation (and hence, genetic differentiation) than older systems (Nosil 2012).

In the present study, we were expecting to recover two discrete genetic groups, in accordance to the genetic cluster species concept (Mallet 1995). This concept is very useful for incipient species in parapatry or under sympatric conditions (Coyne & Orr 1998). However, as previous authors have shown, there is a problem related to the question of how much genetic differentiation is required to recognize a species, since we could recognize them with one or some hundreds of loci that could underly divergent adaptation, reduce gene flow, cause reproductive isolation or reduce hybrid fitness (Presgraves 2003);(Coyne & Orr 1998; Panova *et al.* 2006; Nosil & Schluter 2011). Another problem with the genetic cluster is the hierarchical nature of evolution, thus, there could be different levels of genetic structure and there is not a genetic threshold apparent between species or higher taxonomical levels (Coyne & Orr 1998; Hausdorf & Hennig 2010).

Still, the complete lack of any signs of a genetic differentiation based on a highly sensitive test system, such as allele frequency comparisons between polymorphic microsatellite loci, is to our knowledge unprecedented. We would expect at least some differentiation caused by assortative mating even if the system was extremely young. In the well-studied case of Cichlids in Lake Ejagham (Cameroon), species divergence started not much more than 1000–2000 years ago and in spite of continued gene flow between them, they show differentiation at the microstatellite level (Schliewen *et al.* 2001; Poelstra *et al.* 2018). Hence, the alternative is that the two morphs represent an extreme case of morphological plasticity, where two major morphs develop out of a single gene pool, whereby the developmental trajectory to form one or the other morph would be determined early during ontogeny. This is how different morphs arise in social insects (Ross & Keller 1995), but this is usually coupled with a monopolization of reproduction in one of the morphs only. This is not evident in our case, as both morphs can be found to be breeding and there is also no known social interaction between them.

Conclusions

The two morphs present in Lake Catemaco were originally considered as different genera (Contreras-Balderas & Rivera-Teillery 1983), and currently these two morphs are considered as different species (i.e. *A. aeneus* and *A. caballeroi sensu* (Schmitter-Soto 2017)). Based on the

biological species concept (Mayr 1963) or on the genotypic cluster species concept (Mallet 1995), our two lacustrine morphs are either in an extremely early phase of speciation or represent two extreme morphotypes derived from a single gene pool. In this regard, elucidating the genetic basis underlying speciation is a long-standing goal of evolutionary biology, particularly in terms of delimiting a continuous process in a restricted temporal window (Simpson 1951; De Queiroz 2007). Our lacustrine morphs system gives a unique opportunity to understand the genetic basis of how morphological divergence could persist in the presence of gene flow. Even when we failed to recover the two morphs as reproductively isolated, this model provides a unique opportunity to characterize those factors that would promote the ecological divergence.

Declarations

Ethics approval and consent to participate

Fish were collected under permits SGPA/DGVS/05464/, SAGARPA (México), the fish were euthanized following conditions stated in the collecting permit.

Consent for publication

Not applicable

Availability of data and materials

The datasets during and/or analysed during the current study available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica, UNAM, project number IN212419, primary supporting sampling collection and molecular data analyses. Additionally, Max Plank Institute for Evolutionary Biology in Plön Germany provides reagents and sequencing support in the present proposal.

Abbreviations

ML = Maximum likelihood.

BI = Bayesian inference.

SDS = Sodium dodecyl sulfate

EDTA = Ethylenediaminetetraacetic acid

NaCl = Sodium chloride

F_{ST} = one of the Wright's F-statistics, corresponding to a measure of population differentiation due to genetic structure.

F_{IS} = one of the Wright's F-statistics, corresponding to an inbreeding coefficient.

MCMC = Markov chain Monte Carlo

Authors' contributions

CPOG conceived and designed the study, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored drafts of the paper and approved the final draft. EGG analyzed the data, authored drafts of the paper and approved the final draft. DT designed the study, contributed reagents/materials/analysis tools, authored drafts of the paper and approved the final draft. IG conceived and designed the study, contributed reagents/materials/analysis tools, authored drafts of the paper and approved the final draft.

Acknowledgments

The authors thank Carlos Pedraza for field and lab assistance and for his helpful suggestions on an early version of the manuscript. To Lourdes Alcaraz and Silvia Perea for their helpful suggestions on the early version of the article and lab assistance. We acknowledge funding sources from Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (IN212419) to CPOG and Max Plank Institute for Evolutionary Biology in Plön Germany. Fish were collected under permits SGPA/DGVS/05464/, SAGARPA (México).

References

1. Barluenga, M., Stölting, K.N., Salzburger, W., Muschick, M. & Meyer, A. (2006). Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, 439, 719.

2. Beerli, P. & Felsenstein, J. (1999). Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, 152, 763-773.

3. Beerli, P. & Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences*, 98, 4563-4568.

4. Beerli, P. & Palczewski, M. (2010). Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics*.

5. Beheregaray, L., Chae, J., Chao, N. & Caccone, A. (2005). Characterization of microsatellite loci for the Amazonian rummy-nose tetra, Hemigrammus bleheri (Teleostei, Characidae). *Molecular Ecology Notes*, 5, 536-537.

6. Belkhir, K. (2004). 1996-2004 GENETIX 4.05, logiciel sous Windows TM pour la genetique des populations. *http://www. univmontp2. fr/~ genetix/genetix/genetix. htm*.

7. Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. (2000). Genetix, a windowsTM based software for population genetic analyses. laboratoire génome, populations, interactions CNRS UMR 5000. *Université de Montpellier II, Montpellier, France. http://www. univ-montp2. fr* genetix/genetix. htm*.

8. Bradic, M., Beerli, P., Garcia-de Leon, F.J., Esquivel-Bobadilla, S. & Borowsky, R.L. (2012). Gene flow and population structure in the Mexican blind cavefish complex (Astyanax mexicanus). *BMC evolutionary biology*, 12, 9.

9. Contreras-Balderas, S. & Rivera-Teillery, R. (1983). Bramocharax (Catemaco) caballeroi subgen. et. sp. nv. del Lago de Catemaco, Veracruz, México.(pisces: characidae).

10. Coyne, J.A. & Orr, A.H. (1998). The evolutionary genetics of speciation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 353, 287-305.

11. De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic biology*, 56, 879-886.

12. Dieckmann, U., Doebeli, M., Metz, J.A. & Tautz, D. (2004). *Adaptive speciation*. Cambridge University Press.

13. Doebeli, M. & Dieckmann, U. (2003). Speciation along environmental gradients. *Nature*, 421, 259.

14. Elipot, Y., Hinaux, H., Callebert, J., Launay, J.-M., Blin, M. & Rétaux, S. (2014). A mutation in the enzyme monoamine oxidase explains part of the Astyanax cavefish behavioural syndrome. *Nature communications*, 5, 3647.

15. 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.

16. Excoffier, L. & Lischer, H.E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, 10, 564-567.

17. Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices, version 2.9. 3. *http://www2. unil. ch/popgen/softwares/fstat. htm*.

18. Gross, J.B., Borowsky, R. & Tabin, C.J. (2009). A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish Astyanax mexicanus. *PLoS genetics*, 5, e1000326.

19. Guinand, B. (1996). Use of a multivariate model using allele frequency distributions to analyse patterns of genetic differentiation among populations. *Biological Journal of the Linnean Society*, 58, 173-195.

20. Guo, S.W. & Thompson, E.A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, 361-372.

21. Hatfield, T. & Schluter, D. (1999). Ecological speciation in sticklebacks: environment-dependent hybrid fitness. *Evolution*, 53, 866-873.

22. Hausdorf, B. & Hennig, C. (2010). Species delimitation using dominant and codominant multilocus markers. *Systematic Biology*, 59, 491-503.

23. Hori, M. (1993). Frequency-dependent natural selection in the handedness of scale-eating cichlid fish. *Science*, 260, 216-219.

24. Jeffery, W.R. (2009). Regressive evolution in Astyanax cavefish. *Annual review of genetics*, 43, 25-47.

25. Kass, R.E. & Raftery, A.E. (1995). Bayes factors. *Journal of the american statistical association*, 90, 773-795.

26. Magalhaes, I.S., Ornelas-Garcıa, C.P., Leal-Cardin, M., Ramírez, T. & Barluenga, M. (2015). Untangling the evolutionary history of a highly polymorphic species: introgressive hybridization and high genetic structure in the desert cichlid fish Herichtys minckleyi. *Molecular ecology*, 24, 4505-4520.

27. Mallet, J. (1995). A species definition for the modern synthesis. *Trends in Ecology & Evolution*, 10, 294-299.

28. Mayr, E. (1963). Animal species and evolution. *Animal species and evolution.*

29. McGaugh, S.E., Gross, J.B., Aken, B., Blin, M., Borowsky, R., Chalopin, D. *et al.* (2014). The cavefish genome reveals candidate genes for eye loss. *Nature communications*, 5, 5307.

30. Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, 583-590.

31. Nei, M. & Chesser, R.K. (1983). Estimation of fixation indices and gene diversities. *Annals of human genetics*, 47, 253-259.

32. Nosil, P. (2012). *Ecological speciation*. Oxford University Press.

33. Nosil, P., Harmon, L.J. & Seehausen, O. (2009). Ecological explanations for (incomplete) speciation. *Trends in ecology & evolution*, 24, 145-156.

34. Nosil, P. & Schluter, D. (2011). The genes underlying the process of speciation. *Trends in Ecology & Evolution*, 26, 160-167.

35. Ornelas-García, C.P., Domínguez-Domínguez, O. & Doadrio, I. (2008). Evolutionary history of the fish genus Astyanax Baird & Girard (1854)(Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC evolutionary biology*, 8, 340.

36. Ornelas-García, C.P., Bastir, M. & Doadrio, I. (2014). Morphometric variation between two morphotypes within the Astyanax Baird and Girard, 1854 (Actinopterygii: Characidae) genus, From a Mexican tropical lake. *Journal of morphology*, 275, 721-731.

37. Ornelas-García, C.P., Córdova-Tapia, F., Zambrano, L., Bermúdez-González, M.P., Mercado-Silva, N., Mendoza-Garfias, B. *et al.* (2018). Trophic specialization and morphological divergence between two sympatric species in Lake Catemaco, Mexico. *Ecology and evolution*, 8, 4867-4875.

38. Panova, M., Hollander, J. & Johannesson, K. (2006). Site-specific genetic divergence in parallel hybrid zones suggests nonallopatric evolution of reproductive barriers. *Molecular Ecology*, 15, 4021-4031.

39. Poelstra, J.W., Richards, E.J. & Martin, C.H. (2018). Speciation in sympatry with ongoing secondary gene flow and a potential olfactory trigger in a radiation of Cameroon cichlids. *Molecular ecology*, 27, 4270-4288.

40. Presgraves, D.C. (2003). A fine-scale genetic analysis of hybrid incompatibilities in Drosophila. *Genetics*, 163, 955-972.

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

42. Protas, M., Conrad, M., Gross, J.B., Tabin, C. & Borowsky, R. (2007). Regressive evolution in the Mexican cave tetra, Astyanax mexicanus. *Current biology*, 17, 452-454.

43. Protas, M.E., Hersey, C., Kochanek, D., Zhou, Y., Wilkens, H., Jeffery, W.R. *et al.* (2006). Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature genetics*, 38, 107.

44. Raymond, M. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.*, 86, 248-249.

45. Ross, K.G. & Keller, L. (1995). Ecology and evolution of social organization: insights from fire ants and other highly eusocial insects. *Annual Review of Ecology and Systematics*, 26, 631-656.

46. Schliewen, U., Rassmann, K., Markmann, M., Markert, J., Kocher, T. & Tautz, D. (2001). Genetic and ecological divergence of a monophyletic cichlid species pair under fully sympatric conditions in Lake Ejagham, Cameroon. *Molecular Ecology*, 10, 1471-1488.

47. Schluter, D. (1996). Ecological causes of adaptive radiation. *The American Naturalist*, 148, S40-S64.

48. Schmitter-Soto, J.J. (2017). A revision of Astyanax (Characiformes: Characidae) in Central and North America, with the description of nine new species. *Journal of Natural History*, 51, 1331-1424.

49. Simpson, G.G. (1951). The species concept. *Evolution*, 5, 285-298.

50. Sonnenberg, R., Nolte, A.W. & Tautz, D. (2007). An evaluation of LSU rDNA D1-D2 sequences for their use in species identification. *Frontiers in zoology*, 4, 6.

51. Thibert-Plante, X. & Hendry, A. (2009). Five questions on ecological speciation addressed with individual-based simulations. *Journal of evolutionary biology*, 22, 109-123.

52. Van Oosterhout, C., Hutchinson, W.F., Wills, D.P. & Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4, 535-538.

53. West-Eberhard, M.J. (2003). *Developmental plasticity and evolution*. Oxford University Press.

Tables

Table 1. Genetic variation of Characid system at twelve microsatellite loci. n: number of individuals analyzed, A_R: allelic richness, Na: number of alleles, H_o: observed heterozygosity, He: expected heterozygosity, F_{IS}: inbreeding index.

Population		Locus													
		211B	215D	235D	30C	6A	NYU13	142C	16C	207F	229B	NYU14	NYU29	Mean	SD
Astyanax	**AR**	14.7	15.0	20.8	17.7	17.0	23.1	18.7	17.1	17.5	13.8	9.4	22.6	16.9	3.7
n=199	**Na**	18.0	17.0	22.0	21.0	20.0	25.0	20.0	19.0	20.0	16.0	12.0	25.0	19.6	1.1
	Ho	0.8	0.9	0.9	0.8	0.6	0.8	0.6	0.9	0.8	0.9	0.7	0.8	0.8	0.0
	He	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.6	0.9	0.9	0.0
	F_{IS}	0.1	0.0	0.0	0.0	***0.3***	***0.1***	***0.3***	0.0	0.1	0.0	0.0	***0.2***	0.1	0.0
Bramocharax	**AR**	14.9	13.9	20.0	17.8	16.0	25.0	17.9	16.0	17.9	11.9	9.8	20.9	16.4	3.9
n=94	**Na**	15.0	14.0	20.0	18.0	16.0	25.0	18.0	16.0	18.0	12.0	10.0	21.0	16.9	1.2
	Ho	0.8	0.8	0.9	0.9	0.6	0.7	0.6	0.9	0.8	0.9	0.7	0.7	0.8	0.0
	He	0.9	0.9	0.9	0.9	0.8	0.9	0.9	0.9	0.9	0.9	0.6	0.9	0.9	0.0
	F_{IS}	0.0	***0.0***	0.0	0.0	0.2	***0.2***	***0.3***	0.0	0.1	0.0	0.0	0.2	0.1	0.0
Maquinas river	**AR**	4.0	5.0	2.0	3.0	4.0	4.0	4.0	2.0	3.0	6.0	2.0	4.0	3.4	1.2
n=30	**Na**	4.0	5.0	2.0	3.0	4.0	4.0	4.0	2.0	3.0	6.0	2.0	4.0	3.6	0.4
	Ho	0.7	0.7	0.6	0.5	0.6	0.4	0.3	0.4	0.3	0.6	0.6	0.7	0.5	0.0
	He	0.7	0.7	0.5	0.5	0.7	0.4	0.6	0.4	0.5	0.7	0.5	0.6	0.6	0.0
	F_{IS}	0.0	0.0	-0.1	-0.1	0.0	0.2	***0.5***	0.0	0.4	0.2	-0.2	-0.1	0.1	0.1

Significant departures from Hardy-Weinberg equilibrium are shown in bold after sequential Bonferroni correction.

Table 2. Model comparison of gene flow models for Astyanax and Bramocharax

Model	Ln mL	Ln Bayes Factor	Model probability	*Ne*
1) Panmictic	**-383825.44**	-247769.74	0.00	18,791.675
2) *Astyanax to Bramocharax*	-281039.67	-42198.2	0.00	13,958.25
3) *Bramocharax to Astyanax*	-266380.39	-12879.64	0.00	16,458.25

In bold the best model.

Supplementary Files Legend

Table S1. Collecting localities and specimens analyzed.

Table S2 Microsatellites information in *Astyanax* genus.

Table S3. Linkage disequilibrium Test.

Table S4. Genetic variation of Characid system at twelve microsatellite loci. n: number of individuals analyzed, AR: allelic richness, Na: number of alleles, Ho: observed heterozygosity, He: expected heterozygosity, FIS: inbreeding index. Significant departures from Hardy-Weinberg equilibrium are shown in bold after sequential Bonferroni correction.

Table S5. Hierarchical AMOVA between morphs and between geographic regions.

Figures

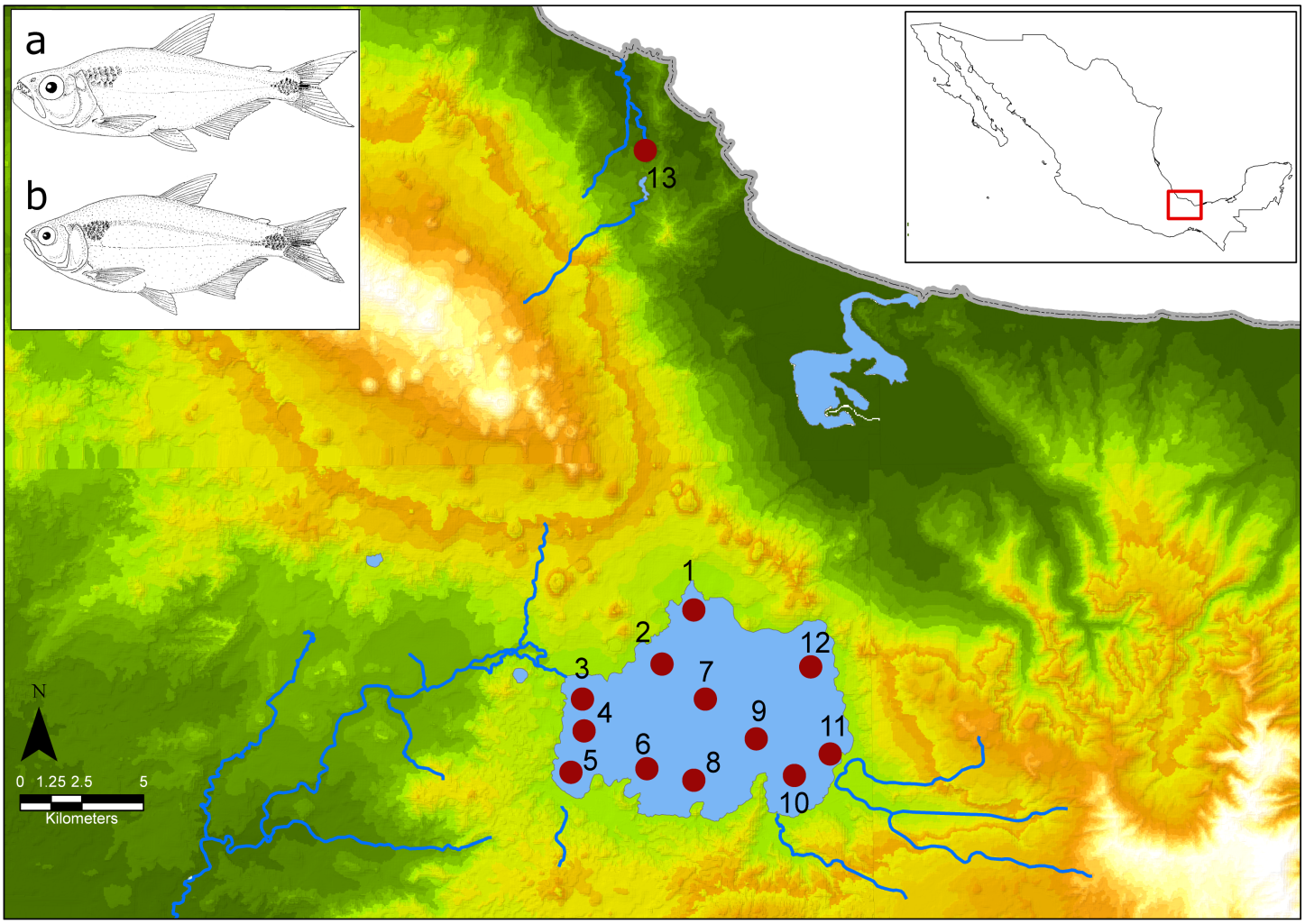

Figure 1

Map of localities - Map of the sampled localities in the Catemaco Lake. a) Bramocharax morph and b) Astyanax morph. 1) Changos Island, 2) Agatepec Island, 3) Finca, 4) Maxacapan, 5) La Victoria, 6) Pozolapan, 7) Catemaco, 8) Ecuniapan, 9) Mimiagua, 10) Las Margaritas, 11) Cuetzalapan, 12) Oxochapan and 13) Maquinas River.

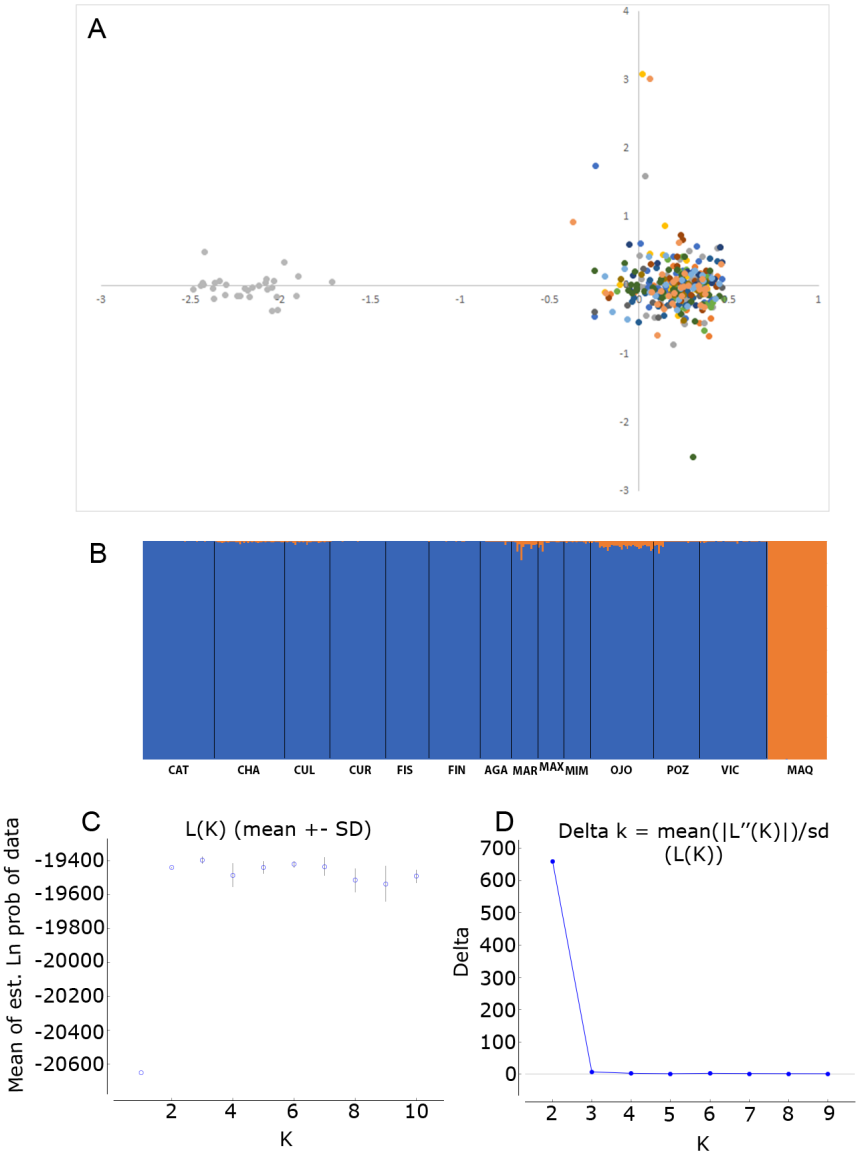

Figure 2

Genetic structure based on microsatellite data - (A) Factorial correspondence analysis for 12 microsatellite loci including the population from river Maquinas. (B) Bayesian population assignment based on 12 microsatellite loci with the software STRUCTURE for sampling locations from Lake Catemaco (10) and Maquinas river, showing the grouping with the highest ΔK (for K = 2). (C) Plot of mean likelihood L(K) and variance per K value from STRUCTURE, and the Evanno et al. (2005) plot for detect the K number of groups that best fit the data. (D) Delta K for STRUCTURE analysis. CAT= Center of Catemaco Lake, CHA= Isla Changos, CUL= Cuetzalapan, CUR= Cuetzalapan river, FIS= La Finca, AGA= Isla Agatepec, MAR= Margaritas, MAX= Maxacapan, MIM= Mimiagua, OJO= Oxochapan, POZ= Pozolapan, VIC= La Victoria and MAQ= Maquinas river.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupportingInformation1.xlsx