Received: 15 December 2019 | Revised: 31 July 2021 | Accepted: 11 August 2021

DOI: 10.1111/mec.16139

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Pervasive admixture and the spread of a large-lipped form in a cichlid fish radiation

Will Sowersby[1,2] | **José Cerca**[3,4,5] | **Bob B. M. Wong**[1] | **Topi K. Lehtonen**[1,6,7] |
David G. Chapple[1] | **Mariana Leal-Cardín**[8,9] | **Marta Barluenga**[8] |
Mark Ravinet[10,11,12]

[1]School of Biological Sciences, Monash University, Melbourne, Victoria, Australia
[2]Department of Biology, Osaka City University, Osaka, Japan
[3]Frontiers of Evolutionary Zoology Research Group, Natural History Museum, University of Oslo, Oslo, Norway
[4]Department of Environmental Science, Policy and Management, University of California, Berkeley, Berkeley, California, USA
[5]Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology, Trondheim, Norway
[6]Department of Biology, University of Turku, Turku, Finland
[7]Organismal and Evolutionary Biology, University of Helsinki, Helsinki, Finland
[8]Department of Biodiversity and Evolutionary Biology, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain
[9]Universidad de Alcalá de Henares, Madrid, Spain
[10]Centre for Ecological and Evolutionary Synthesis, University of Oslo, Oslo, Norway
[11]Division of Population Genetics, National Institute of Genetics, Mishima, Japan
[12]School of Life Sciences, University of Nottingham, Nottingham, UK

Correspondence
Will Sowersby, Department of Biology, Osaka City University, Osaka, Japan.
Email: will.sowersby@sci.osaka-cu.ac.jp

José Cerca, Frontiers of Evolutionary Zoology Research Group, Natural History Museum, University of Oslo, Oslo, Norway.
Email: jose.cerca@gmail.com

Funding information
SYNTHESYS Project

Abstract

Adaptive radiations have proven important for understanding the mechanisms and processes underlying biological diversity. The convergence of form and function, as well as admixture and adaptive introgression, are common in adaptive radiations. However, distinguishing between these two scenarios remains a challenge for evolutionary research. The Midas cichlid species complex (*Amphilophus* spp.) is a prime example of adaptive radiation, with phenotypic diversification occurring at various stages of genetic differentiation. One species, *A. labiatus*, has large fleshy lips, is associated with rocky lake substrates, and occurs patchily within Lakes Nicaragua and Managua. By contrast, the similar, but thin-lipped, congener, *A. citrinellus*, is more common and widespread. We investigated the evolutionary history of the large-lipped form, specifically regarding whether the trait has evolved independently in both lakes from ancestral thin-lipped populations, or via dispersal and/or admixture events. We collected samples from distinct locations in both lakes, and assessed differences in morphology and ecology. Using RAD-seq, we genotyped thousands of SNPs to measure population structure and divergence, demographic history, and admixture. We found significant between-species differences in ecology and morphology, local intraspecific differences in body shape and trophic traits, but only limited intraspecific variation in lip shape. Despite clear ecological differences, our genomic approach

---

Marta Barluenga and Mark Ravinet contributed equally to this study.
Will Sowersby and José Cerca contributed equally to this study.

uncovered pervasive admixture between the species and low genomic differentiation, with species within lakes being genetically more similar than species between lakes. Taken together, our results suggest a single origin of large-lips, followed by pervasive admixture and adaptive introgression, with morphology being driven by local ecological opportunities, despite ongoing gene-flow.

KEYWORDS
admixture, cichlid, ecological specialization, parallel evolution, RAD-seq, resource trait

1 | INTRODUCTION

Comparable phenotypic traits can evolve in different populations that have been exposed to similar selective pressures and ecological challenges (Colosimo et al., 2005; Futuyma, 1998; Gillespie et al., 2020; Schluter et al., 2004; Stern, 2013). This repeated evolution of convergent traits is widespread across taxa, particularly in the context of adaptive radiations, where similar ecotypes often demonstrate clear associations between morphological traits and niche specialization (Gavrilets & Losos, 2009; Gillespie, 2004; Grant & Grant, 2008; Muschick et al., 2012; Schluter, 2000; Simões et al., 2016). Iconic examples include *Anolis* lizards, in which similar ecomorphs have evolved independently on Caribbean islands (Losos et al., 1998; Mahler et al., 2013), and the repeated evolution of limnetic and benthic three-spined stickleback pairs, after independent events of colonization of freshwater habitats by marine ancestors (Bell, 1974; Walker & Bell, 2000). Importantly, adaptive radiations have proven critical for developing our understanding of the mechanisms and processes that underlie population divergence and speciation (Coyne & Orr, 2004; Doebeli et al., 2005; Marques et al., 2019; Rundle & Nosil, 2005; Schluter, 2000; Sobel et al., 2010).

The extent of diversification in a given system is hypothesized to be largely determined by ecological opportunity (in turn driven by niche availability), levels of interspecific competition, and habitat heterogeneity (Losos, 2010; Schluter, 2000). For instance, diversification is predicted to be strongest in populations that inhabit variable or complex habitats with limited interspecific competition, while high intraspecific competition promotes the exploitation of new resources (see Gavrilets & Losos, 2009). Under this scenario, diversification is driven by ecological specialization and divergent natural selection, with adaptation to different niches potentially leading to increasing reproductive isolation (Coyne & Orr, 2004; Skulason & Smith, 1995). Indeed, both comparative (Mahler et al., 2010; Nosil & Reimchen, 2005; Parent & Crespi, 2009) and experimental (Meyer & Kassen, 2007; Rainey & Travisano, 1998) evidence support the significance of ecological opportunity in driving diversification. However, not all lineages undergo diversification even when ecological opportunities are available (Arbogast et al., 2006; Lovette et al., 2002) and knowledge of the conditions that promote the repeated evolution of adaptive traits has therefore become an important component of evolutionary research.

The independent and parallel evolution of analogous phenotypic traits implies that populations may respond to equivalent environmental conditions in similar ways, and provides strong evidence that selection plays a key role in the process of adaptation (Arendt & Reznick, 2008; Berner et al., 2008, 2009; Futuyma, 1998; Schluter et al., 2004). Convergence between closely related lineages in form and function does not necessarily reflect independent evolutionary events (i.e., arising via de novo mutation). Alternative explanations include independent selection on standing genetic variation (Colosimo et al., 2005; Johannesson et al., 2012), the adaptive retention or removal of labile ancestral traits (see Bernatchez, 2004; Østbye et al., 2005), and introgression of adaptive alleles among populations (Berner & Salzburger, 2015; Enciso-Romero et al., 2017). In the latter case, the evolution of a given trait or morphotype can occur once and then enters a new population during dispersal. If dispersal or gene flow among populations is infrequent and selection on adaptive alleles generates a degree of genetic differentiation via reproductive isolation, it could be difficult to distinguish the scenario from parallel evolution caused by de novo mutation. Hence, documenting cases in which independent repeated evolution has occurred remains critical for understanding the interaction between ecological opportunity, selection, hybridization and/or admixture in generating diversity.

One group of organisms that has demonstrated a propensity to diversify is the cichlid fishes (Cichlidae). Indeed, diversification of cichlids within the East African Great Lakes constitutes one of the most well-known and most diverse extant vertebrate adaptive radiations (McGee et al., 2020; Meier et al., 2018; Muschick et al., 2011; Salzburger, 2018; Seehausen, 2006). Across cichlid lineages, multiple adaptive phenotypes have evolved repeatedly (Muschick et al., 2012), both within (Rüber et al., 1999) and between lakes (Kocher et al., 1993; Meyer, 1993). More recently, the role of hybridization in the rapid adaptive radiations of these fishes has been identified as potentially important. Specifically, gene flow and admixture may act as important sources of genetic variation, which in turn, is subject to selection and can be reassembled into different divergent phenotypes when ecological opportunity arises (Marques et al., 2019; McGee et al., 2020; Meier et al., 2017, 2019; Svardal et al., 2019). Less well documented is the evolutionary origin of cichlids in the Neotropics, where lineages have also radiated rapidly (albeit at a smaller scale) and display highly variable trophic morphology (Barlow & Munsey, 1976; Barluenga & Meyer, 2004, 2010; Elmer et al., 2010; Kautt et al., 2020). For example, the Midas cichlid species complex (*Amphilophus* spp., Günther, 1864) is a young adaptive radiation, spanning the two Nicaraguan great lakes (Nicaragua and Managua) and a range of smaller crater lakes, where their diversification is observed at various stages (Barluenga et al., 2006; Elmer et al., 2010; Kautt et al., 2012, 2020), and provides some of the strongest evidence for sympatric speciation in nature (Barluenga et al., 2006; Elmer et al., 2009).

A distinctive ecotype with hypertrophied lips has evolved independently multiple times in distant cichlid lineages, in both Africa (Arnegard & Snoeks, 2001; Colombo et al., 2013; Henning et al., 2017; Oliver & Arnegard, 2010; Salzburger et al., 2005) and the Americas (Burress, 2015; Colombo et al., 2013; Elmer, Kusche, et al., 2010; Elmer, Lehtonen, et al., 2010; Klingenberg et al., 2003; Machado-Schiaffino et al., 2017). Large lips appear to be an evolutionary innovation for utilizing underexploited resources and trophic niches, specifically by facilitating suction feeding to extract invertebrates (primarily crustaceans) from the crevices that form in between rocks (Barlow & Munsey, 1976; Baumgarten et al., 2015; Colombo et al., 2013; Machado-Schiaffino et al., 2017; Seehausen, 1996). One member of the Midas cichlid species complex, *A. labiatus* (the red devil), fills this role and has evolved the large lip trait. This species is endemic to the two great lakes of Nicaragua (although similar forms do exist in the nearby smaller crater lakes), where it has a patchy distribution, associated with rocky habitats (Barlow, 1976).

In the great lakes, the large-lipped *A. labiatus* co-occurs with a more common, but otherwise remarkably similar, thin-lipped congener, *A. citrinellus* (the Midas cichlid). However, the two species appear to mate assortatively and do differ in their ecological niche (Barlow, 1976; Barlow & Munsey, 1976; Barluenga & Meyer, 2004; Baumgarten et al., 2015; Colombo et al., 2013; Machado-Schiaffino et al., 2017). Interestingly, despite clear genomic differentiation in a locus associated with lip size, there is very little reported overall genomic divergence between the large and thin-lipped species in the two great lakes (Kautt et al., 2020). Moreover, large lip forms can appear remarkably rapidly (even within historical time-frames), as seen in certain adjoining crater lakes, which have been colonized by thin-lipped congeners, often from the great lakes (Elmer, Kusche, et al., 2010; Elmer, Lehtonen, et al., 2010; Machado-Schiaffino et al., 2017). Introgression within these crater lakes, and also with *Amphilophus* populations from the older great lakes, has been argued to play an important role in shaping ecological adaptation and sympatric speciation (Kautt et al., 2012, 2016, 2020; Machado-Schiaffino et al., 2017). Indeed, independent colonization events by *Amphilophus* from the great lakes have resulted in the evolution of multiple adaptive radiations in these smaller, more isolated systems (Elmer, Kusche, et al., 2010). However, despite *Amphilophus* cichlids being recognized as a model system for diversification via ecological adaptation, surprisingly little is known about the evolutionary history of these source populations in the great lakes themselves, including what role, if any, admixture has played in the evolution and adaptive radiation of ecological traits. Moreover, the number of locations where *Amphilophus* have been sampled has seemingly, to date, been rather limited relative to the size of the great lakes. We therefore do not have a clear understanding of the level of differentiation between these species, especially considering how readily cichlids hybridize and form geographical variation in body morphology, trophic traits and diet (sensu Konings, 2016). A greater understanding of the evolutionary history of the *Amphilophus* source populations in the great lakes may also provide important insights into the adaptive radiations of these cichlids in nearby crater lakes.

Here, our aim was to investigate the evolutionary history of the large-lipped *A. labiatus*. Specifically, we investigated whether the large-lipped species has evolved and spread to exploit unoccupied and isolated rocky habitat patches between and within the Nicaraguan great lakes via dispersal and/or admixture, or if the large lip trait has repeatedly evolved independently (e.g., via de novo mutation) from thin-lipped populations (sensu Rüber et al., 1999). To achieve this, we used a combination of ecological, geometric morphometric, and genomic data, and investigated the demographic history, population structure and extent of genomic differentiation among these divergent cichlid species from across their range, to better understand the evolutionary history of the *Amphilophus* species complex.

2 | MATERIALS AND METHODS

2.1 | Sampling procedure

To investigate the ecological, morphological and genetic relationship between different trophic phenotypes, we collected a total of 420 specimens ($n = 192$ *A. labiatus* and $n = 228$ *A. citrinellus*) across six distinct locations in Lake Nicaragua (Lago Cocibolca, $n = 339$ specimens) and two locations in the smaller Lake Managua (Lago Xolotlán, $n = 81$ specimens), in different field trips from 2011 to 2019 (Figure S1; Table S1). The research, including collection and sampling procedures, complied with all relevant State laws (Nicaragua: Ministeriodel Ambiente y los Recursos Naturales Nicaragua (MARENA), Permit no. 013-102013; Australia: Animal Ethics Committee of Monash University (BSCI/2012/23). All specimens were collected using a combination of gillnets and handlines, often with the aid of local fishermen. Each fish was sampled using a standard procedure, that is, it was humanely killed with a lethal dose of clove oil solution (unless purchased already dead from a fisherman), photographed (left lateral size), its pharyngeal jaw was removed, fin and muscle tissue samples were taken for DNA and stable isotope analysis, and, lastly, its entire gut was removed for diet analysis. All tissue and gut samples were preserved in ethanol and taken to the laboratory at the Museo Nacional de Ciencias Naturales CSIC, Madrid, Spain.

2.2 | Geometric morphometric analysis

We assessed the size of the body, head, and lips of a subset individuals ($n = 165$ *A. labiatus* and $n = 203$ *A. citrinellus*) from all locations

(n = 81 in Lake Managua and n = 287 in Lake Nicaragua) using landmark-based geometric morphometric analysis. The majority of samples were collected during a 2013 field season (n = 230), with additional specimens, collected between 2011–2019, included in the analysis to boost the sample size from locations that were relatively underrepresented in the primary (2013) survey effort. In separate analyses, four landmarks were placed on the lips, 10 landmarks on the head, and 15 landmarks on the body, at characteristic points on images of each individual (TPS DIG 2.12, Rohlf, 2012; sensu Klingenberg et al., 2003, Magalhaes et al., 2015; see Figure S2). Next, we computed partial warps (PW) and reduced these PWs to relative warps (RW), equivalent to principal components in a principal component analysis (PCA) (TPSRELW 1.46, Rohlf, 2011), which were used as the explanatory factor in further statistical analyses (described below). Further, to visualize variation in lip, head, and body shape between and within species we conducted a separate canonical variate analysis (CVA) for each of the three traits, per species, in MORPHOJ 1.07A (Klingenberg, 2011).

We ran separate univariate Bayesian mixed models to assess differences in trait shape (deviation from the consensus) in respect to the two main axes of lip, head, and body shape variation extracted from the landmark positions (i.e., RW1 and RW2). The first six models, three with RW1 and three with RW2 of lip, head, or body shape as a response variable, were fitted with species (*A. labiatus* or *A. citrinellus*) and lake (Lake Nicaragua or Lake Managua) as fixed effects, standard body length (cm) as a covariate, and individual ID added as a random effect. We repeated this process with landmark data from *A. labiatus* and *A. citrinellus* separately, with models fitted with the RW1 or RW2 of lip, head, or body values from either *A. labiatus* or *A. citrinellus* as the response variable, with collection location as the fixed effect, standard body length as a covariate, and individual added as a random effect.

Lastly, we repeated the above process to assess shape variation in the lower pharyngeal jaw in 222 individuals (n = 105 *A. labiatus* and n = 117 *A. citrinellus*). Here, eight landmarks were placed at identical characteristic points on each jaw (see Figure S3).

2.3 | Assessment of stomach and gut content

Gut contents were removed from the stomach and intestinal tracts of every individual, with the content of the gut confidently identified in 216 individuals (n = 102 *A. labiatus*, n = 114 *A. citrinellus*). All food items were identified under a dissection microscope into one or more of the following categories: mollusks, crustaceans, insects, zooplankton, fish (remains, scales, eggs), plant material, algae, detritus and silt (e.g., sand), and other. We determined the percentage that each food item constituted, in addition to the overall fullness of the gut itself. To reduce the dimensionality of the diet variable space we conducted a PCA of the food items identified in the gut content analysis. We also measured the length of the gut, relative to standard body length, in 221 individuals. We compared gut length relative to standard body length between *A. labiatus* and *A. citrinellus* using a *t - test* (the dependent variable met the assumptions for this test).

2.4 | Stable isotope analysis

A subset of individuals was selected haphazardly, to ensure an adequate number of individuals from each species and location, for stable isotope analysis. Muscle tissue samples (n = 49 *A. labiatus*, n = 71 *A. citrinellus*) were dried (60 °C for 24 h) and crushed in preparation for the purpose. Solid samples (1 mg ± 0.5 mg) were then analysed for 13 carbon ($\delta^{13}C$) and 15 nitrogen ($\delta^{13}N$) signatures at the Department of Plant Sciences Stable Isotope Facility at UC Davis, California. To assess whether *A. labiatus* and *A. citrinellus* differ in diet and ecological niche between and within the great lakes, we ran two separate Bayesian models. First, with carbon ($\delta^{13}C$) values as the response variable, species and lake fitted as the fixed effects and individual added as a random effect. The second model contained nitrogen ($\delta^{13}N$) values as the response variable, species and lake fitted as the fixed effects and individual added as a random effect.

2.5 | Principal component analysis of ecological data

To reduce the dimensionality of the variable space we conducted a PCA of the ecological variables (i.e., stable isotope values (N and C), standard body length, lower pharyngeal jaw shape, lip, head, and body shape) on a subset of individuals (n = 111; n = 64 *A. citrinellus* and n = 47 *A. labiatus*). Specifically, we selected those individuals that had been sampled for stable isotope analysis, to limit the number of missing values. Any missing variables were handled in the PCA analysis by using the “mice package” in R. Inspection of the imputed data distribution matched the original data distribution.

2.6 | DNA extraction, RAD library preparation, sequencing and quality filtering

Genomic DNA was extracted from a subset of individual samples using DNeasy Blood & Tissue Kit (Qiagen, USA). Following DNA extraction, samples were digested using the Sbf1 enzyme following a standard RAD-sequencing protocol (as described in Roesti et al., 2012). We pooled 40 (barcoded) individuals into five libraries (total n = 200). Each library contained a mixture of individuals from each of the sampled populations to account for downstream batch effects. Libraries were sequenced (single-end) on separate Illumina HiSeq2500 lanes (200 bp reads) at the ETH Zurich Department of Biosystems Science and Engineering in Basel, Switzerland. During initial sequencing, one library (Library 4) failed, and was resequenced using the rapid-run option, resulting in lower coverage for these individuals. An overview of the bioinformatic procedures can be found in the Supporting Information (Figure S4). Library quality summaries were generated using FASTQC (Andrews, 2010) to guide downstream filtering. Raw reads were filtered and demultiplexed using the process_radtags module of STACKS v 1.30 (Catchen et al., 2011, 2013). To ensure only high-quality reads were retained, all reads were trimmed to 175 bp and any read with an average Phred quality score <20 was removed.

2.7 | Read alignment, variant calling and data filtering

High-quality reads were aligned to the *Amphilophus* reference genome assembly version 5 (www.ncbi.nlm.nih.gov/bioproject/PRJEB6974/; Elmer et al., 2014) using BWA-MEM 0.7.10 (Li, 2013). The resulting BAM files were then sorted. After this, BAMs were realigned around insertion-deletion polymorphisms to prevent false positive variant calls. Base calling was then done for variant and invariant sites using the GATK 3.7 HAPLOTYPECALLER (DePristo et al., 2011) on individuals before joint genotyping across the entire data set. The resulting variant call format file (VCF) was then filtered to remove indels and annotated with GATK filter thresholds. Since there are no standard values for these parameters, we used vcfrandomsample from vcflib (Garrison, 2016) to randomly select 100,000 variants and then examined their distributions to set appropriate thresholds, as per the GATK best practice guidelines. We therefore filtered all sites with a quality:depth ratio (QD) <5, Fisher strand bias (FS) >5, mapping quality (MQ) of <40, MQRanksum <–10.0 and read position rank sum <–5. Following this, we applied a more stringent set of filters on our called variants; we only allowed variants with a depth of coverage between 15–200x, and with a proportion of missing data below 25% using VCFTOOLS v0.1.13 (Danecek et al., 2011). Finally, considering the effect of missing data in RAD-seq data sets (Cerca et al., 2020, 2021) we estimated "missingness" per individual using VCFTOOLS (--missing-indv) and removed 17 individuals which had at least 50% missing data (variant call format 1; Figure S4). Additionally, we generated a second data set, where, in addition to the aforementioned filters, all sites with a minor allele frequency of <0.01 were removed (variant call format 2; Figure S4).

2.8 | Population structure analysis

Since analyses of population structure assume independence among sites, we first performed linkage pruning on our variants (Variant Call Format 3; Figure S4). To do this, we calculated linkage disequilibrium within a 50 Kb sliding window, using a step of 10 variant sites and removing any sites exceeding an r^2 threshold exceeding 0.1 using PLINK v1.90B5.2 (Chang et al., 2015; Gaunt et al., 2007; Taliun et al., 2014). We then partitioned genetic variation across orthogonal axes using a PCA, included in the R package "ADEGENET" (Jombart, 2008; Jombart & Ahmed, 2011; R Core Team, 2017). We further explored population structure using a parameter-based population cluster model implemented in the software ADMIXTURE 1.3 (Alexander et al., 2009), running analyses from 1–9 populations (*K*). We obtained an estimate of the best population (*K*) by generating standard error scores using leave-one-out cross validation and selecting the value of *K* with the lowest log likelihood score.

2.9 | Phylogenetic analysis

We studied the population history of both species using a phylogenetic network and a phylogenetic tree based on the variant data (variant call format 2; Figure S4). For the phylogenetic network we converted the variant call format file using VCF2PHYLIP v2.0 (Ortiz, 2019), and loaded the resulting PHYLIP file in SPLITSTREE v4 (Huson & Bryant, 2006). For the phylogenetic tree we used SNAPP (Bryant et al., 2012), included as part of BEAST2 (Bouckaert et al., 2014). SNAPP uses the coalescent model to infer trees based on single-nucleotide polymorphisms, focusing on integrating possible gene trees in opposition to finding a maximum credibility tree. To run SNAPP efficiently, we selected 4000 random SNPs from the linkage pruned VCF and kept only 10 individuals per population. Populations were assigned based on the decomposition of genetic divergence including *A. labiatus*-Lake Managua (L-Ma), *A. citrinellus*-Lake Managua (C-Ma), *A. labiatus*-Puerto Diaz (L-PD), *A. labiatus*-Lake Nicaragua (L-Nic), *A. citrinellus*-Lake Nicaragua (L-Nic). We ran an MCMC for 933,000 generations, sampling every 1000 generations and obtaining an ESS posterior probability of 424 after a 10% burnin cutoff.

2.10 | Population divergence

To detect parallel population divergence, we estimated F_{ST} using ARLEQUIN (Excoffier & Lischer, 2010; Lischer & Excoffier, 2012), by doing pairwise comparisons on the populations detected by the PCA. Since the overall objective was to test for parallel evolution *vs* introgression, and considering that individuals from Puerto Diaz in Lake Nicaragua have considerably different allelic variation when compared to the remaining populations (see PCA), we excluded these individuals from the population divergence analyses. Specifically, after removing individuals from Puerto Diaz, we focused on the four main clusters identified by the PCA: *A. labiatus* from Lake Nicaragua; *A. labiatus* from Lake Managua; *A. citrinellus* from Lake Nicaragua; *A. citrinellus* from Lake Managua. Additionally, we obtained estimates of divergence using AMOVA within and among populations using ARLEQUIN, and by using 1000 permutations.

2.11 | Effective population size and demographic analyses

We obtained estimates of effective population sizes using the linkage disequilibrium (LD) method included in NEESTIMATOR v2.1 (Do et al., 2014) to determine N_e parameter space in coalescent simulations (variant call format 1; Figure S4). We conducted demographic analyses using coalescent simulations with FASTSIMCOAL v2 (variant call format 1; Figure S4; Excoffier et al., 2013). FASTSIMCOAL2 relies on arbitrary, user-defined demographic scenarios, and uses a coalescent-simulation framework to simulate different site frequency spectrums (SFS) for different scenarios. Then, the observed SFS was compared to the estimated SFS, and a range of parameters, including time of coalescence, population size and gene flow, is inferred from the most likely scenario. To obtain the SFS of these populations, we used a data set without a minimum allele frequency cutoff (see above). To generate the SFS we used the "easySFS.py" (https://github.com/isaacovercast/easySFS), projecting each population (--proj) to six individuals. For these analyses, we excluded *A. labiatus* individuals from the site Puerto Diaz, similar to the differentiation analysis, and thus focused on *A. labiatus* from Lake Nicaragua (excluding Puerto Diaz); *A. labiatus* from Lake Managua; *A. citrinellus* from Lake Nicaragua; *A. citrinellus* from Lake Managua. In total, we included 14 models of various plausible scenarios, including gene flow at different time periods, dispersal, and hybrid origin (see Figure 1).

2.12 | Genetic origin of the lip trait

Using previously identified potential candidate-genes (Manousaki et al., 2013), we procured the genetic origin of the lip-trait by determining genetic divergence, as measured by scans of F_{ST} (pairwise: *A. citrinellus* vs. *A. labiatus*), in scaffolds where these genes occur (Ravinet et al., 2017). Specifically, we downloaded the six differentially expressed lip genes, as identified by Manousaki et al. (2013), including Calpain-9 gene candidate (GI: 348533281), the Apolipoprotein D-like gene candidate (GI: 348522161), a hpothetical protein LOC100692391 (GI: 348542207), the four-and-a-half LIMdomain protein 2-like (GI: 348531006), and the GTPase IMAP family member 8-like gene candidate (GI: 348545456). We determined the scaffold where these genes occur using BLAST. We were unable to find the differentially expressed Myelin-oligodendrocyte glycoproteinprecursor gene candidate (GI: 100305242). For the five identified scaffolds, we obtained estimates of F_{ST} along moving windows of 50,000 bp, with a step size of 10,000 bp, by using the Weir-Cockerham implementation from vcftools and plotting these over R.

3 | RESULTS

3.1 | Geometric morphometric analysis

Analysis of the landmark data revealed a significant difference between the mean lip shape of *A. labiatus* and *A. citrinellus* (RW1 82.38%; estimate = 0.246 (95% CI = 0.219, 0.277) $P_{MCMC} < 0.001$, Table S2; but not on RW2 13.80%, Table S3), with *A. labiatus* having larger lips compared to *A. citrinellus* (as per Colombo et al., 2013; Klingenberg et al., 2003; Figure S5a). Within *A. labiatus*, there were no significant differences in lip shape across locations (Tables S4–S5). Similarly, we found no significant differences along the major axis of lip shape variation within *A. citrinellus* across locations, whereas *A. citrinellus* individuals differed between La Virgen and the rest of the locations with respect to the minor axis of lip shape variation (RW2; estimate = 0.136 (95% CI = 0.033, 0.232) $P_{MCMC} = 0.011$, Tables S6–S7). Visual inspection of CVA plots, which maximize the separation between groups, show that overall little variation in lip shape exists within either *A. citrinellus* or *A. labiatus* across locations (Figure 2c,f).

Amphilophus labiatus had a significantly narrower mean head shape than *A. citrinellus* (RW1 37.59%; estimate = 0.068 (95% CI = 0.061, 0.074) $P_{MCMC} < 0.001$, Table S8; RW2 15.96%; estimate = -2.002^{-02} (95% CI = -2.716^{-02}, -1.345^{-02}) $P_{MCMC} < 0.001$, Figure S6a, Table S9). There was no significant difference in *A. labiatus* mean head shape across locations (Tables S10–S11). In contrast, we found that *A. citrinellus* from the La Virgen population differed in mean head shape compared to other locations (RW1; estimate = 0.061 (95% CI = 0.027, 0.095) $P_{MCMC} < 0.0001$, Tables S12; but not on RW2, Table S13). Visual inspection of CVA plots suggests *A. labiatus* from Lake Managua tend to have rounder shaped heads than *A. labiatus* from Lake Nicaragua (Figure 2b). *Amphilophus citrinellus* from La Virgen appear to have narrower heads and a larger mouth, compared to other *A. citrinellus* populations (Figure 2e).

Amphilophus labiatus and *A. citrinellus* differed significantly in mean body shape (RW1 23.98%; estimate = 0.028 (95% CI = 0.023, 0.031) $P_{MCMC} < 0.0001$, Table S14; RW2 17.42%; estimate = 0.015 (95% CI = 0.011, 0.02) $P_{MCMC} < 0.0001$, Table S15; Figures S7a–S7b), with *A. citrinellus* having dorsoventrally an overall higher body compared to *A. labiatus* (Figure S7). There were no significant differences in mean *A. labiatus* body shape between locations (Tables S16, S17). Similarly, *A. citrinellus* body shape did not significantly differ between locations (Tables S18,S19). Visual inspection of CVA plots suggests that *A. labiatus* body shape varies nonsignificantly between the lakes, with individuals from Lake Managua tending to have taller bodies, dorsoventrally (Figure 2a). *Amphilophus citrinellus* individuals from La Virgen show a nonsignificant trend of having shorter bodies and larger mouth/lip regions compared to other populations ($P_{MCMC} = 0.08$; Figure 2e).

Mean lower pharyngeal jaw shape (Figure S8) differed significantly between *A. citrinellus* and *A. labiatus* (RW1 44.22%, estimate = 0.041 (95% CI = 0.031, 0.05) $P_{MCMC} < 0.001$, Table S21; RW2 15.10%, estimate = 0.891 (95% CI = 0.236, 1.33) $P_{MCMC} = 0.004$); Table S22) and between specimens from the two lakes (estimate = −0.028 (95% CI = −0.042, −0.012) $P_{MCMC} < 0.0001$; Table S20). Specifically, Lake Managua contained a greater proportion of individuals with slender papilliform jaw types, compared to Lake Nicaragua, where jaw shape appeared to be more varied (Figure S9). However, we have a smaller representation of pharyngeal jaws from Lake Managua, and our sample could therefore be missing variation in jaw shape from that lake. *Amphilophus labiatus* have more slender papilliform shaped jaws than *A. citrinellus*, and overall, less variation in jaw shape compared to *A. citrinellus* (Figures

FIGURE 1 Demographic analysis based on the genomic data (RADseq). On the left, likelihood of different models is presented, with the models being drawn on the right. For each model, there is a tree-like representation, and coalescence events are highlighted with horizontal, dashed blue lines. Arrows represent the direction of gene flow, with double-headed arrows representing symmetric gene flow. Cn, *A. citrinellus* in Lake Nicaragua; Ln, *A. labiatus* in Lake Nicaragua; Cm, *A. citrinellus* in Lake Managua; Lm, *A. labiatus* in Lake Managua

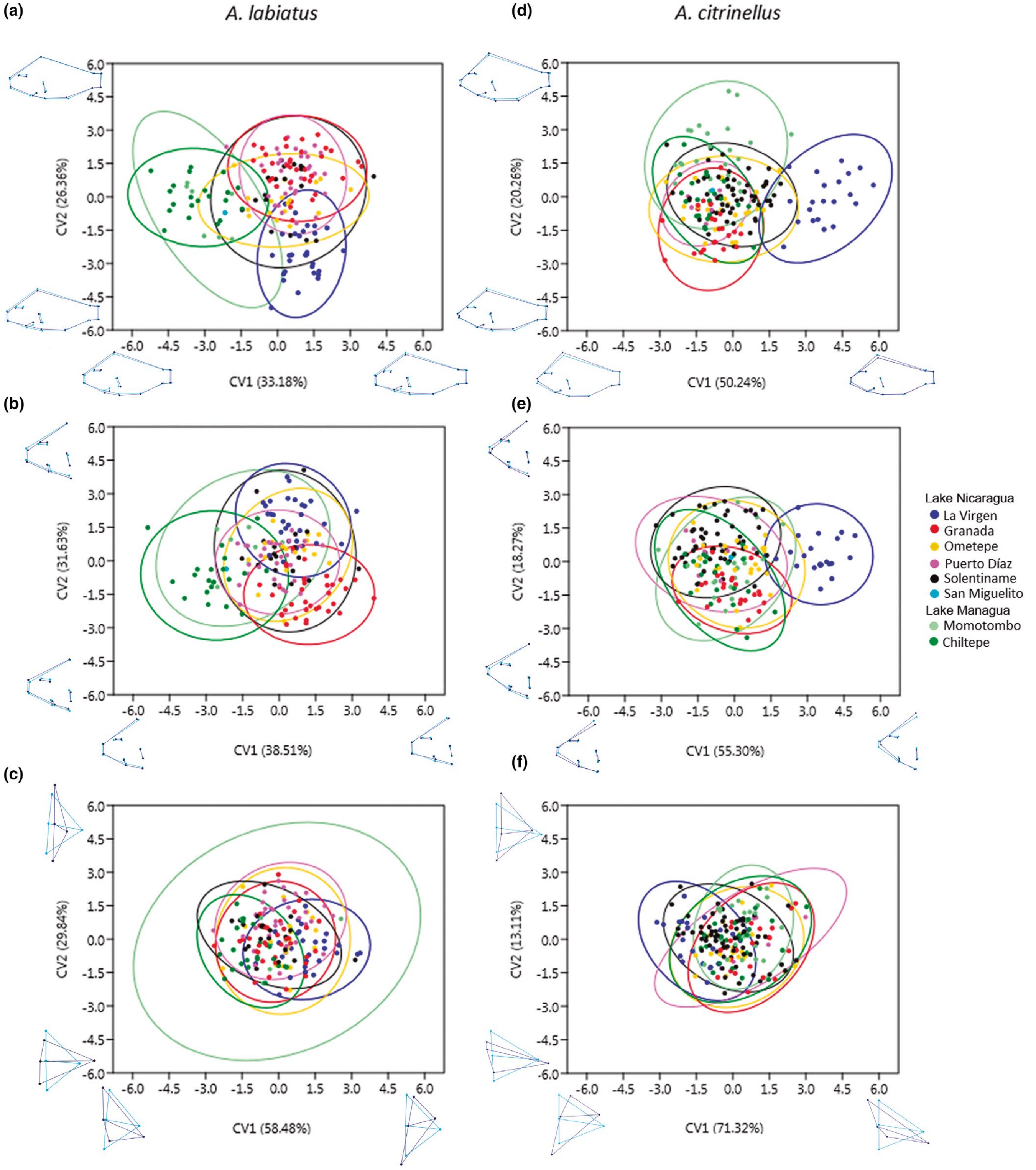


FIGURE 2 Canonical variate analysis (CVA) plots displaying patterns of shape variation within *A. labiatus* (left) and *A. citrinellus* (right) bodies (a, d), heads (b, e), and lips (c, f). CVA is used to analyse group structure in multivariate data and maximally separates predefined groups of interest (here locations). Wireframes (not to scale) are orientated anterior to posterior, dark blue lines represent the shape at the extreme of the axis, light blue lines represent the mean shape along each axis

S9–S10). We do note that, while not directly addressed in this current study, jaw shape in *A. citrinellus* appears to be bimodal (either molariform or papilliform), with jaw shape being anecdotally associated with differences in diet (M. Barluenga, unpublished data). Therefore, shape variation in the lower pharyngeal jaw of *A. citrinellus* will largely depend on the degree of molarization occurring in each lake.

3.2 | Assessment of stomach and gut content

In *A. labiatus*, the most common gut content item was crustaceans (Lake Nicaragua 18.34%, Lake Managua 18%; Figure 2; Figures S11). The most common food item found in the gut content of *A. citrinellus* was molluscs (Lake Nicaragua 23.78%, Lake Managua 20%; Figure S11). We note that many samples had a high proportion of unidentifiable digested material. To reduce the dimensionality of the diet data, we conducted a PCA of the diet of both *A. labiatus* and *A. citrinellus*. The first PC explained 21.4% of the variation and the second PC explained 19.5% of the variation in the diet of the two species (Figure S12). Visual inspection of the PCA demonstrates that diet is quite varied across individuals within both species, but as predicted, snails (mollusks) load most clearly with *A. citrinellus* while crabs (crustaceans) load more with *A. labiatus*. In addition, *A. citrinellus* have a significantly longer mean gut length adjusted for body length compared to *A. labiatus* (*t* test, $t_{218} = 2.0382$, $p = 0.042$; Figure S13).

3.3 | Stable isotope analysis

There was a significant difference in mean δ^{13}C values between *A. labiatus* and *A. citrinellus* (estimate = 0.891 (95% CI = 0.23, 1.33) $P_{MCMC} = 0.004$, Figure 3), with more depleted δ^{13}C values in *A. citrinellus*. There was no difference within species in mean δ^{13}C values between Lake Nicaragua and Lake Managua (Table S22). However, *A. labiatus* from the La Virgen location (Lake Nicaragua) do show a trend of lower δ^{13}C values compared to *A. labiatus* from other locations (Figure 3). δ^{15}N values did not significantly differ between the two species, or between the two great lakes (Table S23, Figure 3).

3.4 | Ecomorphological PCA

To reduce the dimensionality of the data we conducted a PCA combining ecological and geometric morphometric variables. The first PC explained 39.67% of the variation and the second PC explained 15.96% of the variation in the ecomorphological data (Figure 4). Visual inspection of the PCA (see Figure 4) shows a high degree of individual variation within both species and locations. However, overall *A. labiatus* and *A. citrinellus* are primarily separated along PC1, with a degree of overlap between the two in the ecomorphological variable space. Interestingly, the majority of the ecomorpho variables included in the PCA (i.e., δ^{13}C values, lower pharyngeal jaw shape, lip shape, head shape and body shape) load together along PC1, which is the PC that mostly separates *A. labiatus* and *A. citrinellus*. In contrast, body length and δ^{15}N values load strongest on PC2, the axis that also explains much of the variation within species, which is probably driven by age-based differences – that is, ontogenetic shifts in diet and size.

3.5 | Genomic data

We obtained a total of 1,467,644,584 illumina reads after sequencing 198 individuals over five Illumina lanes. After mapping to a draft version of the *A. citrinellus* reference genome, calling and filtering variants, we retrieved a data set comprising 895,111 sites. After filtering for minimum allele frequency, coverage, linkage disequilibrium and missing data, we obtained a data set comprising 16,030 SNPs with an average coverage of 76×, which was used for population genomic analyses.

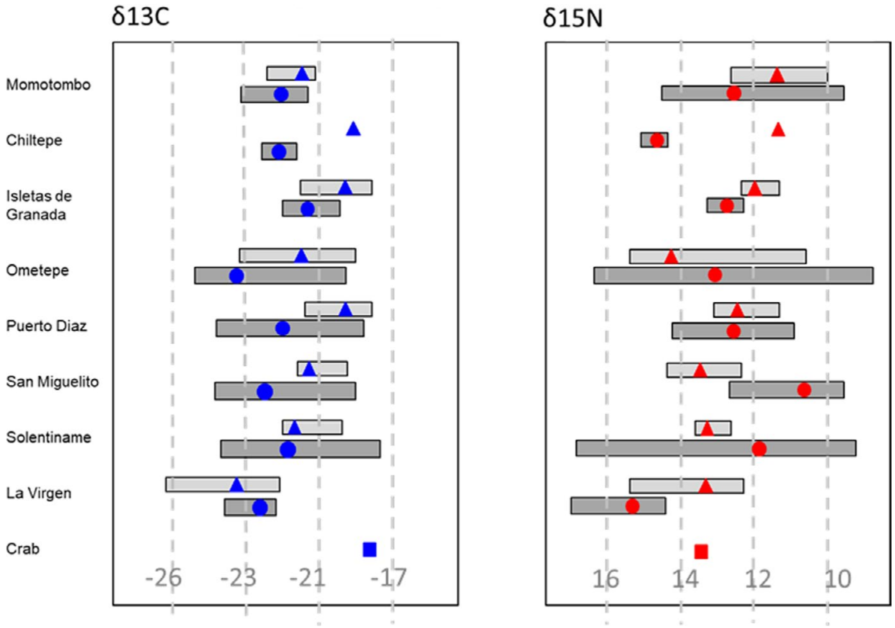


FIGURE 3 Mean stable isotope values obtained from *A. labiatus* (triangles) and *A. citrinellus* (circles) muscle tissue across locations, δ^{13}C (blue) and δ^{15}N (red), grey bars = range

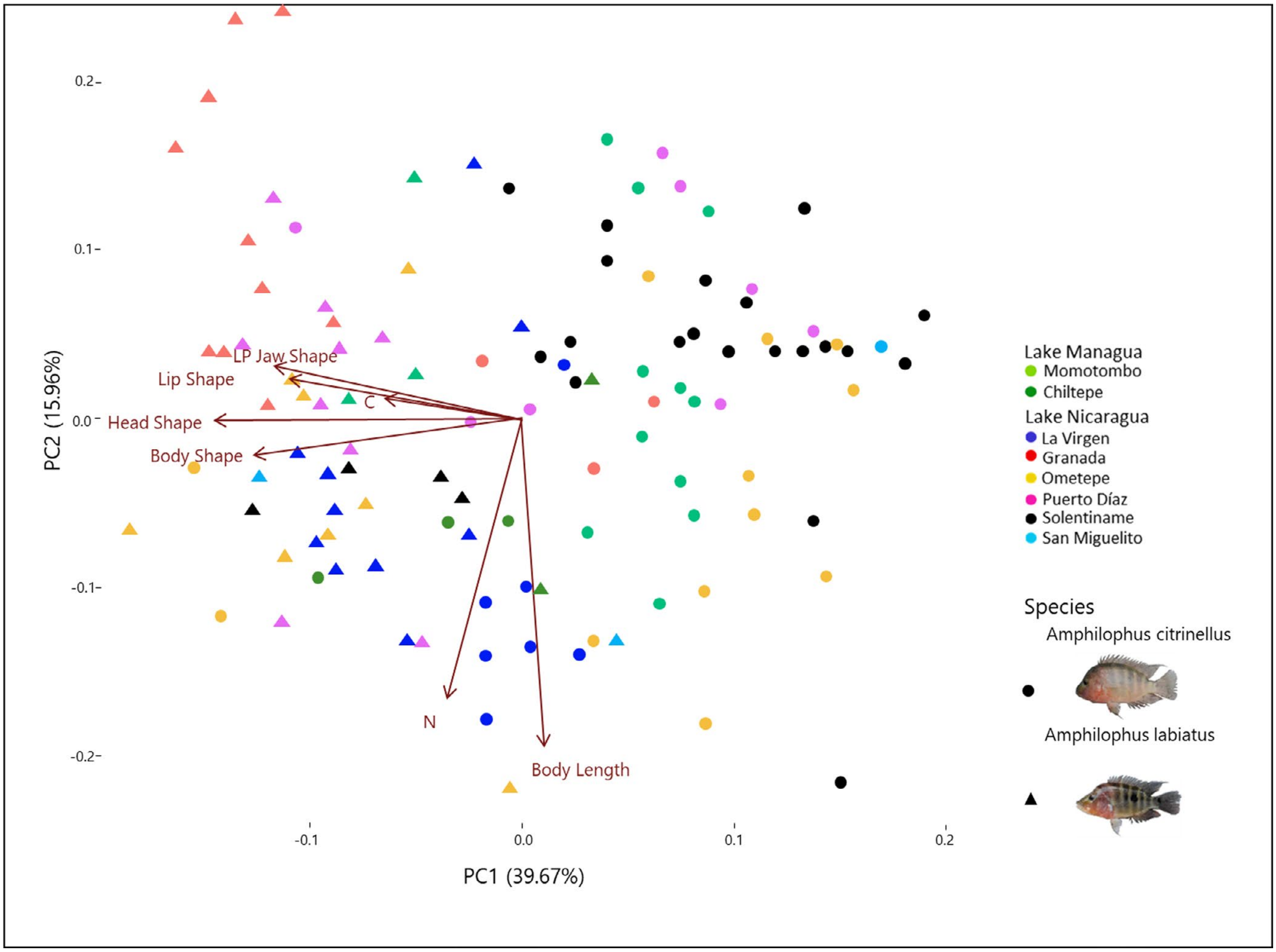


FIGURE 4 Principal component analysis (PCA) of the ecological variables collected from *A. labiatus* (triangles) and *A. citrinellus* (circles) across two locations in Lake Managua and six locations in Lake Nicaragua. PC1 explains 39.67% and PC2 15.96% of the variation in the data

3.6 | Population structure

We decomposed genetic variation in principal component axes, the first three of which accounted for a total of only 6.4% of the variation in the data set. The first major axis explained 2.9% of the variation in the data, separating individuals from both lakes (Figure 5). The second axis explained 2% of the observed variation, and separated *A. labiatus* at the locality Puerto Diaz from other Lake Nicaragua populations (Figure 5). The species are moderately differentiated on the second axis, with this separation being more perceptible in fish of Lake Nicaragua, which may suggest that the lake is either at a later stage of divergence or experiences less admixture. Interestingly, four individuals from Lake Nicaragua were clustered with individuals from Lake Managua, and four from Lake Managua clustered with individuals from Lake Nicaragua, suggesting dispersal.

Admixture and dispersal between lakes is also evident from the ADMIXTURE analysis, which shows individuals with mixed ancestry. In agreement with the PCA, a subset of individuals ($n = 8$ in total) in each lake were assigned to the cluster corresponding to the majority of individuals from the other lake – suggesting recent dispersal events (Figure 6). Cross-validation error of ADMIXTURE analyses suggests that $K = 2$ is the best fit, reflecting mostly lake structure, rather than species (Figure S14). This result implies that individuals within lakes are more closely related to one another than they are to the individuals from the same species in different lakes, meaning that *A. labiatus* and *A. citrinellus* are poorly differentiated species. Furthermore, while there is evidence for dispersal between lakes, a large proportion of both *A. citrinellus* and *A. labiatus* individuals from Lake Nicaragua showed some level of Lake Managua ancestry—suggesting asymmetric dispersal, concordant with the flow of the river (Tipitapa), which intermittently connects both lakes.

In agreement with the pca and admixture analyses, the phylogenetic analysis found that divergent species within lakes are more closely related to each other, than species between lakes (Figures 5–7). The branch leading to the species from Lake Managua was strongly supported (posterior probability = 1). However, the level of support for branches separating lineages in Lake Nicaragua was less clear with a posterior support of 0.88 for all Lake Nicaraguan

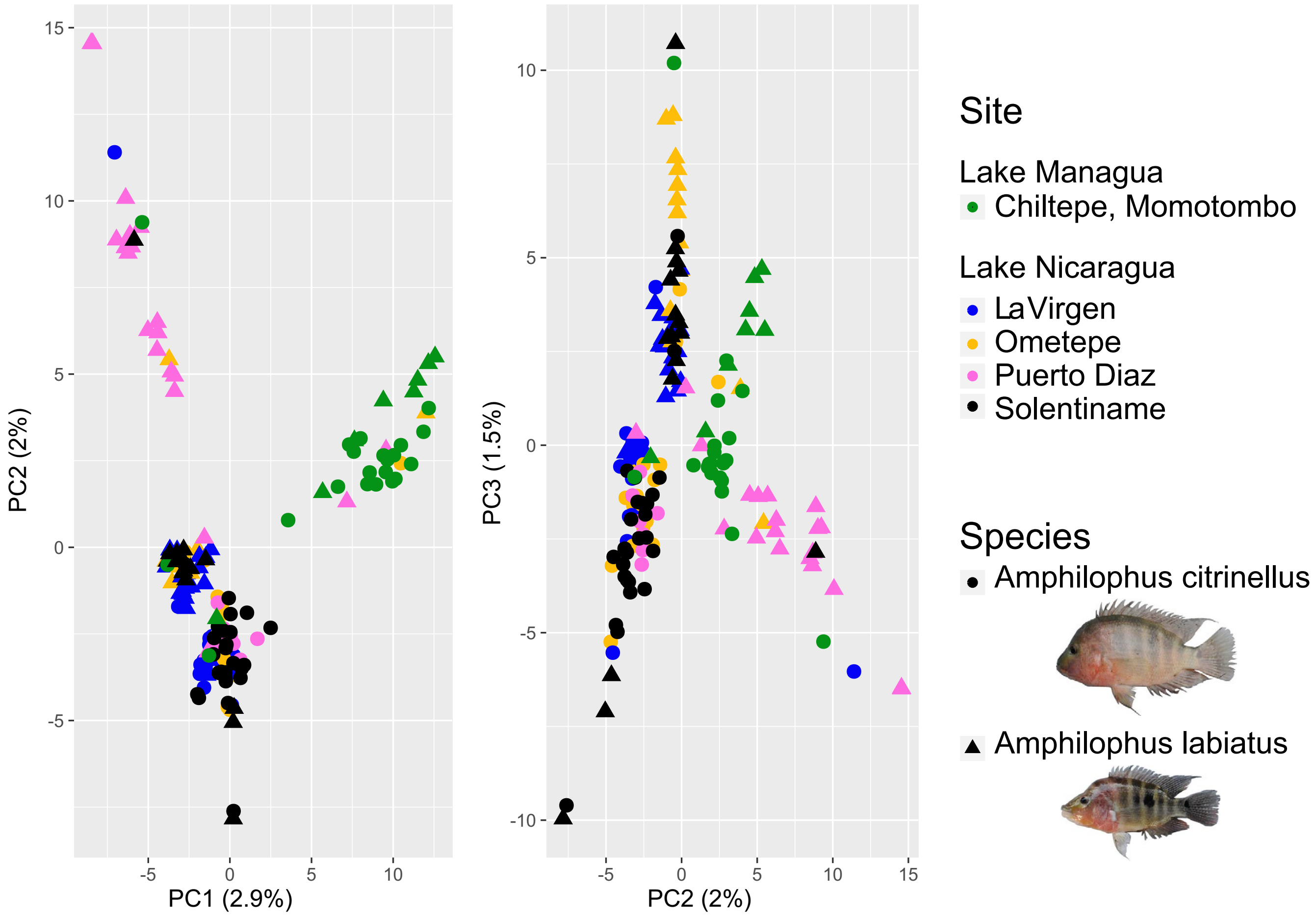


FIGURE 5 Principal component analysis (PCA) of genomic data (RADseq) from *A. labiatus* (triangles) and *A. citrinellus* (circles) across two locations in Lake Managua and six locations in Lake Nicaragua. On the left PC1 versus PC2, and on the right PC1 versus PC3. PC1 explains 2.9% of the variation, PC2 2% and PC3 1.5%

groups (i.e., a split between *A. labiatus* from Puerto Diaz and all other groups from this lake). *Amphilophus labiatus* and *A. citrinellus* from the remaining sites in the lake appear as sister taxa, with a lower posterior support of 0.67. In the phylogenetic network it is evident that individuals from different species and lakes are closely related, as the network branches mostly from the middle (Figure 7b). While specimens from Lake Managua form a fairly defined cluster, they are not far removed from the centre of the network, and four specimens from Lake Nicaragua are nested within the "Managua branches" (Figure 7b).

3.7 | Population divergence

Mean genome wide F_{ST} estimates based on 16,030 SNPs suggest that genomic differentiation between *A. labiatus* and *A. citrinellus* is lowest in Lake Nicaragua ($F_{ST} = 0.01$; see also Tables S24 and S25). In comparison, differentiation between species in Lake Managua is over two times greater ($F_{ST} = 0.02$; Table 1). Mean F_{ST} between *A. citrinellus* in the two lakes is of a similar magnitude ($F_{ST} = 0.037$), but the highest levels of differentiation are seen between the two *A. labiatus* groups, where mean $F_{ST} = 0.055$. AMOVA analysis suggests that a strikingly high percentage of the variation (97.36%) occurs within populations, while only 2.63% occurs between populations.

3.8 | Effective population size and demography

We estimated the effective population size (N_e) using the linkage disequilibrium method (LD), which indicated N_e was generally low. For *A. citrinellus* individuals N_e was 51 and 14 in Lake Nicaragua and Lake Managua, respectively, and it was 46 and five for *A. labiatus* in Lake Nicaragua and Lake Managua (Table 2). When broken into populations, NeEstimator suggested effective population sizes for *A. citrinellus* within Lake Nicaragua of 12 individuals in Ometepe, 14 in La Virgen, 19 in Solentiname, and six in Puerto Diaz; and for *A. labiatus* nine individuals in Ometepe, 12 in La Virgen, 11 in Solentiname, and 14 in Puerto Diaz (Table 2).

We next assessed the demographic history of the system by testing the presence/absence and prevalence of gene flow throughout the evolutionary history of the radiation. Likelihood assessment of coalescent-based simulations in fastsimcoal2 support the "Dispersal

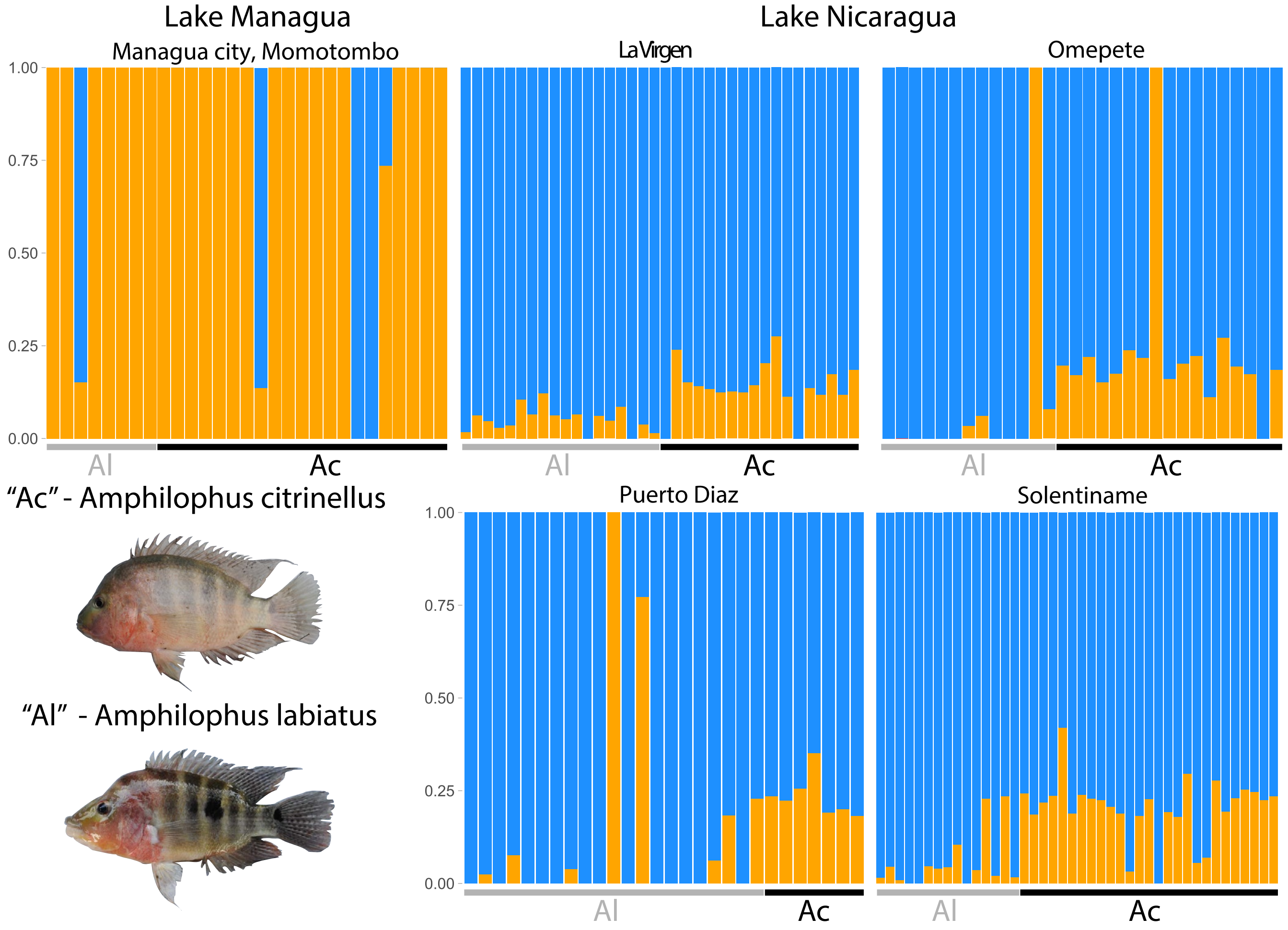


FIGURE 6 ADMIXTURE analysis showing the most supported number of clusters (K2). Specimens from Managua are displayed in the left-most column. Specimens from Lake Nicaragua (La Virgen, Omepete, Puerto Diaz, Solentiname) are displayed on the central and right columns. Each bar represents a different individual "Ac", *A. citrinellus*, "Al", *A. labiatus*), with both species being present at every site

gene flow" (likelihood = −231.1647), "Hybrid origin *A. labiatus* Managua no gene flow" (likelihood = −284.6498) and "All gene flow" (likelihood = −300.7759; Figure 1, Table 3). AIC assessment suggests that "Dispersal gene flow", "Hybrid origin Nicaragua no gene flow" and "Ancestral gene flow" are the best supported models (Figure S15). Notably, both model choice methods indicate that the "Dispersal gene flow" model is the best supported. This scenario suggests that the species in different lakes have a more recent common ancestor with one another than either does with the alternative species in the same lake. Thus, under this model, the species have a common origin but have then undergone gene flow with the alternative ecotype/species when isolated together in each of the two lakes. Similarly, under this model, the estimates for the coalescence time of populations are 1642 generations ago (coalescence between Lake Nicaragua *A. labiatus* and Lake Managua *A. labiatus*), the second occurring 21,816 generations ago (coalescence between Lake Nicaragua *A. citrinellus* and Lake Managua *A. citrinellus*) and the third having occurred 253,208 generations ago (coalescence between the ancestor of *A. labiatus* and *A. citrinellus*). These agree with the estimated ages for the colonization of the lake, approximately 500,000 years ago (Bussing, 1976). From this model, we estimated population sizes of 1,221 individuals for *A. citrinellus* in Lake Nicaragua, 76 individuals for *A. labiatus* in Nicaragua, 261 individuals for *A. citrinellus* in Lake Managua, and 370 individuals for *A. labiatus* in Managua.

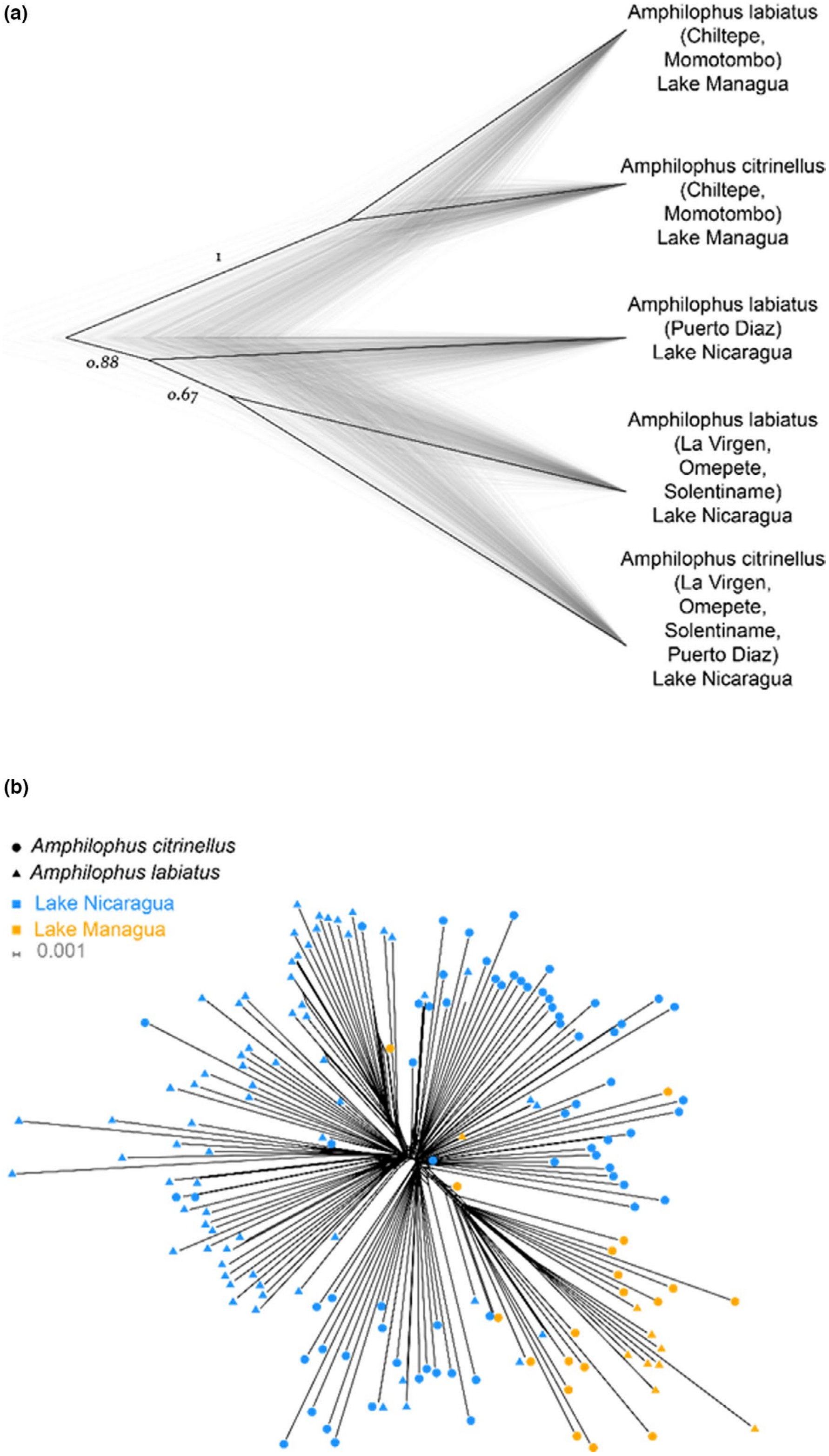
3.9 | Genetic origin of the lip trait

Our attempt to identify loci under selection was unsuccessful. For the five examined genes, we found only evidence for increased divergence in Lake Managua specimens on scaffold 01000607, the one corresponding to the four-and-a-half gene (Figure S16). However, this F_{ST} peak sits ~1.3 Mb away from the gene on this scaffold and thus cannot be taken as evidence for selection without denser SNP-sampling.

4 | DISCUSSION

Across populations, we found remarkable genetic similarity between the large-lipped form *A. labiatus* and the thin-lipped *A.*

FIGURE 7 Phylogenetic analyses of the genomic data (RAD-seq). (a) Phylogenetic tree based on 4000 randomly selected single nucleotide polymorphisms (SNPs). The black line represents a maximum credibility tree obtained from 934 sampled trees (grey trees). In this analysis, we selected five tips (*A. labiatus* and *A. citrinellus* from Lakes Managua and Nicaragua, *A. labiatus* from Puerto Diaz in Lake Nicaragua), based on the principal component analysis (see Figure 4). Posterior support is provided nearby the branches. (b) Phylogenetic network based on the genomic data. Species are represented as triangles (*A. labiatus*) and circles (*A. citrinellus*), and lakes are represented in blue (Lake Nicaragua) and orange (Lake Managua)



citrinellus. Indeed, we found genetic variation suggesting that *A. labiatus* and *A. citrinellus* specimens occurring within the same lake are more closely related to each other than conspecifics in the other lake. This result suggests that there is pervasive admixture and homogenization of genomic differentiation between the poorly differentiated species within both lakes. Our analysis of the evolutionary history of the large-lipped species suggests that this trait has a single origin, most probably in Lake Managua, and the trait then spread via dispersal and adaptive introgression within and between the lakes. Alternatively, the genomic basis for phenotypic divergence may have arisen from standing genetic variation in both lakes and divergence between the species may be relatively recent. In any case, the genomic basis of the phenotypic differences requires further research.

TABLE 1 Mean pairwise F_{ST} estimates for species within and between lakes

	*A. citrinellus*Lake Managua	*A. citrinellus*Lake Nicaragua	*A. labiatus*Lake Managua
*A. citrinellus*Lake Managua	–	–	–
*A. citrinellus*Lake Nicaragua	0.037	–	–
*A. labiatus*Lake Managua	0.020	0.042	–
*A. labiatus*Lake Nicaragua	0.042	0.010	0.055

Notes: All F_{ST} values were significant ($p < .001$) after 110 permutations.

TABLE 2 Effective population size (N_e) estimates. The top panel shows a species-level analysis and the bottom a Nicaragua population-level analysis. Results obtained with the linkage-disequilibrium method included as part of NeEstimator

Lake	*A. citrinellus*	*A. labiatus*
Managua	14	5
Nicaragua	51	46
Locations in Lake Nicaragua		
Ometepe	12	9
La Virgen	14	12
Solentiname	19	11
Puerto Diaz	6	14

Admixture occurs when alleles are incorporated into existing lineages via gene flow and recombination. Admixture can play an important role in adaptation and speciation, for example, constraining these processes via homogenization of the genome or promoting them via the introduction of new variation. Our analysis of the evolutionary history of the large-lipped *A. labiatus* suggests pervasive admixture in this system, but shows no evidence supporting independent evolutionary events (e.g., from de novo mutations). Introgression also explains the best supported demographic model from our SFS inference. Specifically, the favored model, "dispersal gene flow", suggested by likelihood and AIC contradicts our estimated population-level phylogeny. Such a contradiction is expected in cases where extensive admixture and subsequent introgression occurs, generating patterns of reticulated evolution, which we clearly observed in our network analysis. Consequently, our data suggest that the most parsimonious explanation is that large lips, an innovative trophic trait, probably originated once and then spread through both lakes via dispersal, potentially driven by adaptive introgression. An alternative interpretation of the "dispersal gene flow" model is that the species have repeatedly arisen due to standing genetic variation present in both lakes, but that this has occurred so recently that there is a substantial amount of shared polymorphism among the species and lakes. More generally, shared polymorphism can be difficult to distinguish from gene flow (Cruickshank & Hahn, 2014). However, the two best supported models after the dispersal model include a hybrid origin for the species in Lake Managua or substantial gene flow among all species and both lakes. Overall, it seems likely that admixture has occurred throughout the evolutionary history of these fish (see Kautt et al., 2020). In contrast to a recent study by Kautt et al. (2020), in which haplotype statistics suggested that the large-lipped trait originated in Lake Nicaragua, our data, collected from a wide range of locations, suggest that the trait most likely originated in Lake Managua. Specifically, our admixture analysis shows an asymmetric pattern of ancestry, with individuals from Lake Nicaragua having a greater proportion of ancestry from Lake Managua compared to what individuals from Lake Managua have from Lake Nicaragua. Interestingly, this pattern aligns with the topography of the region, with flood waters and temporary rivers typically flowing from Lake Managua into Lake Nicaragua. More broadly, our results complement other recent studies that have found that admixture, introgression, and selection on standing genetic variation can play a role in adaptation to ecological opportunities and adaptive radiation (Marques et al., 2019; McGee et al., 2020; Meier et al., 2017, 2019; Svardal et al., 2019).

Despite a low level of genetic differentiation suggesting that these are poorly differentiated species, we found clear morphological and ecological segregation between species across and between the two lakes. Specifically, *A. labiatus* has larger lips, a narrower head and body, smaller relative gut length, a diet dominated by crustaceans, and significantly different stable isotope values (δ^{13}C). While some of these differences have been observed previously (Astorqui, 1971; Colombo et al., 2013; Klingenberg et al., 2003; Manousaki et al., 2013), our study demonstrates that this ecological and morphological differentiation between the species is widespread throughout the two great lakes. Moreover, we found that *A. labiatus* and *A. citrinellus* differed significantly in their δ^{13}C signatures, but not δ^{13}N, implying that the two species have markedly different diets and consume prey items that originate from different primary producers (probably from different microhabitats), while they nevertheless maintain similar trophic positions in local food webs (Caut et al., 2009). Given the genetic similarity and the history of pervasive admixture between the species (e.g., ~3% of genetic variation occurring between populations), the radiation of the large lip trophic trait appears to be driven by divergent selective pressures associated with habitat use (i.e., rocky patches) and foraging preferences (i.e., crustaceans). Our results suggest that variability in available prey items probably drives selection for differences in trophic morphology, in accordance with recent evidence from other fish systems (Collar et al., 2005; Martin & Wainwright, 2011; Price et al., 2011; Wainwright, 2007). Despite the differences in morphology and ecology between the species, we could only identify one outlier region on the probed scaffolds, but it seems to be far away (~1.3 Mb) from the gene of interest. It is likely that the few genomic markers

TABLE 3 Results of the demographic analyses. Information on population size for each considered population, time of coalescence, and average likelihood per model

Model name	Population size				Coalescence (in generation times)			Average likelihood
	A. citrinellus Lake Nicaragua	*A. labiatus* Lake Nicaragua	*A. citrinellus* Lake Managua	*A. labiatus* Lake Managua	First coalescent event	Second coalescent event	Third coalescent event	
All gene flow	256	46	1,071	1,323	1,297	6,370	361,598	−300.78
Ancestral gene flow	30	134	49	1,494	409,536	556,937	626,379	−508.28
Dispersal gene flow	1,221	76	261	370	1,642	21,816	253,208	−231.16
Dispersal no gene flow	30	304	1,242	986	466,235	567,384	591,269	−506.49
Hybrid origin *A. labiatus* - Managua gene flow	116	338	578	1,262	831	7,176	406,371	−343.04
Hybrid origin *A. labiatus* - Managua no gene flow	1,367	1,080	1,564	1,675	107	9,134	411,704	−284.65
Hybrid origin *A. labiatus* -Nicaragua gene flow	60	786	156	1,401	5,585	100,926	482,551	−429.34
Hybrid origin *A. labiatus* - Nicaragua no gene flow	129	188	141	1319	330,866	490,286	605,478	−434.36
Managua gene flow	211	884	641	923	144,687	531,669	621,043	−433.36
Modern lineages gene flow	1,056	47	1,030	487	5,947	21,496	559,221	−322.23
Nicaragua gene flow	92	123	393	1,343	2,313	41,033	493,985	−415.73
No gene flow	845	278	1,446	1,700	169	11,170	390,861	−324.33
Pulse of gene flow	139	21	1,078	660	191	11,082	421,609	−403.58
Sympatric gene flow	448	523	1,130	259	10,757	422,289	482,477	−361.31

obtained by RAD-seq, together with the history of pervasive admixture, reduced the power of F_{ST} -based scans of selection to detect divergence in our study.

Interestingly, despite significant differences in many of the measured ecological and morphometric variables, we also observed some overlap between the species in the ecomorphological variable space, and detect a large degree of individual variation within species and locations. This overlap is probably driven by the admixture in this system, while it is also the case that ecological opportunity alone does not always lead to extensive adaptive radiations (Yoder et al., 2010; Stroud & Losos, 2016). For example, ecologically derived selection can often result in only partial genomic divergence between ecotypes (Hendry, 2009; Hendry et al., 2009). Dispersal to a new environment and the acquisition of a key innovation are important sources of ecological opportunity (Simpson, 1949). The large-lipped *A. labiatus* has been regarded as an evolutionary innovation that increases the adaptive potential to utilize underexploited resources and exploit new trophic niches (Burress, 2015; Machado-Schiaffino et al., 2014; Manousaki et al., 2013; Oliver & Arnegard, 2010). Upon encountering a novel environment, organisms may experience strong directional selection, for example, the three-spined stickleback (*Gasterosteus aculeatus*) has repeatedly adapted, most likely from standing genetic variation, to freshwater conditions (Albert et al., 2008; Barrett & Schluter, 2008; Ólafsdóttir et al., 2007). Similarly, *A. citrinellus* and *A. labiatus* are consistently found associated with different substrates, the former typically associated with sandy habitats and the latter mostly found in rocky zones (M. Barluenga, personal observation).

Interspecific competition is widely considered an important driver of adaptation to underexploited ecological opportunities (Schluter, 2000; Yoder et al., 2010). Relative to other large lakes where cichlids have radiated rapidly, the diversity of fish species in the two Nicaraguan great lakes is comparably lower (Bussing, 1976), although we lack the complete picture on the distribution and abundance of many species in this region (Härer et al., 2017). In combination with relatively low levels of interspecific competition, high admixture and introgression between the species could homogenize much of the genome, overcoming reproductive isolation as a result of divergent selection, relaxing selection on niche specialization and, ultimately, diversification (Barluenga et al., 2006). A common feature of adaptive radiations is the speed at which adaptations to ecological opportunities arise (Barrett & Schluter, 2008). However, less well known are the longer-term consequences of ecological opportunity for diversification, speciation, and adaptive radiation (Yoder et al., 2010). In accordance with previous studies (Barluenga & Meyer, 2010; Kautt et al., 2012, 2020; Machado-Schiaffino et al., 2017), we demonstrate genetic similarities between thin and large-lipped species, with our genetic results suggesting that these species are poorly differentiated or are at the early stages of divergence (Geiger et al., 2010; Kautt et al., 2020). The Nicaraguan great lakes formed relatively recently (within ~1 million years) and evidence suggests that cichlid fishes have colonized these lakes only around 500,000 years ago (Bussing, 1976). Our demographic analyses are roughly in agreement with the above numbers and show that the species complex speciated around 253,000 generations ago (with generation time presumed to be ~2 years). Nonetheless, using whole-genome sequences, Kautt et al. (2020) estimated the age of the species complex to be around 16,700 years old, coincident with a volcanic eruption in the Lake Managua area. These discrepancies between the different studies highlight that the evolution of the species complex may have been shaped by a history of bottlenecks and introgression and requires further investigation. In any case, the close proximity of the two great lakes and periodic flooding events (Barlow & Munsey, 1976) have probably resulted in migrants between the lakes, in addition to continued gene flow between the species, and on-going selection of trophic traits. The age of the radiation, together with the wide genomic differentiation seen in the smaller crater lakes (Kautt et al., 2020), leads us to conclude that differentiation and reproductive isolation are incomplete in the great lakes, despite the tendency for assortative mating within species (Machado-Schiaffino et al., 2017). It may be that ecological conditions, that is, the size of the greater lakes, the distribution of rocky patches, as well as regular dispersal between lakes, are leading admixture and introgression toward “stalled divergence” (sensu Kautt et al., 2020).

Interestingly, morphology and diet varied between the locations within the lakes. For example, both species from La Virgen tended to differ from their conspecifics elsewhere in Lake Nicaragua in both morphology and diet (according to $\delta^{13}C$ signature). Resource based divergence can occur quickly (Barrett & Schluter, 2008; Elmer, Lehtonen, et al., 2010; Hendry et al., 2007; Sowersby et al., 2020), leading to resource partitioning and the evolution of convergent trophic morphologies (Rüber et al., 1999). However, we did not observe any genetic differentiation with La Virgen and other locations. Furthermore, the ecomorphological differences we observed in La Virgen were relatively small, and variation in lip shape across locations was overall very low. Likewise, we found that *A. labiatus* from Puerto Diaz in Lake Nicaragua had a high degree of genetic differentiation from other *A. labiatus*. The genetic differentiation between *A. labiatus* from Puerto Diaz and elsewhere in Lake Nicaragua was not coupled with detectable differences in ecology or morphology. Therefore, we cannot rule out the possibility that genetic variation from an unknown source is entering the Puerto Diaz population, leading to substantial difference in allele frequencies compared to other populations. Given Puerto Diaz's position on the north east coast of the lake, contact could be occurring with *Amphilophus* populations inhabiting nearby rivers draining into Lake Nicaragua (e.g., River Malacatoya), or with recent migrants coming from Lake Managua into Lake Nicaragua (via the Tipitapa River, which flows intermittently). Indeed, we have presented one of the first detailed genomic analysis of the source populations for crater lake radiations, and our results, together with previous studies assessing crater lakes, demonstrate that admixture is pervasive in the *Amphilophus* system (Barluenga & Meyer, 2010; Geiger et al., 2010; Kautt et al., 2016; Machado-Schiaffino et al., 2017; Kautt et al., 2020). Moreover, comparable lacustrine/riverine introgression has also been shown in Cameroonian cichlids (Poelstra et al., 2018) and in other fish species (Berner et al., 2009; Ravinet et al., 2013). Finally, we note that, within *A. citrinellus*, the distribution of lower pharyngeal jaw types (molariform or papilliform) appears bimodal and associated with preferences for specific food items (snails vs. algae), as well as with differences in body morphology (M. Barluenga unpublished data). This putative specialization within *A. citrinellus* has probably increased variation in the ecomorphological space within locations detected in the current study, and further research may well identify this variation within *A. citrinellus* as another example of diversification via ecological adaptation in the *Amphilophus* radiation.

4.1 | Conclusion

Our results suggest pervasive admixture in this system, with the parallel evolution of the large-lipped species from thin-lipped populations unlikely occurring due to separate de novo mutations. Instead, adaptive introgression and repeated selection from standing genetic variation is more likely, occurring when ecological opportunity is present (sensu Meier et al., 2017; Dean et al., 2019; Marques et al., 2019). This scenario implies that the level of convergence in the *Amphilophus* system may be lower than previously estimated, particularly if dispersal or gene flow among these populations has been infrequent and selection on adaptive alleles has generated genetic differentiation. More generally, our results reveal a complicated evolutionary history in this system and highlight the importance of considering admixture when investigating the evolution of repeated phenotypic traits.

ACKNOWLEDGEMENTS

Nataly Hidalgo Aranzamendi translated ethics and permit documents. Carlos Garita Alvarado and several local fishermen helped with specimen collection in the field. Melinda Hofmann, Gema Aguilera García, and Christian Beisel assisted in the laboratory, with the preparation of samples for analysis. Several undergraduate students at Monash University helped prepare images for geometric morphometric analysis. Walter Salzburger, Marius Rösti, Bernd Egger, and Alysha Heimberg provided advice in regard to genetic analysis. Staff at the DNA Database of Japan (DDBJ) assisted with the use of the NIG supercomputer to conduct population genomic analyses. The Norwegian National E-Infrastructure provided high-performance computing and storage via the respective projects NN9408K and NS9408K. This is the NHM Evolutionary Genomics laboratory's contribution number 27. Funding was also provided by a Monash University travel grant (to WS) and from the SYNTHESYS Project (http://www.synthesys.info/) financed by the European Community Research Infrastructure Action under the FP7 "Capacities" Program (to TKL and MB).

AUTHOR CONTRIBUTIONS

Will Sowersby, José Cerca, Bob B. M. Wong, Topi K. Lehtonen, Mark Ravinet and Marta Barluenga designed the study. Will Sowersby, Bob B. M. Wong, Topi K. Lehtonen and Marta Barluenga collected the data. Will Sowersby, José Cerca, Mariana Leal-Cardín and Mark Ravinet analysed the data. Will Sowersby and José Cerca wrote the first draft and all authors contributed to subsequent drafts and approved the final manuscript.

DATA AVAILABILITY STATEMENT

All ecological and morphological data have been submitted to the Dryad Digital Repository (Sowersby et al., 2021; doi.org/10.5061/dryad.q83bk3jj7). Genomic data have been submitted to the European Nucleotide Archive (project accession no. PRJEB46748). Scripts and code used for the genomic analyses have been made available at https://github.com/jcerca/Papers/tree/main/amphilophus_MolEcol

ORCID

Will Sowersby https://orcid.org/0000-0002-9774-5935
José Cerca https://orcid.org/0000-0001-7788-4367
Bob B. M. Wong https://orcid.org/0000-0001-9352-6500
Topi K. Lehtonen https://orcid.org/0000-0002-1372-9509
David G. Chapple https://orcid.org/0000-0002-7720-6280
Mariana Leal-Cardín https://orcid.org/0000-0001-5640-6366
Mark Ravinet https://orcid.org/0000-0002-2841-1798

REFERENCES

Albert, A. Y. K., Sawaya, S., Vines, T. H., Knecht, A. K., Miller, C. T., Summers, B. R., Balabhadra, S., Kingsley, D. M., & Schluter, D. (2008). The genetics of adaptive shape shift in stickleback: Pleiotropy and effect size. *Evolution*, *62*, 76–85.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109.

Andrews, S. (2010) fastqc: A Quality Control Tool for High Throughput Sequence Data [Online]. Retrieved from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arbogast, B. S., Drovetski, S. V., Curry, R. L., Boag, P. T., Seutin, G., Grant, P. R., Grant, B. R., & Anderson, D. J. (2006). The origin and diversification of Galapagos mockingbirds. *Evolution*, *60*, 370–382. https://doi.org/10.1111/j.0014-3820.2006.tb01113.x.

Arendt, J., & Reznick, D. (2008). Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in Ecology and Evolution*, *23*, 26–32. https://doi.org/10.1016/j.tree.2007.09.011.

Arnegard, M. E., & Snoeks, J. (2001). New three-spotted cichlid species with hypertrophied lips (Teleostei: Cichlidae) from the deep waters of Lake Malaŵi/Nyasa, Africa. *Copeia*, *2001*, 705–717. https://doi.org/10.1643/0045-8511(2001)001(0705:NTSCSW)2.0.CO;2.

Astorqui, I. (1971). Peces de la cuenca de los grandes lagos de Nicaragua. *Revista De Biología Tropical*, *19*, 7–57.

Barlow, G. W. (1976). The midas cichlid in Nicaragua. In T. B. Thorson (Ed.), *Investigations of the Ichthyofauna of Nicaraguan Lakes*. University of Nebraska Press.

Barlow, G. W., & Munsey, J. W. (1976). The Red Devil-midas-arrow cichlid species complex in Nicaragua. In T. B. Thorson (Ed.), *Investigations of the Ichthyofauna of Nicaraguan Lakes*. University of Nebraska Press.

Barluenga, M., & Meyer, A. (2004). The Midas cichlid species complex: incipient sympatric speciation in Nicaraguan cichlid fishes? *Molecular Ecology*, *13*, 2061–2076. https://doi.org/10.1111/j.1365-294X.2004.02211.x.

Barluenga, M., & Meyer, A. (2010). Phylogeography, colonization and population history of the Midas cichlid species complex (*Amphilophus* spp.) in the Nicaraguan crater lakes. *BMC Evolutionary Biology*, *10*, 326. https://doi.org/10.1186/1471-2148-10-326.

Barluenga, M., Stölting, K. N., Salzburger, W., Muschick, M., & Meyer, A. (2006). Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, *439*, 719–723. https://doi.org/10.1038/nature04325.

Barrett, R. D. H., & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, *23*, 38–44. https://doi.org/10.1016/j.tree.2007.09.008.

Baumgarten, L., Machado-Schiaffino, G., Henning, F., & Meyer, A. (2015). What big lips are good for: on the adaptive function of repeatedly evolved hypertrophied lips of cichlid fishes. *Biological Journal of the Linnean Society*, *115*, 448–455. https://doi.org/10.1111/bij.12502.

Bell, M. A. (1974). *Reduction and Loss of the Pelvic Girdle in Gasterosteus (Pisces): A Case of Parallel Evolution*. Natural History Museum.

Bernatchez, L. (2004). Ecological theory of adaptive radiation: Empirical assessment from coregonine fishes (Salmoniformes). In A. P. Hendry, & S. C. Stearns (Eds.), *Evolution illuminated: Salmon and their relatives* (pp. 175–207). Oxford University Press.

Berner, D., Adams, D. C., Grandchamp, A.-C., & Hendry, A. P. (2008). Natural selection drives patterns of lake–stream divergence in stickleback foraging morphology. *Journal of Evolutionary Biology*, *21*, 1653–1665. https://doi.org/10.1111/j.1420-9101.2008.01583.x.

Berner, D., Grandchamp, A.-C., & Hendry, A. P. (2009). Variable progress toward ecological speciation in parapatry: Stickleback across eight lake-stream transitions. *Evolution*, *63*, 1740–1753. https://doi.org/10.1111/j.1558-5646.2009.00665.x.

Berner, D., & Salzburger, W. (2015). The genomics of organismal diversification illuminated by adaptive radiations. *Trends in Genetics: TIG*, *31*, 491–499. https://doi.org/10.1016/j.tig.2015.07.002.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, *10*, e1003537. https://doi.org/10.1371/journal.pcbi.1003537.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, *29*, 1917–1932. https://doi.org/10.1093/molbev/mss086.

Burress, E. D. (2015). Cichlid fishes as models of ecological diversification: patterns, mechanisms, and consequences. *Hydrobiologia*, *748*, 7–27. https://doi.org/10.1007/s10750-014-1960-z.

Bussing, W. (1976). Geographic distribution of the San Juan ichthyofauna of Central America with remarks on its origin and ecology. In T. B. Thorson (Ed.), *Investigations of the Ichthyofauna of Nicaraguan Lakes*. University of Nebraska Press.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). STACKS: Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*, 171–182.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). STACKS: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140. https://doi.org/10.1111/mec.12354.

Caut, S., Angulo, E., & Courchamp, F. (2009). Variation in discrimination factors (Δ15N and Δ13C): The effect of diet isotopic values and applications for diet reconstruction. *Journal of Applied Ecology*, *46*, 443–453.

Cerca, J., Maurstad, M. F., Rochette, N. C., Rivera-Colón, A. G., Rayamajhi, N., Catchen, J. M., & Struck, T. H. (2021). Removing the bad apples: a simple bioinformatic method to improve loci-recovery in de novo RADseq data for non-model organisms. *Methods in Ecology & Evolution*, *12*, 805–817. https://doi.org/10.1111/2041-210X.13562.

Cerca, J., Rivera-Colón, A., Ferreira, M. S., Ravinet, M., Nowak, M. D., Catchen, J. M., & Struck, T. H. (2020). Incomplete lineage sorting and ancient admixture, and speciation without morphological change in ghost-worm cryptic species. *PeerJ*, *9*, e10896. https://peerj.com/articles/10896/

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, *4*, 7. https://doi.org/10.1186/s13742-015-0047-8.

Collar, D. C., Near, T. J., & Wainwright, P. C. (2005). Comparative analysis of morphological diversity: Does disparity accumulate at the same rate in two lineages of centrarchid fishes? *Evolution*, *59*, 1783–1794. https://doi.org/10.1111/j.0014-3820.2005.tb01826.x.

Colombo, M., Diepeveen, E. T., Muschick, M., Santos, M. E., Indermaur, A., Boileau, N., Barluenga, M., & Salzburger, W. (2013). The ecological and genetic basis of convergent thick-lipped phenotypes in cichlid fishes. *Molecular Ecology*, *22*, 670–684. https://doi.org/10.1111/mec.12029.

Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal Jr, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R. M., Schluter, D., & Kingsley, D. M. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*, *307*, 1928–1933. https://doi.org/10.1126/science.1107239.

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sinauer Associates Inc.

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*, 3133–3157. https://doi.org/10.1111/mec.12796.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Dean, L. L., Magalhaes, I. S., Foote, A., D'Agostino, D., McGowan, S., & MacColl, A. D. C. (2019). Admixture between ancient lineages, selection, and the formation of sympatric stickleback species-pairs. *Molecular Biology and Evolution*, *36*, 2481–2497. https://doi.org/10.1093/molbev/msz161.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*, 491–498. https://doi.org/10.1038/ng.806.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NEESTIMATOR v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, *14*, 209–214. https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.12157

Doebeli, M., Dieckmann, U., Metz, J. A. J., & Tautz, D. (2005). What we have also learned: adaptive speciation is theoretically plausible. *Evolution*, *59*, 691–695. https://doi.org/10.1554/04-154.

Elmer, K. R., Fan, S., Kusche, H., Luise Spreitzer, M., Kautt, A. F., Franchini, P., & Meyer, A. (2014). Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nature Communications*, *5*, 5168. https://doi.org/10.1038/ncomms6168.

Elmer, K. R., Kusche, H., Lehtonen, T. K., & Meyer, A. (2010). Local variation and parallel evolution: morphological and genetic diversity across a species complex of neotropical crater lake cichlid fishes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *365*, 1763–1782. https://doi.org/10.1098/rstb.2009.0271.

Elmer, K. R., Lehtonen, T. K., Kautt, A. F., Harrod, C., & Meyer, A. (2010). Rapid sympatric ecological differentiation of crater lake cichlid fishes within historic times. *BMC Biology*, *8*, 60. https://doi.org/10.1186/1741-7007-8-60.

Elmer, K. R., Lehtonen, T. K., & Meyer, A. (2009). Color assortative mating contributes to sympatric divergence of Neotropical cichlid fish. *Evolution*, *63*, 2750–2757. https://doi.org/10.1111/j.1558-5646.2009.00736.x.

Enciso-Romero, J., Pardo-Díaz, C., Martin, S. H., Arias, C. F., Linares, M., McMillan, W. O., Jiggins, C. D., & Salazar, C. (2017). Evolution of novel mimicry rings facilitated by adaptive introgression in tropical butterflies. *Molecular Ecology*, *26*, 5160–5172. https://doi.org/10.1111/mec.14277.

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLOS Genetics*, *9*, e1003905. https://doi.org/10.1371/journal.pgen.1003905.

Excoffier, L., & Lischer, H. E. L. (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. https://doi.org/10.1111/j.1755-0998.2010.02847.x.

Futuyma, D. J. (1998). *Evolution, Third Edition*, 3rd ed. Sinauer Associates Inc.

Garrison, E. (2016). VCFLIB, *a simple C++ library for parsing and manipulating VCF files*. Retrieved from https://github.com/vcflib/vcflib

Gaunt, T. R., Rodríguez, S., & Day, I. N. (2007). Cubic exact solutions for the estimation of pairwise haplotype frequencies: implications for linkage disequilibrium analyses and a web tool "CubeX". *BMC Bioinformatics*, *8*, 428. https://doi.org/10.1186/1471-2105-8-428.

Gavrilets, S., & Losos, J. B. (2009). Adaptive radiation: contrasting theory with data. *Science*, *323*, 732–737. https://doi.org/10.1126/science.1157966.

Geiger, M. F., McCrary, J. K., & Schliewen, U. K. (2010). Not a simple case – A first comprehensive phylogenetic hypothesis for the Midas cichlid complex in Nicaragua (Teleostei: Cichlidae: Amphilophus). *Molecular Phylogenetics and Evolution*, *56*, 1011–1024. https://doi.org/10.1016/j.ympev.2010.05.015.

Gillespie, R. (2004). Community assembly through adaptive radiation in Hawaiian spiders. *Science*, *303*, 356–359. https://doi.org/10.1126/science.1091875.

Gillespie, R. G., Bennett, G. M., De Meester, L., Feder, J. L., Fleischer, R. C., Harmon, L. J., Hendry, A. P., Knope, M. L., Mallet, J., Martin, C., Parent, C. E., Patton, A. H., Pfennig, K. S., Rubinoff, D., Schluter, D., Seehausen, O., Shaw, K. L., Stacy, E., Stervander, M., … Wogan, G. O. U. (2020). Comparing adaptive radiations across space, time, and taxa. *Journal of Heredity*, *111*, 1–20. https://doi.org/10.1093/jhered/esz064.

Grant, P. R., & Grant, B. (2008). *How and why species multiply: The radiation of Darwin's Finches*. Princeton University Press.

Härer, A., Torres-Dowdall, J., & Meyer, A. (2017). The imperiled fish fauna in the Nicaragua Canal zone. *Conservation Biology*, *31*, 86–95. https://doi.org/10.1111/cobi.12768

Hendry, A. P. (2009). Speciation. *Nature*, *458*, 162–164. https://doi.org/10.1038/458162a.

Hendry, A. P., Huber, S. K., De León, L. F., Herrel, A., & Podos, J. (2009). Disruptive selection in a bimodal population of Darwin's finches. *Proceedings of the Royal Society B: Biological Sciences*, 276753–276759. http://doi.org/10.1098/rspb.2008.1321.

Hendry, A. P., Nosil, P., & Rieseberg, L. H. (2007). The speed of ecological speciation. *Functional Ecology*, *21*, 455–464. https://doi.org/10.1111/j.1365-2435.2007.01240.x.

Henning, F., Machado-Schiaffino, G., Baumgarten, L., & Meyer, A. (2017). Genetic dissection of adaptive form and function in rapidly-speciating cichlid fishes. *Evolution*, *71*, 1297–1312. https://doi.org/10.1111/evo.13206.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030.

Johannesson, K., Forslund, H., Capetillo, N., Kautsky, L., Johansson, D., Pereyra, R. T., & Råberg, S. (2012). Phenotypic variation in sexually and asexually recruited individuals of the Baltic Sea endemic macroalga *Fucus radicans*: in the field and after growth in a common-garden. *BMC Ecology*, *12*, 2. https://doi.org/10.1186/1472-6785-12-2.

Jombart, T. (2008). ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129.

Jombart, T., & Ahmed, I. (2011). ADEGENET 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071. https://doi.org/10.1093/bioinformatics/btr521.

Kautt, A. F., Elmer, K. R., & Meyer, A. (2012). Genomic signatures of divergent selection and speciation patterns in a 'natural experiment', the young parallel radiations of Nicaraguan crater lake cichlid fishes. *Molecular Ecology*, *21*, 4770–4786. https://doi.org/10.1111/j.1365-294X.2012.05738.x.

Kautt, A. F., Kratochwil, C. F., Nater, A., Machado-Schiaffino, G., Olave, M., Henning, F., Torres-Dowdall, J., Härer, A., Hulsey, C. D., Franchini, P., Pippel, M., Myers, E. W., & Meyer, A. (2020). Contrasting signatures of genomic divergence during sympatric speciation. *Nature*, *588*, 106–111. https://doi.org/10.1038/s41586-020-2845-0.

Kautt, A. F., Machado-Schiaffino, G., & Meyer, A. (2016). Multispecies outcomes of sympatric speciation after admixture with the source population in two radiations of Nicaraguan crater lake cichlids. *PLOS Genetics*, *12*, e1006157. https://doi.org/10.1371/journal.pgen.1006157.

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, *11*, 353–357. https://doi.org/10.1111/j.1755-0998.2010.02924.x.

Klingenberg, C. P., Barluenga, M., & Meyer, A. (2003). Body shape variation in cichlid fishes of the *Amphilophus citrinellus* species complex. *Biological Journal of the Linnean Society*, *80*, 397–408. https://doi.org/10.1046/j.1095-8312.2003.00246.x.

Kocher, T. D., Conroy, J. A., McKaye, K. R., & Stauffer, J. R. (1993). Similar morphologies of cichlid fish in Lakes Tanganyika and Malawi are due to convergence. *Molecular Phylogenetics and Evolution*, *2*, 158–165. https://doi.org/10.1006/mpev.1993.1016.

Konings, A. (2016). *Malawi Cichlids in their natural habitat*, 5th ed. Cichlid Press.

Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv:1303.3997 [q-Bio]. https://arxiv.org/abs/1303.3997

Lischer, H. E. L., & Excoffier, L. (2012). PGDSPIDER: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*, 298–299. https://doi.org/10.1093/bioinformatics/btr642.

Losos, J. B. (2010). Adaptive radiation, ecological opportunity, and evolutionary determinism. American Society of Naturalists E. O. Wilson award address. *American Naturalist*, *175*, 623–639. https://doi.org/10.1086/652433.

Losos, J. B., Jackman, T. R., Larson, A., Queiroz, K., Rodríguez Schettino, L. (1998). Contingency and determinism in replicated adaptive radiations of island lizards. *Science*, *279*, 2115–2118. https://doi.org/10.1126/science.279.5359.2115.

Lovette, I. J., Bermingham, E., & Ricklefs, R. E. (2002). Clade-specific morphological diversification and adaptive radiation in Hawaiian songbirds. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*, 37–42.

Machado-Schiaffino, G., Henning, F., & Meyer, A. (2014). Species-specific differences in adaptive phenotypic plasticity in an ecologically relevant trophic trait: hypertrophic lips in Midas cichlid fishes. *Evolution*, *68*, 2086–2091. https://doi.org/10.1111/evo.12367.

Machado-Schiaffino, G., Kautt, A. F., Torres-Dowdall, J., Baumgarten, L., Henning, F., & Meyer, A. (2017). Incipient speciation driven by hypertrophied lips in Midas cichlids fish? *Molecular Ecology*, *26*, 2348–2362.

Magalhaes, I. S., Ornelas-Garcıa, C. P., Leal-Cardin, M., Ramírez Abásolo, 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.

Mahler, D. L., Ingram, T., Revell, L. J., & Losos, J. B. (2013). Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science*, *341*, 292–295. https://doi.org/10.1126/science.1232392.

Mahler, D. L., Revell, L. J., Glor, R. E., & Losos, J. B. (2010). Ecological opportunity and the rate of morphological evolution in the diversification of greater Antillean anoles. *Evolution*, *64*, 2731–2745. https://doi.org/10.1111/j.1558-5646.2010.01026.x.

Manousaki, T., Hull, P. M., Kusche, H., Machado-Schiaffino, G., Franchini, P., Harrod, C., Elmer, K. R., & Meyer, A. (2013). Parsing parallel evolution: Ecological divergence and differential gene expression in the adaptive radiations of thick-lipped Midas cichlid fishes from Nicaragua. *Molecular Ecology*, *22*, 650–669. https://doi.org/10.1111/mec.12034.

Marques, D. A., Meier, J. I., & Seehausen, O. (2019). A combinatorial view on speciation and adaptive radiation. *Trends in Ecology and Evolution*, *34*, 531–544. https://doi.org/10.1016/j.tree.2019.02.008.

Martin, C. H., & Wainwright, P. C. (2011). Trophic novelty is linked to exceptional rates of morphological diversification in two adaptive radiations of Cyprinodon pupfish. *Evolution*, *65*, 2197–2212. https://doi.org/10.1111/j.1558-5646.2011.01294.x.

McGee, M. D., **Borstein**, S. R., **Meier**, J. I., Marques, D. A., Mwaiko, S., Taabu, A., Kishe, M. A., O'Meara, B., Bruggmann, R., Excoffier, L., & Seehausen, O. (2020). The ecological and genomic basis of explosive adaptive radiation. *Nature*, *586*, 75–79. https://doi.org/10.1038/s41586-020-2652-7.

Meier, J. I., Marques, D. A., Mwaiko, S., Wagner, C. E., Excoffier, L., & Seehausen, O. (2017). Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature Communications*, *8*, 1–11. https://doi.org/10.1038/ncomms14363.

Meier, J. I., Marques, D. A., Wagner, C. E., Excoffier, L., & Seehausen, O. (2018). Genomics of parallel ecological speciation in Lake Victoria cichlids. *Molecular Biology and Evolution*, *35*, 1489–1506. https://doi.org/10.1093/molbev/msy051.

Meier, J. I., Stelkens, R. B., Joyce, D. A., Mwaiko, S., Phiri, N., Schliewen, U. K., Selz, O. M., Wagner, C. E., Katongo, C., & Seehausen, O. (2019). The coincidence of ecological opportunity with hybridization explains rapid adaptive radiation in Lake Mweru cichlid fishes. *Nature Communications*, *10*, 5391.

Meyer, A. (1993). Phylogenetic relationships and evolutionary processes in East African cichlid fishes. *Trends in Ecology and Evolution*, *8*, 279–284. https://doi.org/10.1016/0169-5347(93)90255-N.

Meyer, J. R., & Kassen, R. (2007). The effects of competition and predation on diversification in a model adaptive radiation. *Nature*, *446*, 432–435. https://doi.org/10.1038/nature05599.

Muschick, M., Barluenga, M., Salzburger, W., & Meyer, A. (2011). Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. *BMC Evolutionary Biology*, *11*, 116. https://doi.org/10.1186/1471-2148-11-116.

Muschick, M., Indermaur, A., & Salzburger, W. (2012). Convergent evolution within an adaptive radiation of cichlid fishes. *Current Biology*, *22*, 2362–2368. https://doi.org/10.1016/j.cub.2012.10.048.

Nosil, P., & Reimchen, T. E. (2005). Ecological opportunity and levels of morphological variance within freshwater stickleback populations. *Biological Journal of the Linnean Society*, *86*, 297–308. https://doi.org/10.1111/j.1095-8312.2005.00517.x.

Ólafsdóttir, G. Á., Snorrason, S. S., & Ritchie, M. G. (2007). Postglacial intra-lacustrine divergence of Icelandic threespine stickleback morphs in three neovolcanic lakes. *Journal of Evolutionary Biology*, *20*, 1870–1881. https://doi.org/10.1111/j.1420-9101.2007.01375.x.

Oliver, M. K., & Arnegard, M. E. (2010). A new genus for *Melanochromis labrosus*, a problematic Lake Malawi cichlid with hypertrophied lips (Teleostei: Cichlidae). *Ichthyological Exploration Freshwaters*, *21*, 209–232.

Ortiz, E. M. (2019) vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis (Version v2.0). Zenodo.

Østbye, K., Naesje, T. F., Bernatchez, L. et al (2005). Morphological divergence and origin of sympatric populations of European whitefish (*Coregonus lavaretus* L.) in Lake Femund. *Norway. Journal of Evolutionary Biology*, *18*, 683–702.

Parent, C. E., & Crespi, B. J. (2009). Ecological opportunity in adaptive radiation of Galápagos endemic land snails. *The American Naturalist*, *174*, 898–905. https://doi.org/10.1086/646604.

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. https://doi.org/10.1111/mec.14784.

Price, S. A., Holzman, R., Near, T. J., & Wainwright, P. C. (2011). Coral reefs promote the evolution of morphological diversity and ecological novelty in labrid fishes. *Ecology Letters*, *14*, 462–469. https://doi.org/10.1111/j.1461-0248.2011.01607.x.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rainey, P. B., & Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature*, *394*, 69–72. https://doi.org/10.1038/27900.

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M. A. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, *30*, 1450–1477. https://doi.org/10.1111/jeb.13047.

Ravinet, M., Prodöhl, P. A., & Harrod, C. (2013). Parallel and nonparallel ecological, morphological and genetic divergence in lake–stream stickleback from a single catchment. *Journal of Evolutionary Biology*, *26*, 186–204.

Roesti, M., Salzburger, W., & Berner, D. (2012). Uninformative polymorphisms bias genome scans for signatures of selection. *BMC Evolutionary Biology*, *12*, 94. https://doi.org/10.1186/1471-2148-12-94.

Rohlf, F. J. (2011). *tpsRegr, version 1.41*. Department of Ecology and Evolution. State University of New York at Stony Brook.

Rohlf, F. J. (2012). *tpsDig, version 2.12*. Department of Ecology and Evolution. State University of New York at Stony Brook.

Rüber, L., Verheyen, E., & Meyer, A. (1999). Replicated evolution of trophic specializations in an endemic cichlid fish lineage from Lake Tanganyika. *Proceedings of the National Academy of Sciences*, *96*, 10230–10235. https://doi.org/10.1073/pnas.96.18.10230.

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*, 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x.

Salzburger, W. (2018). Understanding explosive diversification through cichlid fish genomics. *Nature Reviews Genetics*, *19*, 705–717. https://doi.org/10.1038/s41576-018-0043-9.

Salzburger, W., Mack, T., Verheyen, E., & Meyer, A. (2005). Out of Tanganyika: Genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evolutionary Biology*, *5*, 17.

Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford University Press.

Schluter, D., Clifford, E. A., Nemethy, M., & McKinnon, J. S. (2004). Parallel evolution and inheritance of quantitative traits. *The American Naturalist*, *163*, 809–822. https://doi.org/10.1086/383621.

Seehausen, O. (1996). Distribution of and reproductive isolation among color morphs of a rock-dwelling Lake Victoria cichlid (*Haplochromis nyererei*). *Ecology of Freshwater Fish*, *5*, 195–202. https://doi.org/10.1111/j.1600-0633.1996.tb00133.x.

Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *273*, 1987–1998.

Simões, M., Breitkreuz, L., Alvarado, M., Baca, S., Cooper, J. C., Heins, L., Herzog, K., & Lieberman, B. S. (2016). The evolving theory of evolutionary radiations. *Trends in Ecology and Evolution*, *31*, 27–34. https://doi.org/10.1016/j.tree.2015.10.007.

Skulason, S., & Smith, T. B. (1995). Resource polymorphisms in vertebrates. *Trends in Ecology and Evolution*, *10*, 366–370. https://doi.org/10.1016/S0169-5347(00)89135-1.

Sobel, J. M., Chen, G. F., Watt, L. R., & Schemske, D. W. (2010). The biology of speciation. *Evolution*, *64*, 295–315. https://doi.org/10.1111/j.1558-5646.2009.00877.x.

Sowersby, W., Cerca, J., Wong, B. B. M., Lehtonen, T., Chapple, D. G., Leal-Cardín, M., Barluenga, M., & Ravinet, M. (2021). Pervasive admixture and the spread of a large-lipped form in a cichlid fish radiation. *Dryad Digital Repository*. https://doi.org/10.5061/dryad.q83bk3jj7.

Sowersby, W., Lehtonen, T. K., Ravinet, M., Barluenga, M., & Wong, B. B. M. (2020). Resource trait specialisation in an introduced fish population with reduced genetic diversity. *Biological Invasions*, *22*, 2447–2460. https://doi.org/10.1007/s10530-020-02264-y.

Stern, D. L. (2013). The genetic causes of convergent evolution. *Nature Reviews Genetics*, *14*, 751–764. https://doi.org/10.1038/nrg3483.

Stroud, J. T., & Losos, J. B. (2016). Ecological Opportunity and Adaptive Radiation. *Annual Review of Ecology, Evolution, and Systematics*, *1*, 507–532. https://doi.org/10.1146/annurev-ecolsys-121415-032254

Svardal, H., Quah, F. X., Malinsky, M. et al (2019). Ancestral hybridisation facilitated species diversification in the Lake Malawi cichlid fish adaptive radiation. *Molecular Biology and Evolution*, *37*, 1100–1113. https://academic.oup.com/mbe/article/37/4/1100/5671705

Taliun, D., Gamper, J., & Pattaro, C. (2014). Efficient haplotype block recognition of very long and dense genetic sequences. *BMC Bioinformatics*, *15*, 10. https://doi.org/10.1186/1471-2105-15-10.

Wainwright, P. C. (2007). Functional versus morphological diversity in macroevolution. *Annual Review of Ecology, Evolution, and Systematics*, *38*, 381–401. https://doi.org/10.1146/annurev.ecolsys.38.091206.095706.

Walker, J. A., & Bell, M. A. (2000). Net evolutionary trajectories of body shape evolution within a microgeographic radiation of threespine sticklebacks (*Gasterosteus aculeatus*). *Journal of Zoology*, *252*, 293–302. https://doi.org/10.1111/j.1469-7998.2000.tb00624.x

Yoder, J. B., Clancey, E., Des roches, S., Eastman, J. M., Gentry, L., Godsoe, W., Hagey, T. J., Jochimsen, D., Oswald, B. P., Robertson, J., Sarver, B. A. J., Schenk, J. J., Spear, S. F., & Harmon, L. J. (2010). Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, *23*, 1581–1596. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1420-9101.2010.02029.x

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Sowersby, W., Cerca, J., Wong, B. B. M., Lehtonen, T. K., Chapple, D. G., Leal-Cardín, M., Barluenga, M., & Ravinet, M. (2021). Pervasive admixture and the spread of a large-lipped form in a cichlid fish radiation. *Molecular Ecology*, 00, 1–21. https://doi.org/10.1111/mec.16139