

DR. GONZALO MACHADO-SCHIAFFINO (Orcid ID : 0000-0002-4049-3247)

Received Date : 05-Jul-2016
Revised Date : 22-Dec-2016
Accepted Date : 11-Jan-2017
Article type : Original Article

Incipient speciation driven by hypertrophied lips in Midas cichlids fish?

GONZALO MACHADO-SCHIAFFINO*, ANDREAS F. KAUTT^{*,1}, JULIAN TORRES-
DOWDALL^{*,1}, LUKAS BAUMGARTEN^{*,†}, FREDERICO HENNING* and AXEL
MEYER*

**Chair of Zoology and Evolutionary Biology, Department of Biology, University of
Konstanz, Universitätsstrasse 10, 78457, Konstanz, Germany*

Correspondence: Gonzalo Machado-Schiaffino or Axel Meyer, Fax: +49
7531883018; E-mail: gonzalo.machado-schiaffino@uni-konstanz.de or
axel.meyer@uni-konstanz.de

¹These authors contributed equally to this work

Current address:

*†Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, 50829,
Köln, Germany*

This article has been accepted for publication and undergone full peer review but has not
been through the copyediting, typesetting, pagination and proofreading process, which may
lead to differences between this version and the Version of Record. Please cite this article as
doi: 10.1111/mec.14029

This article is protected by copyright. All rights reserved.

Abstract

Sympatric speciation has been debated in evolutionary biology for decades. Although it has gained in acceptance recently, still only a handful of empirical examples are seen as valid (e.g. crater lake cichlids). In this study, we disentangle the role of hypertrophied lips in the repeated adaptive radiations of Nicaraguan crater lake cichlid fish. We assessed the role of disruptive selection and assortative mating during the early stages of divergence and found a functional trade-off in feeding behavior between thick- and thin-lipped ecotypes suggesting that this trait is a target of disruptive selection. Thick-lipped fish perform better on non-evasive prey at the cost of a poorer performance on evasive prey. Using enclosures in the wild, we found that thick-lipped fish perform significantly better in rocky than in sandy habitats. We found almost no mixed pairs during two breeding seasons and hence significant assortative mating. Genetic differentiation between ecotypes seems to be related to the time since colonization, being subtle in L. Masaya (1600 generations ago) and absent in the younger L. Apoyeque (<600 generations ago). Genome-wide differentiation between ecotypes was higher in the old source lakes than in the young crater lakes. Our results suggest that hypertrophied lips might be promoting incipient sympatric speciation through divergent selection (ecological divergence in feeding performance) and non-random mating (assortative mating) in the young Nicaraguan crater lakes. Nonetheless, further manipulative experiments are needed in order to confirm the role of hypertrophied lips as the main cue for assortative mating.

Keywords: Sympatric speciation, crater lakes, functional trade-off, intraspecific ecological divergence

Introduction

The likelihood of speciation in the presence of gene flow has been debated for decades (Mayr 1942; Coyne & Orr 2004; Gavrillets 2004; Bolnick & Fitzpatrick 2007; Fitzpatrick *et al.* 2008; Nosil 2012; Feder *et al.* 2013; Meyer & Kautt 2014). For sympatric speciation by natural selection to happen, typically three main conditions have to be fulfilled: (1) strong disruptive selection on ecotypes, (2) strong assortative mating by ecotypes and (3) the build-up of linkage disequilibrium between the loci responsible for the ecological traits and mate choice (Bolnick & Fitzpatrick 2007).

Although still no agreement has been reached about how often speciation with gene flow occurs, it is generally accepted that divergence in sympatry is facilitated when the same trait that is under divergent natural selection is also involved in assortative mating (Gavrillets 2004; Papadopulos *et al.* 2011; Servedio *et al.* 2011; Martin 2013).

To better understand what roles divergent selection and assortative mating play at the earliest stages of speciation one needs to identify and investigate case studies where both of these processes might be at work. Crater lake cichlid fish are widely regarded as one of the most convincing examples of sympatric speciation (Schliewen *et al.* 1994; Coyne & Orr 2004; Barluenga *et al.* 2006; Malinsky *et al.* 2015; Kautt *et al.* 2016a) but see (Martin *et al.* 2015) and are an ideal system to assess the potential role of adaptive traits during the early stages of divergence in sympatry.

Midas cichlids (*Amphilophus* sp. species complex) inhabiting a number of small, completely isolated volcanic crater lakes located in Western Nicaragua were colonized recently (less than 24 000 years ago) from the great lakes Managua and Nicaragua (Fig. 1A). This group of Neotropical cichlids is particularly variable in ecomorphological traits such as body shape (limnetic/benthic), pharyngeal jaws (papiliform/molariform), and remarkably hypertrophied lips (thick-/thin-lipped) (Barluenga *et al.* 2006; Elmer *et al.* 2010a; Manousaki *et al.* 2013; Machado-Schiaffino *et al.* 2014).

Interestingly, hypertrophied lips have evolved independently and repeatedly in African and Neotropical cichlids (Henning & Meyer 2014). The fact that thick-lipped ecotypes evolved in parallel strongly supports the hypothesis that this trait is an adaptation (*sensu* Losos 2011) for a particularly specialized type of foraging behavior. It has been suggested that the feeding apparatus (hypertrophied lips in combination with narrow and pointed heads) of thick-lipped species enhances the ability to forage in rocky substrates by facilitating the access to crustaceans and fish larvae that hide between rocks (Barlow & Munsey 1976; Kohda & Tanida 1996; Konings 1998; Arnegard & Snoeks 2001; Oliver & Arnegard 2010; Baumgarten *et al.* 2015). Although this had not been properly tested in Neotropical cichlids yet, it appears to be the case for African cichlids (Baumgarten *et al.* 2015).

Within the Midas cichlid species complex, thick- and thin-lipped ecotypes are present in both great lakes (formally described as *A. labiatus* and *A. citrinellus*, respectively) (Barlow 1976), and in the crater lakes Masaya and Apoyeque (Fig. 1A). Frequency of thick-lipped ecotype is much lower (e.g. 5 to 20%) than thin-lipped one in all these lakes, being the highest (around 20%) in crater lake Apoyeque (Elmer *et al.* 2010b). Interestingly, L. Masaya is one of the oldest (6000 years old) and L.

Apoyeque one of youngest (1800 years old) crater lakes in Nicaragua (Kutterolf *et al.* 2007). Morphological differences between thick- and thin-lipped ecotypes are principally found in the mouth region, (Elmer *et al.* 2010b; Manousaki *et al.* 2013) and are associated with diet differences both in African (Colombo *et al.* 2013) and Neotropical (Manousaki *et al.* 2013) cichlids. Based on stomach contents, both studies found that thin-lipped ecotypes predate mostly on fish whereas thick-lipped ecotypes predate preferentially on invertebrates; suggesting a potential trade-off. Moreover, it has been suggested that these niche differences might drive incipient ecological sympatric speciation at least in crater lake Apoyeque (Elmer *et al.* 2010b).

The young age of the Midas cichlid complex and the repeated colonization of the crater lakes at different points in time presents a unique opportunity to assess the role of ecologically relevant traits, such as hypertrophied lips, during different stages of the speciation process. By comparing the potential functional trade-off, the degree of assortative mating, the pattern of genome-wide genetic differentiation and the demographic history between thick- and thin-lipped ecotypes we aimed to disentangle how disruptive selection and assortative mating might promote or constrain sympatric speciation in the young Nicaraguan Midas cichlid species complex.

Materials and methods

Hypertrophied lip size distribution in wild fish

Standardized lateral photographs of wild caught individuals from each lake were taken in order to quantify lip size. 137, 168, 234 and 89 individuals were collected during field trips in L. Managua, L. Nicaragua, L. Apoyeque and L. Masaya, respectively. Protruded lip area (as measured in Machado-Schiaffino *et al.* (2014))

and body area were measured using ImageJ software (Schneider *et al.* 2012) (see Fig. S1F). Hereafter, we refer to normalized lip area as the protruded lip area corrected by body area and all values are expressed in ratios. Hartigans' Dip tests for unimodality and Kernel density plots were carried out in R (R Development Core Team 2014).

Testing feeding performance

Two independent behavior experiments were conducted in order to compare the feeding performance of wild caught thick- and thin-lipped Midas cichlids from Lake Apoyeque when fish are exposed to sessile or evasive prey. For both experiments fish were starved for one day prior to the experiment to keep them at a similar level of satiation. Fish were transferred into a 200L experimental tank one day before the beginning of each experiment for acclimation. Standardized photographs were taken before the experiments in order to measure both lip area, head angle, lip length, standard length and body area for each individual. All measurements were size-corrected by standard length or body area.

As a first experiment, 15 wild caught thick- and 16 thin-lipped juvenile fish (mean standard length = 86.6 mm, SD = 7.1) from Lake Apoyeque were individually exposed to mosquito larvae equidistantly placed (every 3 mm) along an acrylic structure (see Fig. 2A) consisting of a continuously decreasing angle (angle range = 60° to 15°; total height = 16 cm) following Henning *et al.* (2017). Each acrylic device was kept in the tank for at least six hours. After that time, the distance from the base (angle = 60°) of the acrylic device to the furthestmost removed mosquito larvae was recorded as a proxy for the minimum foraging angle for each fish.

For the second experiment, the same 15 wild caught thick- and 15 thin-lipped juvenile fish from Lake Apoyeque were individually exposed to five free-swimming fish fry. Each experimental fish was recorded from above with a video camera for at least 45 minutes. Only trials in which fish were able to catch at least three fry were considered to obtain a better estimates of the average number of attempts. But note that results were qualitatively similar if the minimum number was set to two. After the experiments the number of failures and successful attacks were quantified from the video recordings.

Testing feeding performance in situ: Enclosures

Reciprocal transplant enclosure experiments in the wild have been successfully applied in fish to test fitness of morphologically divergent species in different ecological niches (Hatfield & Schluter 1999; Hendry *et al.* 2002). However, to our knowledge, enclosure experiments have not been extensively used in cichlid fish to test fitness trade-off associated with ecologically relevant traits. In the present study, wild-caught thick- and thin-lipped Midas cichlids from Lake Apoyeque were allocated to enclosures in two well differentiated habitats (rocky and sandy) within the lake in order to test whether the performance between ecotypes (measured as difference in growth) varies depending on the habitat. A total of 40 enclosures were built using chicken-wire mesh. Each enclosure had 1.125 m³ of volume (0.9 x 2.5 x 0.5 meters). All enclosures were attached to the ground (approximately at two meters depth) and equally distributed in the rocky and sandy area. The material employed to build the enclosures was flexible and allowed to fix them in close contact with the ground even in the rocky habitat characterized by crevices among rocks. Single fish were introduced into each enclosure in order to avoid the effect of competition. A total of

ten enclosures per morph (thick- and thin-lipped) and habitat (rocky and sandy) were included in this experiment. Juvenile individuals were chosen in order to ensure that substantial growth was still possible. Size matching (by weight) was performed to ensure that starting weight of the fish in the enclosures was similar for each ecotype and habitat (ANOVA *F-statistic* = 0.1528, *df* = 3, *p* = 0.698; starting weights: overall mean = 25 g, SD = 7.64).

All fish were weighted before the beginning of the experiment and after seven weeks. Standardized pictures were taken for all individuals. Lip area, body area, and standard length were measured. Differences in weight between morph and habitat were tested with non-parametric Kruskal-Wallis tests due to a lack of normality and homoscedasticity of the data even after transformation. Multiple comparisons were performed with Dunn's test implemented in R (R Development Core Team 2014) and *p* values were adjusted using Bonferroni corrections.

Assortative mating

Pair-composition censuses were performed in Lake Apoyeque during two breeding seasons, in December 2013 and 2014. Two divers performed the census independently, starting from the same point of the lake and following opposite directions in order to avoid counting the same nest twice. In total, around 1.5 km (one quarter of the total perimeter of the lake) were covered. All nests with offspring and both parents were included in the analyses. After the identification of a nest, a close examination was carried out in order to identify the phenotypes of both parents. Despite the fact that it was not possible to quantify lip area for the parents, their ecotypes were easily determined without ambiguity.

Additionally, the degree of assortative mating was determined for fish from Lake Nicaragua. Due to the size of the lake and its turbidity, pair-composition censuses in the wild were not possible for this lake. Thus, mate-choice experiments were performed under laboratory conditions. All fish included in this experiment were F1s from multiple crosses of wild-caught individuals. A total of 25 thick- (*A. labiatus*) and 25 thin-lipped (*A. citrinellus*) adult individuals were kept under similar conditions (temperature, light and substrate) in a very large tank (7000 L). In order to reduce the number of potential variables associated with pair formation, only dark morphs from both species were used. After seven days of acclimation fish were allowed to interact, the tank was periodically inspected, and the composition of every formed pair that stayed together for at least three days was recorded. Note that spawning took place within the first three days after pair-formation in most cases. After that period, the pair was removed and replaced with individuals of the same ecotype, size, and sex. Thus, the frequency of ecotypes in the tanks was kept constant during the whole experiment. Standardized photographs from the lateral and dorsal side were taken for all individuals that successfully formed a pair in order to measure the association between male and female lip size.

Genome-wide differentiation

To quantify how genetic variation is partitioned across genomes between ecotypes and lakes, we performed restriction-site associated DNA-sequencing (RAD-seq) based population genomic analyses. Between 19 and 72 wild caught individuals per ecotype (thick- and thin-lipped) and lake (Managua, Nicaragua, Masaya and Apoyeque) were included in the genomic libraries (Table S1, Supporting information). Most of these fish were collected between 2010 and 2013 and are

stored in AM's collection at the University of Konstanz. A small piece of pectoral fin from each sample was used to extract genomic DNA with a Zymo kit.

Double digest RADseq libraries were prepared following Peterson *et al.* (2012) and Kautt *et al.* (2016a). Briefly, 600 ng of genomic DNA per sample was double-digested using the restriction enzymes PstI and MspI for 3 hours at 37 °C. After that, P1 and P2 adapters (Recknagel *et al.* 2013) were ligated to the digested DNA using T4 ligase for 30 minutes at room temperature. A total of 307 individually barcoded samples were pooled in six libraries. Size selection for each library was performed using Pippin Prep technology (Sage Science, Beverly, MA) setting the size-range from 320 to 500bp. Finally, genomic libraries were single-end sequenced (100 bp length) in four lanes on an Illumina HiSeq 2000.

Raw sequence reads were demultiplexed using STACKS v.1.2.9 (Catchen *et al.* 2011; Catchen *et al.* 2013). Only high sequencing quality reads, with correct barcodes and unambiguous RAD site were retained. Demultiplexed reads were mapped to an "in house" anchored *A. citrinellus* reference genome (Elmer *et al.* 2014) using BWA (Li & Durbin 2009). SNPs and genotype calling were performed with STACKS. Loci deviating from HWE or showing more than three polymorphic sites (SNPs) in one population were excluded. Further, loci exhibiting a SNP in the last two positions of a read were excluded, because of an excess number of SNPs at these two positions. A locus was included in the analyses only if it was genotyped in at least 10 individuals (20 alleles) per population. On average $74\,197 \pm 18\,049$ (sd) loci were obtained per individual with a mean coverage of 13.94 ± 4.21 (sd) reads per locus and individual (Table S1, Supporting information).

Population genomic statistics were obtained using the *populations* program implemented in STACKS. Genome-wide F_{st} values were calculated among all pair-

wise combinations of populations using ARLEQUIN software with default settings (Excoffier & Lischer 2010), and the significance was tested using 10 000 permutations. Population structure was investigated with the model-based software ADMIXTURE (Alexander *et al.* 2009) and model-free Principal Component Analyses (PCAs) implemented in the EIGENSOFT package (Patterson *et al.* 2006). Support for the number of genetic clusters was determined based on the implemented cross-validation error technique or Tracy-Widom statistics, respectively. The phylogenetic tree was built using PHYLIP (Felsenstein 1993) based on a distance matrix calculated from the frequency data of 1000 randomly selected markers. Statistical support of nodes was estimated with 1000 bootstrap replicates. Note that only one SNP per locus was used for analyses concerning population structure or the phylogenetic tree to account for non-independence of linked sites. Additionally, individual-based phylogenetic networks were built using SPLITSTREE v.4.13.1 (Huson & Bryant 2006). Individual genotype calls were transformed from variant call format (VCF) to Nexus format using custom scripts and networks were built using the NeighborNet method based on uncorrected P distances.

Demographic inference

Demographic parameters were estimated by simulation and fitting of different demographic models to the information contained in the multidimensional minor site frequency spectrum (SFS) using FASTSIMCOAL2 (Excoffier *et al.* 2013). Briefly, FASTSIMCOAL2 uses coalescent simulations to obtain the expected SFS for a certain specified model and then optimizes each demographic parameter in the model so as to maximize the likelihood. Demographic models can be arbitrarily complex, including a variety of demographic events such as migration, population

size changes, population splits and admixture, and there is no restriction in the number of analyzed populations. The relative fit of different demographic models to the data was evaluated by means of the Akaike Information Criterion (AIC).

Loci presumably located in coding regions were identified via a BLAST search against a compilation of transcriptomic data from various species and tissues of Midas cichlids (Henning *et al.* 2013; Manousaki *et al.* 2013) and excluded. Only one SNP per locus was used for the demographic analyses. The SFS was created as in Kautt *et al.* (2016a). Briefly, data was parsed from VCF files using a custom python script and transformed into the SFS using *daði* (Gutenkunst *et al.* 2009). To account for missing data the SFS was projected down to 25 individuals per lake. The same minimum number of 25 individuals per lake was specified to create the VCF files.

Thus the ratio between monomorphic and polymorphic sites can be correctly inferred and used to convert the parameter estimates into demographic units. Once the presumably best model is chosen the fit to the data can be inferred by using the maximum likelihood point estimates of the model parameters to perform parametric bootstrapping and obtain confidence intervals (Excoffier *et al.* 2013).

Results

Bimodal distribution of hypertrophied lips

The normalized size of lips (lip area corrected by body area) is strongly bimodal in all lake populations except for crater lake Masaya where unimodality (Hartigan's Dip test) could not be rejected probably due to low sample size (Fig. 1B-E). Phenotypic differentiation was more pronounced in the older source lakes (Fig. 1B, C) than in the recently colonized crater lakes (Fig. 1D, E). This pattern is concordant with

theoretical predictions of eco-morphological traits under disruptive selection during the early stages of divergence (Coyne & Orr 2004; Gavrilets 2004).

Disruptive selection acting on lips

Coupled with differences in lips we found pronounced differences in feeding performance – this was tested with wild-caught thick- and thin-lipped ecotypes from Lake Apoyeque (Fig. 2). Thick-lipped fish performed better than thin-lipped fish when food was attached (sessile) and had to be scraped off an acrylic structure with a continuously decreasing angle (Fig. 2A), mimicking crevices in the natural habitat. Although thick- and thin-lipped fish are different both in terms of lips and head shape, feeding performance when food was sessile was significantly associated only with hypertrophied lips (Fig. S1, Supporting information). These findings were validated also in experiments with thick- (*A. labiatus*) and thin-lipped (*A. citrinellus*) fish from great lake Nicaragua (Fig. S2A, Supporting information).

However, wild-caught thin-lipped fish performed significantly better than thick-lipped fish when feeding on evasive prey in the water column (Figs 2B and Fig. S2B, Supporting information), indicating a trade-off. Here, feeding success was negatively associated with lip size.

This trade-off was also assessed in terms of fitness consequences with experiments in the crater lakes: using single-fish enclosures ($n = 40$) in sandy or rocky habitats in crater lake Apoyeque we monitored the change in weight of the fish as a proxy for fitness. Despite the fact that all fish had a similar weight at the beginning of the experiment (ANOVA F -statistic = 0.1528, $df = 3$, $p = 0.698$; starting weights: overall mean = 25 g, $SD = 7.64$), the weight change after seven weeks differed significantly (Kruskal-Wallis chi-squared = 10.4717, $df = 3$, p -value = 0.0149)

in the predicted direction between ecotypes and habitats (Fig. 2C). Although most fish lost weight in the enclosures - probably due to the restricted foraging space -, thick-lipped fish in rocky habitat lost significantly less weight than thick-lipped fish confined to a sandy habitat (Dunn's statistics = -2.540, Bonferroni adjusted $p = 0.033$). Also a trend (Dunn's statistics = -2.199, Bonferroni adjusted $p = 0.083$) towards thick-lipped fish compared with thin-lipped fish was apparent on rocky habitat. Differences between thick- and thin-lipped fish in the enclosures on sandy habitat were not significant (Dunn's statistics = -0.696, Bonferroni adjusted $p = 1.000$), which might be best explained by the observation that evasive prey (e.g. small fish) was free to leave the enclosures, thereby depriving thin-lipped fish from their foraging advantage compared to thick-lipped fish. This is in agreement with the fact that thin-lipped fish predate preferentially on fish compared to thick-lipped fish (Colombo *et al.* 2013; Manousaki *et al.* 2013).

Assortative mating between ecotypes

Given the bimodality and fitness trade-offs of this trait we conducted both field observations and laboratory experiments to determine if ecotypes mate assortatively. Due to the clarity of its water we chose to survey Lake Apoyeque for breeding pairs during two consecutive breeding seasons (2013-2014); the turbid waters of the great lakes and crater lake Masaya precluded us from collecting data on pairs in these lakes. We note that Midas cichlids are seasonally monogamous and pairs defend their brood jointly for weeks. Among a total of 68 breeding pairs, mating was highly significantly non-random (Yates' chi-square = 25.147, $p = 3.46 \times 10^{-6}$) and based on ecotypes (Fig. 3A), supporting the hypothesis that gene flow is restricted between the ecotypes. Due to the fact that thick-lipped fish occur much less frequently (c.a. 5-

20%) than thin-lipped fish in all the lakes where both ecotypes coexist, it would be expected under random mating that most of the pairs including thick-lipped fish would be disassortative; thus the probability of pairing up of two thick-lipped fish is expected to be very low (e.g. 4%, assuming a 20% frequency of thick-lipped fish). However, we found significantly more assortative thick-lipped pairs than expected. The total number of thick-lipped fish included in this survey (11 out of 136, around 8%) was similar to the frequency of this ecotype in the wild (around 20% for crater lake Apoyeque (Elmer *et al.* 2010b)). Interestingly, many of the thin-lipped pairs (20 out of a total of 30 pairs where the kind of substrate was recorded) were found in rocky areas, although they supposedly have a feeding advantage in sandy areas. Thus, both ecotypes tend to breed in rocky areas of the lake suggesting that there is no habitat isolation during breeding season.

In order to validate the strong assortative mating between ecotypes found in the wild, we also performed mate choice experiments in the laboratory in very large tanks with *A. citrinellus* (thin-lipped) and *A. labiatus* (thick-lipped) from Lake Nicaragua. All 25 pairs that formed in these experiments were assortative (Fig. 3B). There was no correlation between lip size (corrected by body area) of the paired individuals within *A. labiatus*, suggesting that the presence of lips alone, rather than their size, might matter most for mate choice (Fig. 3C). Moreover, pair composition was independent of size (weight) for *A. labiatus* and *A. citrinellus* pairs ($r = -0.232$ $p = 0.446$ and $r = -0.207$ $p = 0.518$, respectively).

Recent colonization and weak genome-wide genetic differentiation between ecotypes in the crater lakes

Based on more than 16 740 polymorphic genetic markers, clear genome-wide genetic differentiation was found with each lake population forming a distinct genetic cluster (Figs 1A, and S3A, Supporting information). Using a much smaller genetic data set of microsatellite markers, we had previously suggested that crater lake Masaya was colonized from the nearby great lake Nicaragua (Barluenga & Meyer 2010). Yet, the analysis of this new substantial set of genome-wide markers shows that both crater lakes share a more recent common ancestor with great lake Managua (Fig. 4A). Moreover, individuals from crater lake Masaya tend to cluster closer to great lake Managua (Fig. S3A, Supporting information). Individual-based phylogenetic split networks also suggest that both crater lakes cluster together with great lake Managua and not with Lake Nicaragua (Fig. S4, Supporting information). Altogether, these findings support the notion that we have the interesting opportunity to study the evolution of two independent radiations that were recently, but at different times, founded from the same source population (Lake Managua).

Both crater lakes Apoyeque and Masaya are genetically clearly distinguishable and distinct from Lake Managua (Table S2, Supporting information; Figs 1A, and S4, Supporting information). The genetic differentiation between thick-lipped (*A. labiatus*) and thin-lipped (*A. citrinellus*) species in the source lakes Managua ($F_{ST} = 0.050$, $p < 0.00001$) and Nicaragua ($F_{ST} = 0.019$, $p < 0.00001$) are subtle, but highly significant (see Figs 4B, C, and S3B, C, Supporting information). These findings agree with the strong assortative mating between *A. labiatus* and *A. citrinellus* from Lake Nicaragua that we found in lab experiments (Fig. 3B) and support their status as two biological species. Similarly, we found significant, yet even lower, genome-

wide differentiation between ecotypes in the about 6000 years old crater lake Masaya ($F_{ST} = 0.016$, $p < 0.00001$) (Fig. 4E). However, no genetic differentiation was found between ecotypes in the less than 2000 year old crater lake Apoyeque ($F_{ST} = 0.002$, $p = 0.827$) (Fig. 4D). This lack of genome-wide differentiation is not unexpected at this early stage of speciation without geographical isolation.

The demographic history of the crater lake populations was inferred (Fig. 4F) through coalescent simulations and comparing their fit against the empirical data summarized in the two-dimensional site frequency spectra (SFS) (Excoffier *et al.* 2013). Assuming great lake Managua as the source population, 13 and 17 different models were tested for crater lakes Apoyeque and Masaya, respectively (Table S3, Supporting information).

According to the best model for crater lake Apoyeque the source population (great lake Managua) experienced a population bottleneck about 1870 (95% confidence interval: 1480-2520) generations ago and was reduced to only about 1490 (1040-2290) individuals and it has since been growing exponentially reaching a population size of about 460,270 (0-898,670) individuals at present. The results for the source lakes are similar to Kautt *et al.* (2016a) in which they were used as outgroups for two other crater lake radiations. The confidence intervals around the current population sizes are wide, probably because small differences in the estimated growth rates will lead to a high variance since growth is exponential. Lake Apoyeque was colonized by a small founder population of only about 110 (50-200) individuals around 580 (430-770) generations ago, growing to a current size of 14 720 (1480-32 990) individuals. Continuous migration between the lakes is not supported, but a single admixture event 380 (290-470) generations ago in which the

crater lake population received approximately 16% (8.3%-22.4%) of its gene pool from the source population is strongly supported.

In the best model for crater lake Masaya the estimates for the population dynamics in the source lake (Managua) are similar to above (Fig. 4F). Lake Masaya was colonized only about 1560 (1400-1800) generations ago. Surprisingly, a population size change in L. Masaya is not supported, and it seems thus to have remained relatively stable at about 8610 (7800-9760) individuals throughout its history. Continuous migration from the source into Masaya is not supported, yet Lake Masaya received about 21.0% (14.5%-29.2%) of its gene pool from the source population in an admixture event about 240 (120-400) generations ago. While we consider the model above to be the best model, a different model in which L. Masaya was colonized before the bottleneck in the source lake received a higher support. Yet, considering the geographic isolation of this lake, this model assumes biologically unrealistically high amounts of gene flow (continuous migration of 1.91×10^{-4} and 57.1% admixture event) and a colonization time (6390 generations ago) prior to the formation of the crater lake itself.

In summary, L. Apoyeque was colonized much more recently than L. Masaya and is characterized by a much smaller long-term effective population size. Differences were also visually apparent in a plot of the two-dimensional site frequency spectrum (2D-SFS). While the spectrum of the source population together with L. Apoyeque is characterized by a considerable number of diverging sites, in the case of L. Masaya many more sites fall along the diagonal, indicating only weak signs of divergence (Fig. S6, Supporting information).

Discussion

We find that a conspicuous morphological trait, hypertrophied lips, is associated with genetic differentiation in several repeated radiations of the Nicaraguan Midas cichlid species complex. In several experiments in the field and the laboratory we document a trade-off with respect to feeding behavior between thick- and thin-lipped ecotypes: thick-lipped fish perform better on non-evasive prey attached to narrow angles (“crevices”), but are significantly less successful in feeding on free swimming prey than thin-lipped fish. These findings from laboratory experiments were supported by enclosure studies in crater lake Apoyeque, in which thick-lipped fish performed better in rocky than sandy habitats. That these two types of fish are ecologically distinct was suggested before based on stomach content and stable isotope analyses (Elmer *et al.* 2010b; Manousaki *et al.* 2013), but there was no information on feeding performance. Thus, hypertrophied lips are likely subject to disruptive natural selection. Moreover, our results suggest that this trait might also play a role in assortative mating or is tightly linked to a trait resulting in assortative mating, however, further manipulative experiments are needed to confirm this hypothesis.

Disruptive selection acting on lips

Here, we provide evidence that hypertrophied lips, and not other correlated traits, are subject to divergent natural selection. A bimodal distribution of lip size was found in all populations, except for crater lake Masaya where the sample size was small, with more pronounced differences in the older lakes (great lakes Managua and Nicaragua) than the much younger crater lakes (see Fig. 1B-E). A trade-off in feeding performance was found between thick- and thin-lipped fish: the former are more successful in feeding in small crevices (characteristic of rocky habitat) whereas

the latter feed more efficiently on free swimming fish (sandy/open-water habitat) (Fig. 2). Hybrids between thick- and thin-lipped ecotypes exhibiting an intermediate phenotype (Machado-Schiaffino *et al.* 2014) and are expected to perform worse than either of the more extreme phenotypes on their preferred foods/habitats (Rueffler *et al.* 2006). Therefore, hybrids would be expected to have a lower fitness (Rundle & Whitlock 2001). Unfortunately, we could not perform these experiments on hybrids with an intermediate phenotype due to logistical reasons. Nonetheless, it seems that in both experiments the performance of thick-lipped fish with relatively small lip sizes was intermediate (Fig. 2A-B).

One crucial advantage of hypertrophied lips is that they increase suction power during foraging by sealing cracks and grooves in rocky substrate (Barlow & Munsey 1976; Konings 1998; Oliver & Arnegard 2010; Baumgarten *et al.* 2015). We show that the benefits of hypertrophied lips also engender costs when feeding in the open-water (i.e. are subject to a trade-off; Fig. 2); explaining the pronounced diet differences we previously found where thin-lipped fish tend to prey more on evasive prey (e.g. fish) than thick-lipped individuals (e.g. hard-shelled invertebrates) (Colombo *et al.* 2013; Manousaki *et al.* 2013). The kinetics of the buccal protrusion during prey capture might differ between these ecotypes and would explain the differences in predation success (see also (Holzman *et al.* 2012)). Hypertrophied lips might also be constraining the size and shape of the mouth in thick-lipped fish and create more turbulence in the water, thereby decreasing feeding success (Skorczewski *et al.* 2012) on fish (Fig. 2B). Preliminary analyses furthermore suggest that *A. labiatus* and *A. citrinellus* also differ strongly in mouth shape, due to a difference in gape height, which is in turn caused by differences in the ascending and descending process of the premaxillary bone (Fig. S5, Supporting information).

It would be expected that a more planar and circular mouth shape, as is found in *A. citrinellus*, increases feeding performance by maximizing the capacity of suction feeders to exert hydrodynamic forces on the prey (Skorczewski *et al.* 2012). Overall, thick-lipped fish might increase suction power by sealing cracks and grooves when their mouth is firmly attached to the rocky substrate, meanwhile, thin-lipped fish might increase suction forces when the mouth is not attached to any substrate as is the case when they feed in open water.

Transplant experiments in the field have been shown to be a successful way to test the role of natural selection in divergent environments (Schluter 1995; Soria-Carrasco *et al.* 2014). Using enclosures in the wild (crater lake Apoyeque) we found a trend with thick-lipped fish performing better in the rocky than in the sandy habitat (Fig. 2C). To our knowledge, this is one of the first studies using reciprocal transplant enclosure experiments to test for a fitness trade-off associated with an ecologically relevant trait in cichlid fish. Despite the fact that loss in weight was detected for several enclosures, the weight change after seven weeks differed significantly in the predicted direction between thick-lipped fish confined to the rocky and sandy habitat. It is known that the density of prey is significantly reduced in enclosures (Schluter 1995). However, we decided not to provide supplementary diet in order to avoid any potential bias in growth that might affect our results. Moreover, Midas cichlids growth is relatively slow and, thus, the duration of the experiment might not have been long enough to detect significant growth in all experimental individuals.

Reproductive isolation: assortative mating

The evolution of non-random mating is crucial in speciation driven by selection for local adaptation in the presence of gene flow (Schluter 2000). Assortative mating based on colour has been reported in South American cichlids (Ready *et al.* 2006). Similarly, non-random mating based on a color polymorphism (gold-dark) that is present in some, but not all, species of Midas cichlids has been known for some time (Elmer *et al.* 2009). This suggests that active mate choice might be one of the mechanisms driving divergence in this young system.

We observed that both ecotypes are present at a similar depth in sandy and rocky habitat during the breeding season in Lake Apoyeque and several thin-lipped pairs were found in both habitats. This implies that “like-pairs” are formed independently of the respective preferred feeding habitats. Moreover, laboratory mate choice experiments demonstrated complete assortative mating between *A. citrinellus* and *A. labiatus* (Fig. 3B) suggesting that mate choice is due to direct preference, rather than differences in breeding habitat choice. Not a single disassortative pair was found in the laboratory experiments, suggesting that prezygotic isolating mechanisms predominantly prevent hybridization. A panel of F₂ individuals from a forced cross of thick- and thin-lipped fish suggests, however, that there might be also some postzygotic barriers, since we found that a large number (around 10-20%) of F₂ individuals showed some degree of malformation, particularly in their mouths (data not shown). Although thick- and thin-lipped species from great lake Nicaragua mate completely assortatively in the lab, the significant but relatively low level of genetic differentiation suggests a certain level of gene flow between these species. This is in concordance with previously documented low levels of

genetic differentiation between *A. citrinellus* and *A. labiatus* based on different molecular markers (Barluenga & Meyer 2010; Kautt *et al.* 2012). Although clear evidence for assortative mating between thick- and thin-lipped ecotypes was found both in the wild and in lab experiments, we cannot completely rule out that other visual, behavioural or chemical cues are involved in pair formation. Thus, further manipulative experiments should be performed to test whether hypertrophied lips act as the main mating cue.

Genomic fingerprints of divergence in sympatry

Based on population genomics approaches one can gauge how many genomic regions might contribute to the evolution of reproductive isolation and hence speciation (Ellegren *et al.* 2012; Jones *et al.* 2012; Feder *et al.* 2013). The patterns of genome-wide genetic differentiation that we detected between both ecotypes support a scenario of ecological speciation with gene flow under sympatric conditions (Feder *et al.* 2012; Feder *et al.* 2013); only very few regions of the genome show signs of differentiation between ecotypes in the very recently colonized crater lakes (“L-shaped” frequency distribution of genetic differentiation across loci with most loci having low F_{st} values) compared to the older Lakes Nicaragua and Managua (Fig. S3B, Supporting information). Also the demographic analyses, assuming great lake Managua as the source population, suggested that both crater lakes were colonized very recently (580 and 1600 generations for L. Apoyeque and L. Masaya, respectively). Despite the fact that only a small number of individuals were estimated to have colonized the crater lakes (at least L. Apoyeque) and the much lower frequency of thick-lipped fish in the great lakes (e.g. around 5%), it is likely that thick-lipped fish colonized the crater lakes from the great lakes. Thus,

we think that the presence of the thick-lipped ecotype in the crater lakes is due to the retention of ancestral standing genetic variation, rather than due to *de novo* mutations. However, it seems evident that a strong homogenization event has occurred after the colonization, erasing the genetic differentiation already present between the ecotypes in the great lakes. Only by elucidating the genetic basis of this ecologically relevant trait, would it be possible to test whether hypertrophied lips originated *de novo* in the crater lakes or from standing genetic variation present in the great lakes.

Thick-lipped ecotype could also have been brought into the crater lake by a secondary invasion from the source lakes. In this regard, our best demographic models provide evidence for admixture events in both crater lakes. We note, however, that we used the thin-lipped *A. citrinellus* from L. Managua as a source population in these models. Our rationale for this was that the site frequency spectra of both species (*A. citrinellus* and *A. labiatus*) are almost identical and our data thus was not powerful enough to tease apart whether the colonization and subsequent admixture happened by either one or a mix of both species (see also Kautt *et al.* 2016a). In any case, together with recent evidence for admixture events in Nicaraguan crater lakes Apoyo, Xiloá (Kautt *et al.* 2016a), and Asososca Managua (Kautt *et al.* 2016b) and Cameroonian crater lakes (Martin *et al.* 2015), this might suggest that crater lakes are not as isolated as they were believed to be. In this regard, fish must have somehow come into the crater lakes in the first place and evidence for reasonably small admixture events indicating secondary waves of colonization are not unreasonable. However, we note that distinguishing between more ancient divergence events and high amounts of gene flow and very recent

divergence with little or no gene flow based on genetic data is challenging (Hey *et al.* 2015) and the validity of admixture events will have to be further tested in the future.

Even though some differentiation for a few markers was found, our results show that there is almost no genome-wide genetic differentiation between the young sympatric ecotypes in crater lake Apoyeque and only very weak genomic differentiation in the relatively older crater lake Masaya. Thus, the lack of genomic differentiation found in Lake Apoyeque could be explained by different factors. Firstly, although it is expected that genome-wide differentiation increases as phenotypic divergence increases (Roesti *et al.* 2012; Gagnaire *et al.* 2013; Seehausen *et al.* 2014), divergence could remain heterogeneous across the genome for a long time due to repeated events of interspecific gene flow even after reproductive isolation has become strong (Neafsey *et al.* 2010; Garrigan *et al.* 2012). In this regard, we have found some disassortative pairs in crater lake Apoyeque (Fig. 3A), suggesting that gene flow and recombination between ecotypes is plausibly generating a homogenizing effect (Felsenstein 1981). Thus, ongoing gene flow between populations might homogenize variation in genomic regions not affected by divergent selection or reproductive isolation (Noor & Bennett 2009; Turner & Hahn 2010). Yet, reduced diversity in genomic islands of speciation should also be considered as an alternative explanation (Cruickshank & Hahn 2014). Secondly, due to the localized genomic effect of disruptive selection, divergence during the early stages of ecological speciation is expected to be greater in genomic regions that harbor key quantitative loci (QTL) than it is in regions that have no effect on the phenotypic divergence of the populations (Via 2009). Therefore, adaptation with gene flow tends to result in concentrated genetic architectures with fewer, larger, and more tightly linked divergent alleles (Yeaman & Whitlock 2011). Although the exact

genetic architecture of hypertrophied lips remains unknown, ongoing projects in our laboratory provide evidence that only a few major QTL are involved in this trait in Midas cichlids: we find intermediate phenotypes of F_1 hybrids (Machado-Schiaffino *et al.* 2014) and a high proportion of F_2 individuals that almost completely recover both parental phenotypes (Fig. S7, Supporting information). Thus, it is very probable that due to the reduction of genome complexity associated with RAD-sequencing, considering just a fraction of the whole genome, no markers closely linked to the QTL were present in our data set.

Progress toward sympatric speciation

It has been estimated by Thibert-Plante & Gavrillets (2013), assuming a scenario of relatively weak divergent selection, that 10 000 generations might be needed in order to observe strong reproductive isolation (complete non-random mating). This seems to be the case in thick- and thin-lipped fish in the young crater lakes, where we found a clear trade-off between ecotypes, morphological bimodality, and assortative mating, while our enclosure experiments suggest that the differences between ecotypes in different habitats might not be very strong.

Recently, it has been shown for Ejagham tilapia that even in the presence of strong assortative mating, the process of sympatric speciation might be stalled if disruptive selection is weak (morphological unimodality and minimal ecological divergence; (Martin 2013). Despite the fact that we were not able to quantify the strength of disruptive selection (see (Martin 2012)) in the crater lakes, we found clear morphological bimodality and ecological divergence (also see (Colombo *et al.* 2013; Manousaki *et al.* 2013), suggesting that disruptive selection and assortative mating acting in concert might be strong enough to lead to speciation. This process might be

relatively slow and yet incomplete due to the young age of the crater lakes, small effective population sizes, or complexity of the genetic architectures of trophic morphology. The absence of other sister species pairs in the crater lakes precluded us from testing if complete divergence is possible at the same timescale, as was the case in Lake Ejagham (Martin 2013).

Our results suggest that hypertrophied lips might be promoting incipient sympatric speciation through divergent selection (ecological divergence in feeding performance) and non-random mating (assortative mating) in the young Nicaraguan crater lakes. All this opens the possibility of hypertrophied lips acting as a “magic trait” driving speciation in these young crater lake cichlids, although further manipulative experiments are necessary to confirm its role as the main cue in assortative mating. Nicaraguan crater lakes are a unique natural laboratory for examining, “in real time”, how ecologically relevant traits under selection affect the evolution of reproductive isolation during the earliest stages of speciation in sympatry.

Acknowledgments

Thanks to J. Golcher and A. Haerer for assistance during field work. We also thank C. Chang-Rudolf and D. Monne-Parera for technical support and D. Hulsey for fruitful discussions. We thank MARENA (Ministerio de Ambiente y Recursos Naturales, Nicaragua) for collection permits (DGPN/DB-IC-011-2014). Lab experiments were run under permit number 35-9185.81/G-15/89. This research was funded by an Alexander von Humboldt fellowship, the Deutsche Forschungsgemeinschaft (DFG MA6144/1-1), and the Young Scholar Fund of the University of Konstanz (YSF, 83964814) to GMS. AFK was supported by the

Landesgraduiertenförderung of the State of Baden-Württemberg and the International Max Planck Research School for Organismal Biology. JTD was supported by an EU FP7 Marie Curie Zukunftskolleg Incoming Fellowship Program of the University of Konstanz (grant number 291784) and a grant from the Deutsche Forschungsgemeinschaft (TO914/2-1). FH was funded by a CNPq/DAAD (GDE-290049/2007-5) fellowship. AM was funded by an Advanced grant by the European Research Council (ERC “GenAdap” 293700) and various grants of DFG Deutsche Forschungsgemeinschaft (DFG).

Data Accessibility

All the data used in this study have been deposited in Dryad
doi:10.5061/dryad.kb1pk

References

- Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* **19**, 1655-1664.
- Arnegard ME, Snoeks J (2001) New three-spotted cichlid species with hypertrophied lips (Teleostei : Cichlidae) from the deep waters of Lake Malawi/Nyasa, Africa. *Copeia*, 705-717.
- Barlow GW (1976) The Midas cichlid in Nicaragua. In: *Investigations of the Ichthyology of Nicaraguan Lakes* (ed. Thorson TB), pp. 359-369. University of Nebraska Press, Lincoln.
- Barlow GW, Munsey JW (1976) The red devil-Midas-arrow cichlid species. In: *Investigations of the ichthyology of Nicaraguan lakes* (ed. Thorson TB), pp. 359-369. University of Nebraska Press, Lincoln.
- 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.
- Barluenga M, Stolting KN, Salzburger W, Muschick M, Meyer A (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* **439**, 719-723.
- 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*.
- Bolnick DI, Fitzpatrick BM (2007) Sympatric speciation: Models and empirical evidence. *Annual Review of Ecology Evolution and Systematics* **38**, 459-487.

- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology* **22**, 3124-3140.
- Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping Loci de novo from short-read sequences. *G3 (Bethesda, Md.)* **1**, 171-182.
- Colombo M, Diepeveen ET, Muschick M, *et al.* (2013) The ecological and genetic basis of convergent thick-lipped phenotypes in cichlid fishes. *Molecular Ecology* **22**, 670–684.
- Coyne JA, Orr HA (2004) *Speciation* Sinauer, Sunderland, Massachusetts.
- Cruikshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology* **23**, 3133-3157.
- Ellegren H, Smeds L, Burri R, *et al.* (2012) The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* **491**, 756-760.
- Elmer KR, Fan SH, Kusche H, *et al.* (2014) Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nature Communications* **5**, 8.
- Elmer KR, Kusche H, Lehtonen TK, Meyer A (2010a) 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 B-Biological Sciences* **365**, 1763-1782.
- Elmer KR, Lehtonen TK, Kautt AF, Harrod C, Meyer A (2010b) Rapid sympatric ecological differentiation of crater lake cichlid fishes within historic times. *Bmc Biology* **8**, 60.
- Elmer KR, Lehtonen TK, Meyer A (2009) Color assortative mating contributes to sympatric divergence of neotropical cichlid fish. *Evolution* **63**, 2750-2757.
- Excoffier L, Dupanloup I, Huerta-Sanchez E, Sousa VC, Foll M (2013) Robust Demographic Inference from Genomic and SNP Data. *Plos Genetics* **9**, 17.
- Excoffier L, Lischer HEL (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.
- Feder JL, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends in Genetics* **28**, 342-350.
- Feder JL, Flaxman SM, Egan SP, Comeault AA, Nosil P (2013) Geographic Mode of Speciation and Genomic Divergence. In: *Annual Review of Ecology, Evolution, and Systematics, Vol 44* (ed. Futuyma DJ), pp. 73-97. Annual Reviews, Palo Alto.
- Felsenstein J (1981) Skepticism towards santa rosalia, or why are there so few kinds of animals. *Evolution* **35**, 124-138.
- Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) version 3.6a2. Department of Genetics, University of Washington, Seattle.
- Fitzpatrick BM, Fordyce JA, Gavrillets S (2008) What, if anything, is sympatric speciation? *Journal of Evolutionary Biology* **21**, 1452-1459.
- Gagnaire PA, Pavey SA, Normandeau E, Bernatchez L (2013) The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. *Evolution* **67**, 2483-2497.
- Garrigan D, Kingan SB, Geneva AJ, *et al.* (2012) Genome sequencing reveals complex speciation in the *Drosophila simulans* clade. *Genome Research* **22**, 1499-1511.
- Gavrillets S (2004) *Fitness Landscapes and the Origin of Species* Princeton University Press, Princeton.
- Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *Plos Genetics* **5**, 11.

- Hatfield T, Schluter D (1999) Ecological speciation in sticklebacks: Environment-dependent hybrid fitness. *Evolution* **53**, 866-873.
- Hendry AP, Taylor EB, McPhail JD (2002) Adaptive divergence and the balance between selection and gene flow: Lake and stream stickleback in the misty system. *Evolution* **56**, 1199-1216.
- Henning F, Jones JC, Franchini P, Meyer A (2013) Transcriptomics of morphological color change in polychromatic Midas cichlids. *Bmc Genomics* **14**, 13.
- Henning F, Machado-Schiaffino G, Baumgarten L, Meyer A (2017) Genetic dissection of adaptive form and function in rapidly-speciating cichlid fishes. *Evolution in press*.
- Henning F, Meyer A (2014) The evolutionary genomics of cichlid fishes: explosive speciation and adaptation in the postgenomic era. *Annual Reviews of Genomics and Human Genetics* **15**, 417-441.
- Hey J, Chung Y, Sethuraman A (2015) On the occurrence of false positives in tests of migration under an isolation-with-migration model. *Molecular Ecology* **24**, 5078-5083.
- Holzman R, Collar DC, Mehta RS, Wainwright PC (2012) An integrative modeling approach to elucidate suction-feeding performance. *Journal of Experimental Biology* **215**, 1-13.
- Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**, 254-267.
- Jones FC, Grabherr MG, Chan YF, *et al.* (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55-61.
- Kautt AF, Elmer KR, 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.
- Kautt AF, Machado-Schiaffino G, Meyer A (2016a) Multispecies Outcomes of Sympatric Speciation after Admixture with the Source Population in Two Radiations of Nicaraguan Crater Lake Cichlids. *Plos Genetics* **12**, e1006157.
- Kautt AF, Machado-Schiaffino G, Torres-Dowdall J, Meyer A (2016b) Incipient sympatric speciation in Midas cichlid fish from the youngest and one of the smallest crater lakes in Nicaragua due to differential use of the benthic and limnetic habitats? *Ecology and Evolution* **6**, 5342-5357.
- Kohda M, Tanida K (1996) Overlapping territory of the benthophagous cichlid fish, *Lobochilotes labiatus*, in Lake Tanganyika. *Environmental Biology of Fishes* **45**, 13-20.
- Konings A (1998) *Tanganyika Cichlids in their natural habitats* Cichlid Press, St Leon-Rot.
- Kutterolf S, Freundt A, Perez W, Wehrmann H, Schmincke HU (2007) Late Pleistocene to Holocene temporal succession and magnitudes of highly-explosive volcanic eruptions in west-central Nicaragua. *Journal of Volcanology and Geothermal Research* **163**, 55-82.
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760.
- Losos JB (2011) Convergence, adaptation and constraint. *Evolution* **65**, 1827-1840.
- 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*, 2086-2091.
- Malinsky M, Challis RJ, Tyers AM, *et al.* (2015) Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science* **350**, 1493-1498.
- Manousaki T, Hull PM, Kusche H, *et al.* (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.

- Martin CH (2012) Weak Disruptive Selection and Incomplete Phenotypic Divergence in Two Classic Examples of Sympatric Speciation: Cameroon Crater Lake Cichlids. *American Naturalist* **180**, E90-E109.
- Martin CH (2013) Strong assortative mating by diet, color, size, and morphology but limited progress toward sympatric speciation in a classic example: Cameroon crater lake cichlids. *Evolution* **67**, 2114-2123.
- Martin CH, Cutler JS, Friel JP, *et al.* (2015) Complex histories of repeated gene flow in Cameroon crater lake cichlids cast doubt on one of the clearest examples of sympatric speciation. *Evolution* **69**, 1406-1422.
- Mayr E (1942) *Systematics and the origin of species from the viewpoint of a zoologist* Harvard University Press, Cambridge, MA.
- Meyer A, Kautt AF (2014) Sympatric speciation. In: *Oxford Bibliographies in Evolutionary Biology* (ed. Losos JB). Oxford University Press, New York.
- Neafsey DE, Lawniczak MKN, Park DJ, *et al.* (2010) SNP Genotyping Defines Complex Gene-Flow Boundaries Among African Malaria Vector Mosquitoes. *Science* **330**, 514-517.
- Noor MAF, Bennett SM (2009) Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity* **103**, 439-444.
- Nosil P (2012) *Ecological speciation* Oxford University Press, Oxford, UK.
- Oliver MK, Arnegard ME (2010) A new genus for *Melanochromis labrosus*, a problematic Lake Malawi cichlid with hypertrophied lips (Teleostei: Cichlidae). *Ichthyological Exploration of Freshwaters* **21**, 209-232.
- Papadopoulos AST, Baker WJ, Crayn D, *et al.* (2011) Speciation with gene flow on Lord Howe Island. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 13188-13193.
- Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *Plos Genetics* **2**, 2074-2093.
- Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *Plos One* **7**.
- R Development Core Team (2014) R: a Language and Environment for Statistical Computing.
- Ready JS, Sampao I, Schneider H, *et al.* (2006) Colour forms of Amazonian cichlid fish represent reproductively isolated species. *Journal of Evolutionary Biology* **19**, 1139-1148.
- Recknagel H, Elmer KR, Meyer A (2013) A hybrid genetic linkage map of two ecologically and morphologically divergent Midas cichlid fishes (*Amphilophus* spp.) obtained by massively parallel DNA sequencing (ddRADSeq). *G3-GENES GENOMES GENETICS*.
- Roesti M, Hendry AP, Salzburger W, Berner D (2012) Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. *Molecular Ecology* **21**, 2852-2862.
- Rueffler C, Van Dooren TJM, Leimar O, Abrams PA (2006) Disruptive selection and then what? *Trends in Ecology & Evolution* **21**, 238-245.
- Rundle HD, Whitlock MC (2001) A genetic interpretation of ecologically dependent isolation. *Evolution* **55**, 198-201.
- Schlieffen UK, Tautz D, Paabo S (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* **368**, 629-632.
- Schluter D (1995) Adaptive radiation in sticklebacks - Trade-offs in feeding performance and growth. *Ecology* **76**, 82-90.

- Schluter D (2000) *The Ecology of Adaptive Radiation* Oxford University Press, Oxford.
- Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675.
- Seehausen O, Butlin RK, Keller I, *et al.* (2014) Genomics and the origin of species. *Nature Reviews Genetics* **15**, 176-192.
- Servedio MR, Van Doorn GS, Kopp M, Frame AM, Nosil P (2011) Magic traits in speciation: 'magic' but not rare? *Trends in Ecology & Evolution* **26**, 389-397.
- Skorczewski T, Cheer A, Wainwright PC (2012) The benefits of planar circular mouths on suction feeding performance. *Journal of the Royal Society Interface* **9**, 1767-1773.
- Soria-Carrasco V, Gompert Z, Comeault AA, *et al.* (2014) Stick Insect Genomes Reveal Natural Selection's Role in Parallel Speciation. *Science* **344**, 738-742.
- Thibert-Plante X, Gavrillets S (2013) Evolution of mate choice and the so-called magic traits in ecological speciation. *Ecology Letters* **16**, 1004-1013.
- Turner TL, Hahn MW (2010) Genomic islands of speciation or genomic islands and speciation? *Molecular Ecology* **19**, 848-850.
- Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 9939-9946.
- Yeaman S, Whitlock MC (2011) The genetic architecture of adaptation under migration-selection balance. *Evolution* **65**, 1897-1911.

Author contributions

G.M.S and A.M designed the experiments. G.M.S analyzed the morphological data. G.M.S and J.T.D performed and analyzed the experiments in the field. G.M.S, A.F.K and J.T.D performed and analyzed assortative mating experiments in the lab. G.M.S, L.B and F.H performed and analyzed the feeding performance experiments in the lab. Population genomics data were generated and analyzed by A.F.K and G.M.S. Demographic analyses were performed by A.F.K. G.M.S and F.H designed the crosses and obtained the F2 populations. G.M.S wrote the manuscript with contributions from all the authors.

Figure Legends

Fig. 1 Thin- and thick-lipped populations in the Midas cichlid species complex. (A) Map showing the four lakes where thin- and thick-lipped Midas cichlids coexist. Admixture plot showing clear genetic clustering among the four lakes together with cross-validation errors for admixture runs are shown below the map. The lowest cross-validation error (i.e. highest support) was found for four genetic clusters. (B, C, D, E) Histograms and Kernel density plots of normalized lip area (lip area corrected by body area) in wild-caught individuals from the great lakes Managua (purple) and Nicaragua (green) and the crater lakes Apoyeque (red) and Masaya (blue). Sample sizes (N), Hartigans' Dip tests for unimodality (D) and p-values are shown for each lake.

Fig. 2 Functional trade-off between ecotypes. (A) Correlation between normalized lip area and feeding performance (distance from the base of the acrylic device to the furthest removed mosquito larvae) for wild caught thick- and thin-lipped fish from crater lake Apoyeque exposed to mosquito larvae attached to an acrylic structure with continuously decreasing angle. (B) Correlation between normalized lip area and feeding success ratio (number of eaten prey items divided by the number of attempts) for wild caught thick- and thin-lipped fish from crater lake Apoyeque exposed to free swimming fry. (C) Differential performance (mean change in weight per day in grams and standard deviation) between thick- and thin-lipped ecotypes in rocky and sandy habitats areas in crater lake Apoyeque.

Fig. 3 Assortative mating by ecotypes. (A) Assortative mating in crater lake Apoyeque. Census of pairs during breeding season 2013-2014. Observed number of pairs (red/orange) and expected under random mating (grey) are shown. (B) Strong Assortative mating under laboratory conditions between *A. citrinellus* (thin-lipped) and *A. labiatus* (thick-lipped) from great lake Nicaragua. Observed number of pairs (green or light green) and expected under random mating (grey) are shown. (C) Correlation between normalized lip area of mating females and males under laboratory conditions.

Fig. 4 Genetic relationship and differentiation between ecotypes. (A) Neighbor-joining tree showing that fish from crater lakes Apoyeque (red/orange) and Masaya (blue/light blue) share a more recent common ancestor with those from great lake Managua than great lake Nicaragua (purple/light purple). Thick- and thin-lipped ecotypes from great lake Nicaragua are shown in green and light green, respectively. (B, C) First three main axes of genetic variation (Principal Component Analysis) between ecotypes in the great lakes Managua (purple/light purple) and Nicaragua (green/light green). The first principal component of the PCA was highly significant (p-value ~ 0) for both great lakes. Analyses based on 16 286 and 21 465 polymorphic loci for great lake Managua and Nicaragua, respectively. (D, E) First three main axes of genetic variation (principal components) between ecotypes in the crater lakes Apoyeque (red/orange) and Masaya (blue/light blue). The first principal component of the PCA was highly significant for crater lake Masaya but only slightly significant for L. Apoyeque ($p = 9.34 \times 10^{-29}$ and $p = 0.011$, respectively). Analyses based on 13 547 and 6693 polymorphic loci for crater Masaya and Apoyeque, respectively. (F) Schematic illustrations of the most supported demographic models

of crater lakes Apoyeque and Masaya assuming great lake Managua as the source population.







