1 # Incipient speciation driven by hypertrophied lips in

2 # Midas cichlids fish?

3

4 GONZALO MACHADO-SCHIAFFINO*, ANDREAS F. KAUTT*,[1], JULIAN TORRES-

5 DOWDALL*,[1], LUKAS BAUMGARTEN*,[†], FREDERICO HENNING* and AXEL

6 MEYER*

7 *Chair of Zoology and Evolutionary Biology, Department of Biology, University of

8 Konstanz, Universitätsstrasse 10, 78457, Konstanz, Germany

9

10 Correspondence: Gonzalo Machado-Schiaffino or Axel Meyer, Fax: +49

11 7531883018; E-mail: gonzalo.machado-schiaffino@uni-konstanz.de or

12 axel.meyer@uni-konstanz.de

13 [1]These authors contributed equally to this work

14 Current address:

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

16 Köln, Germany16

17

18

19 ## Abstract

20 **Sympatric speciation has been debated in evolutionary biology for decades.**

21 **Although it has gained in acceptance recently, still only a handful of empirical**

22 **examples are seen as valid (e.g. crater lake cichlids). In this study, we**

23 **disentangle the role of hypertrophied lips in the repeated adaptive radiations of**

24 **Nicaraguan crater lake cichlid fish. We assessed the role of disruptive**

25 **selection and assortative mating during the early stages of divergence and**

26   **found a functional trade-off in feeding behavior between thick- and thin-lipped**

27   **ecotypes suggesting that this trait is a target of disruptive selection. Thick-**

28   **lipped fish perform better on non-evasive prey at the cost of a poorer**

29   **performance on evasive prey. Using enclosures in the wild, we found that**

30   **thick-lipped fish perform significantly better in rocky than in sandy habitats.**

31   **We found almost no mixed pairs during two breeding seasons and hence**

32   **significant assortative mating. Genetic differentiation between ecotypes seems**

33   **to be related to the time since colonization, being subtle in L. Masaya (1600**

34   **generations ago) and absent in the younger L. Apoyeque (<600 generations**

35   **ago). Genome-wide differentiation between ecotypes was higher in the old**

36   **source lakes than in the young crater lakes. Our results suggest that**

37   **hypertrophied lips might be promoting incipient sympatric speciation through**

38   **divergent selection (ecological divergence in feeding performance) and non-**

39   **random mating (assortative mating) in the young Nicaraguan crater lakes.**

40   **Nonetheless, further manipulative experiments are needed in order to confirm**

41   **the role of hypertrophied lips as the main cue for assortative mating.**

42

45

Introduction

47   The likelihood of speciation in the presence of gene flow has been debated for

48   decades (Mayr 1942; Coyne & Orr 2004; Gavrilets 2004; Bolnick & Fitzpatrick 2007;

49   Fitzpatrick *et al.* 2008; Nosil 2012; Feder *et al.* 2013; Meyer & Kautt 2014). For

50   sympatric speciation by natural selection to happen, typically three main conditions

51   have to be fullfilled: (1) strong disruptive selection on ecotypes, (2) strong assortative

52   mating by ecotypes and (3) the build-up of linkage disequilibrium between the loci

53   responsible for the ecological traits and mate choice (Bolnick & Fitzpatrick 2007).

54   Although still no agreement has been reached about how often speciation with gene

55   flow occurs, it is generally accepted that divergence in sympatry is facilitated when

56   the same trait that is under divergent natural selection is also involved in assortative

57   mating (Gavrilets 2004; Papadopulos *et al.* 2011; Servedio *et al.* 2011; Martin 2013).

58   To better understand what roles divergent selection and assortative mating play at

59   the earliest stages of speciation one needs to identify and investigate case studies

60   where both of these processes might be at work. Crater lake cichlid fish are widely

61   regarded as one of the most convincing examples of sympatric speciation (Schliewen

62   *et al.* 1994; Coyne & Orr 2004; Barluenga *et al.* 2006; Malinsky *et al.* 2015; Kautt *et*

63   *al.* 2016a) but see (Martin *et al.* 2015) and are an ideal system to assess the potential

64   role of adaptive traits during the early stages of divergence in sympatry.

65   Midas cichlids (*Amphilophus* sp. species complex) inhabiting a number of

66   small, completely isolated volcanic crater lakes located in Western Nicaragua were

67   colonized recently (less than 24 000 years ago) from the great lakes Managua and

68   Nicaragua (Fig. 1A). This group of Neotropical cichlids is particularly variable in eco-

69   morphological traits such as body shape (limnetic/benthic), pharyngeal jaws

70   (papiliform/molariform), and remarkably hypertrophied lips (thick-/thin-lipped)

71   (Barluenga *et al.* 2006; Elmer *et al.* 2010a; Manousaki *et al.* 2013; Machado-

72   Schiaffino *et al.* 2014).

73   Interestingly, hypertrophied lips have evolved independently and repeatedly in

74   African and Neotropical cichlids (Henning & Meyer 2014). The fact that thick-lipped

75   ecotypes evolved in parallel strongly supports the hypothesis that this trait is an

76   adaptation (*sensu* Losos 2011) for a particularly specialized type of foraging

77   behavior. It has been suggested that the feeding apparatus (hypertrophied lips in

78   combination with narrow and pointed heads) of thick-lipped species enhances the

79   ability to forage in rocky substrates by facilitating the access to crustaceans and fish

80   larvae that hide between rocks (Barlow & Munsey 1976; Kohda & Tanida 1996;

81   Konings 1998; Arnegard & Snoeks 2001; Oliver & Arnegard 2010; Baumgarten *et al.*

82   2015). Although this had not been properly tested in Neotropical cichlids yet, it

83   appears to be the case for African cichlids (Baumgarten *et al.* 2015).

84   Within the Midas cichlid species complex, thick- and thin-lipped ecotypes are

85   present in both great lakes (formally described as *A. labiatus* and *A. citrinellus*,

86   respectively) (Barlow 1976), and in the crater lakes Masaya and Apoyeque (Fig. 1A).

87   Frequency of thick-lipped ecotype is much lower (e.g. 5 to 20%) than thin-lipped one

88   in all these lakes, being the highest (around 20%) in crater lake Apoyeque (Elmer *et*

89   *al.* 2010b). Interestingly, L. Masaya is one of the oldest (6000 years old) and L.

90   Apoyeque one of youngest (1800 years old) crater lakes in Nicaragua (Kutterolf *et al.*

91   2007). Morphological differences between thick- and thin-lipped ecotypes are

92   principally found in the mouth region, (Elmer *et al.* 2010b; Manousaki *et al.* 2013) and

93   are associated with diet differences both in African (Colombo *et al.* 2013) and

94   Neotropical (Manousaki *et al.* 2013) cichlids. Based on stomach contents, both

95   studies found that thin-lipped ecotypes predate mostly on fish whereas thick-lipped

96   ecotypes predate preferentially on invertebrates; suggesting a potential trade-off.

97   Moreover, it has been suggested that these niche differences might drive incipient

98   ecological sympatric speciation at least in crater lake Apoyeque (Elmer *et al.* 2010b).

99   The young age of the Midas cichlid complex and the repeated colonization of

100   the crater lakes at different points in time presents a unique opportunity to assess the

101   role of ecologically relevant traits, such as hypertrophied lips, during different stages

102   of the speciation process. By comparing the potential functional trade-off, the degree

103   of assortative mating, the pattern of genome-wide genetic differentiation and the

104   demographic history between thick- and thin-lipped ecotypes we aimed to

105   disentangle how disruptive selection and assortative mating might promote or

106   constrain sympatric speciation in the young Nicaraguan Midas cichlid species

107   complex.

108

109   **Materials and methods**

110

111   *Hypertrophied lip size distribution in wild fish*

112   Standardized lateral photographs of wild caught individuals from each lake were

113   taken in order to quantify lip size. 137, 168, 234 and 89 individuals were collected

114   during field trips in L. Managua, L. Nicaragua, L. Apoyeque and L. Masaya,

115   respectively. Protruded lip area (as measured in Machado-Schiaffino *et al.* (2014))

116   and body area were measured using ImageJ software (Schneider *et al.* 2012) (see

117   Fig. S1F). Hereafter, we refer to normalized lip area as the protruded lip area

118   corrected by body area and all values are expressed in ratios. Hartigans' Dip tests for

119   unimodality and Kernel density plots were carried out in R (R Development Core

120   Team 2014).

121

122   *Testing feeding performance*

123   Two independent behavior experiments were conducted in order to compare the

124   feeding performance of wild caught thick- and thin-lipped Midas cichlids from Lake

125   Apoyeque when fish are exposed to sessile or evasive prey. For both experiments

126   fish were starved for one day prior to the experiment to keep them at a similar level of

127   satiation. Fish were transferred into a 200L experimental tank one day before the

128   beginning of each experiment for acclimation. Standardized photographs were taken

129   before the experiments in order to measure both lip area, head angle, lip length,

130   standard length and body area for each individual. All measurements were size-

131   corrected by standard length or body area.

132   As a first experiment, 15 wild caught thick- and 16 thin-lipped juvenile fish

133   (mean standard length = 86.6 mm, SD = 7.1) from Lake Apoyeque were individually

134   exposed to mosquito larvae equidistantly placed (every 3 mm) along an acrylic

135   structure (see Fig. 2A) consisting of a continuously decreasing angle (angle range =

136   60º to 15º; total height = 16 cm) following Henning et al. (2017). Each acrylic device

137   was kept in the tank for at least six hours. After that time, the distance from the base

138   (angle = 60º) of the acrylic device to the furthermost removed mosquito larvae was

139   recorded as a proxy for the minimum foraging angle for each fish.

140   For the second experiment, the same 15 wild caught thick- and 15 thin-lipped

141   juvenile fish from Lake Apoyeque were individually exposed to five free-swimming

142   fish fry. Each experimental fish was recorded from above with a video camera for at

143   least 45 minutes. Only trials in which fish were able to catch at least three fry were

144   considered to obtain a better estimates of the average number of attempts. But note

145   that results were qualitatively similar if the minimum number was set to two. After the

146   experiments the number of failures and successful attacks were quantified from the

147   video recordings.

148

149   *Testing feeding performance in situ: Enclosures*

150   Reciprocal transplant enclosure experiments in the wild have been successfully

151   applied in fish to test fitness of morphologically divergent species in different

152   ecological niches (Hatfield & Schluter 1999; Hendry *et al.* 2002). However, to our

153   knowledge, enclosure experiments have not been extensively used in cichlid fish to

154   test fitness trade-off associated with ecologically relevant traits. In the present study,

155   wild-caught thick- and thin-lipped Midas cichlids from Lake Apoyeque were allocated

156   to enclosures in two well differentiated habitats (rocky and sandy) within the lake in

157   order to test whether the performance between ecotypes (measured as difference in

158   growth) varies depending on the habitat. A total of 40 enclosures were built using

159   chicken-wire mesh. Each enclosure had 1.125 m^3 of volume (0.9 x 2.5 x 0.5 meters).

160   All enclosures were attached to the ground (approximately at two meters depth) and

161   equally distributed in the rocky and sandy area. The material employed to build the

162   enclosures was flexible and allowed to fix them in close contact with the ground even

163   in the rocky habitat characterized by crevices among rocks. Single fish were

164   introduced into each enclosure in order to avoid the effect of competition. A total of

165   ten enclosures per morph (thick- and thin-lipped) and habitat (rocky and sandy) were

166   included in this experiment. Juvenile individuals were chosen in order to ensure that

167   substantial growth was still possible. Size matching (by weight) was performed to

168   ensure that starting weight of the fish in the enclosures was similar for each ecotype

169   and habitat (ANOVA *F-statistic* = 0.1528, df = 3, p = 0.698; starting weights: overall

170   mean = 25 g, SD = 7.64).

171   All fish were weighted before the beginning of the experiment and after seven

172   weeks. Standardized pictures were taken for all individuals. Lip area, body area, and

173   standard length were measured. Differences in weight between morph and habitat

174   were tested with non-parametric Kruskal-Wallis tests due to a lack of normality and

175   homoscedasticity of the data even after transformation. Multiple comparisons were

176   performed with Dunn's test implemented in R (R Development Core Team 2014) and

177   p values were adjusted using Bonferroni corrections.

178

179   *Assortative mating*

7

180   Pair-composition censuses were performed in Lake Apoyeque during two breeding

181   seasons, in December 2013 and 2014. Two divers performed the census

182   independently, starting from the same point of the lake and following opposite

183   directions in order to avoid counting the same nest twice. In total, around 1.5 km (one

184   quarter of the total perimeter of the lake) were covered. All nests with offspring and

185   both parents were included in the analyses. After the identification of a nest, a close

186   examination was carried out in order to identify the phenotypes of both parents.

187   Despite the fact that it was not possible to quantify lip area for the parents, their

188   ecotypes were easily determined without ambiguity.

189   Additionally, the degree of assortative mating was determined for fish from

190   Lake Nicaragua. Due to the size of the lake and its turbidity, pair-composition

191   censuses in the wild were not possible for this lake. Thus, mate-choice experiments

192   were performed under laboratory conditions. All fish included in this experiment were

193   F1s from multiple crosses of wild-caught individuals. A total of 25 thick- (*A. labiatus*)

194   and 25 thin-lipped (*A. citrinellus*) adult individuals were kept under similar conditions

195   (temperature, light and substrate) in a very large tank (7000 L). In order to reduce the

196   number of potential variables associated with pair formation, only dark morphs from

197   both species were used. After seven days of acclimation fish were allowed to interact,

198   the tank was periodically inspected, and the composition of every formed pair that

199   stayed together for at least three days was recorded. Note that spawning took place

200   within the first three days after pair-formation in most cases. After that period, the pair

201   was removed and replaced with individuals of the same ecotype, size, and sex. Thus,

202   the frequency of ecotypes in the tanks was kept constant during the whole

203   experiment. Standardized photographs from the lateral and dorsal side were taken

204   for all individuals that successfully formed a pair in order to measure the association

205   between male and female lip size.

206

207 *Genome-wide differentiation*

208 To quantify how genetic variation is partitioned across genomes between ecotypes

209 and lakes, we performed restriction-site associated DNA-sequencing (RAD-seq)

210 based population genomic analyses. Between 19 and 72 wild caught individuals per

211 ecotype (thick- and thin-lipped) and lake (Managua, Nicaragua, Masaya and

212 Apoyeque) were included in the genomic libraries (Table S1, Supporting information).

213 Most of these fish were collected between 2010 and 2013 and are stored in AM's

214 collection at the University of Konstanz. A small piece of pectoral fin from each

215 sample was used to extract genomic DNA with a Zymo kit.

216   Double digest RADseq libraries were prepared following Peterson *et al.* (2012)

217 and Kautt *et al.* (2016a). Briefly, 600 ng of genomic DNA per sample was double-

218 digested using the restriction enzymes PstI and MspI for 3 hours at 37 °C. After that,

219 P1 and P2 adapters (Recknagel *et al.* 2013) were ligated to the digested DNA using

220 T4 ligase for 30 minutes at room temperature. A total of 307 individually barcoded

221 samples were pooled in six libraries. Size selection for each library was performed

222 using Pippin Prep technology (Sage Science, Beverly, MA) setting the size-range

223 from 320 to 500bp. Finally, genomic libraries were single-end sequenced (100 bp

224 length) in four lanes on an Illumina HiSeq 2000.

225   Raw sequence reads were demultiplexed using STACKS v.1.2.9 (Catchen *et*

226 *al.* 2011; Catchen *et al.* 2013). Only high sequencing quality reads, with correct

227 barcodes and unambiguous RAD site were retained. Demultiplexed reads were

228 mapped to an "in house" anchored *A. citrinellus* reference genome (Elmer *et al.*

229 2014) using BWA (Li & Durbin 2009). SNPs and genotype calling were performed

230 with STACKS. Loci deviating from HWE or showing more than three polymorphic

231 sites (SNPs) in one population were excluded. Further, loci exhibiting a SNP in the

232   last two positions of a read were excluded, because of an excess number of SNPs at

233   these two positions. A locus was included in the analyses only if it was genotyped in

234   at least 10 individuals (20 alleles) per population. On average 74 197 ± 18 049 (sd)

235   loci were obtained per individual with a mean coverage of 13.94 ± 4.21 (sd) reads per

236   locus and individual (Table S1, Supporting information).

237   Population genomic statistics were obtained using the *populations* program

238   implemented in STACKS. Genome-wide Fst values were calculated among all pair-

239   wise combinations of populations using ARLEQUIN software with default settings

240   (Excoffier & Lischer 2010), and the significance was tested using 10 000

241   permutations. Population structure was investigated with the model-based software

242   ADMIXTURE (Alexander *et al.* 2009) and model-free Principal Component Analyses

243   (PCAs) implemented in the EIGENSOFT package (Patterson *et al.* 2006). Support for

244   the number of genetic clusters was determined based on the implemented cross-

245   validation error technique or Tracy-Widom statistics, respectively. The phylogenetic

246   tree was built using PHYLIP (Felsenstein 1993) based on a distance matrix

247   calculated from the frequency data of 1000 randomly selected markers. Statistical

248   support of nodes was estimated with 1000 bootstrap replicates. Note that only one

249   SNP per locus was used for analyses concerning population structure or the

250   phylogenetic tree to account for non-independence of linked sites. Additionally,

251   individual-based phylogenetic networks were built using SPLITSTREE v.4.13.1

252   (Huson & Bryant 2006). Individual genotype calls were transformed from variant call

253   format (VCF) to Nexus format using custom scripts and networks were built using the

254   NeighborNet method based on uncorrected P distances.

255

256   *Demographic inference*

257   Demographic parameters were estimated by simulation and fitting of different

258   demographic models to the information contained in the multidimensional minor site

259   frequency spectrum (SFS) using FASTSIMCOAL2 (Excoffier *et al.* 2013). Briefly,

260   FASTSIMCOAL2 uses coalescent simulations to obtain the expected SFS for a

261   certain specified model and then optimizes each demographic parameter in the

262   model so as to maximize the likelihood. Demographic models can be arbitrarily

263   complex, including a variety of demographic events such as migration, population

264   size changes, population splits and admixture, and there is no restriction in the

265   number of analyzed populations. The relative fit of different demographic models to

266   the data was evaluated by means of the Akaike Information Criterion (AIC).

267   Loci presumably located in coding regions were identified via a BLAST search

268   against a compilation of transcriptomic data from various species and tissues of

269   Midas cichlids (Henning *et al.* 2013; Manousaki *et al.* 2013) and excluded. Only one

270   SNP per locus was used for the demographic analyses. The SFS was created as in

271   Kautt *et al.* (2016a). Briefly, data was parsed from VCF files using a custom python

272   script and transformed into the SFS using δαδι (Gutenkunst *et al.* 2009). To account

273   for missing data the SFS was projected down to 25 individuals per lake. The same

274   minimum number of 25 individuals per lake was specified to create the VCF files.

275   Thus the ratio between monomorphic and polymorphic sites can be correctly inferred

276   and used to convert the parameter estimates into demographic units. Once the

277   presumably best model is chosen the fit to the data can be inferred by using the

278   maximum likelihood point estimates of the model parameters to perform parametric

279   bootstrapping and obtain confidence intervals (Excoffier *et al.* 2013).

280

281   **Results**

11

282

283   *Bimodal distribution of hypertrophied lips*

284   The normalized size of lips (lip area corrected by body area) is strongly bimodal in all

285   lake populations except for crater lake Masaya where unimodality (Hartigans' Dip

286   test) could not be rejected probably due to low sample size (Fig. 1B-E). Phenotypic

287   differentiation was more pronounced in the older source lakes (Fig. 1B, C) than in the

288   recently colonized crater lakes (Fig. 1D, E). This pattern is concordant with

289   theoretical predictions of eco-morphological traits under disruptive selection during

290   the early stages of divergence (Coyne & Orr 2004; Gavrilets 2004).

291

292   *Disruptive selection acting on lips*

293   Coupled with differences in lips we found pronounced differences in feeding

294   performance – this was tested with wild-caught thick- and thin-lipped ecotypes from

295   Lake Apoyeque (Fig. 2). Thick-lipped fish performed better than thin-lipped fish when

296   food was attached (sessile) and had to be scraped off an acrylic structure with a

297   continuously decreasing angle (Fig. 2A), mimicking crevices in the natural habitat.

298   Although thick- and thin-lipped fish are different both in terms of lips and head shape,

299   feeding performance when food was sessile was significantly associated only with

300   hypertrophied lips (Fig. S1, Supporting information). These findings were validated

301   also in experiments with thick- (*A. labiatus*) and thin-lipped (*A. citrinellus*) fish from

302   great lake Nicaragua (Fig. S2A, Supporting information).

303   However, wild-caught thin-lipped fish performed significantly better than thick-

304   lipped fish when feeding on evasive prey in the water column (Figs 2B and Fig. S2B,

305   Supporting information), indicating a trade-off. Here, feeding success was negatively

306   associated with lip size*.*

307   This trade-off was also assessed in terms of fitness consequences with

308   experiments in the crater lakes: using single-fish enclosures (n = 40) in sandy or

309   rocky habitats in crater lake Apoyeque we monitored the change in weight of the fish

310   as a proxy for fitness. Despite the fact that all fish had a similar weight at the

311   beginning of the experiment (ANOVA *F-statistic* = 0.1528, df = 3, p = 0.698; starting

312   weights: overall mean = 25 g, SD = 7.64), the weight change after seven weeks

313   differed significantly (Kruskal-Wallis chi-squared = 10.4717, df = 3, p-value = 0.0149)

314   in the predicted direction between ecotypes and habitats (Fig. 2C). Although most

315   fish lost weight in the enclosures - probably due to the restricted foraging space -,

316   thick-lipped fish in rocky habitat lost significantly less weight than thick-lipped fish

317   confined to a sandy habitat (Dunn's statistics = -2.540, Bonferroni adjusted p =

318   0.033). Also a trend (Dunn's statistics = -2.199, Bonferroni adjusted p = 0.083)

319   towards thick-lipped fish compared with thin-lipped fish was apparent on rocky

320   habitat. Differences between thick- and thin-lipped fish in the enclosures on sandy

321   habitat were not significant (Dunn's statistics = -0.696, Bonferroni adjusted p =

322   1.000), which might be best explained by the observation that evasive prey (e.g.

323   small fish) was free to leave the enclosures, thereby depriving thin-lipped fish from

324   their foraging advantage compared to thick-lipped fish. This is in agreement with the

325   fact that thin-lipped fish predate preferentially on fish compared to thick-lipped fish

326   (Colombo *et al.* 2013; Manousaki *et al.* 2013).

327

328   *Assortative mating between ecotypes*

329   Given the bimodality and fitness trade-offs of this trait we conducted both field

330   observations and laboratory experiments to determine if ecotypes mate assortatively.

331   Due to the clarity of its water we chose to survey Lake Apoyeque for breeding pairs

332   during two consecutive breeding seasons (2013-2014); the turbid waters of the great

333   lakes and crater lake Masaya precluded us from collecting data on pairs in these

334   lakes. We note that Midas cichlids are seasonally monogamous and pairs defend

335   their brood jointly for weeks. Among a total of 68 breeding pairs, mating was highly

336   significantly non-random (Yates' chi-square = 25.147, p = 3.46×10^{-6}) and based on

337   ecotypes (Fig. 3A), supporting the hypothesis that gene flow is restricted between the

338   ecotypes. Due to the fact that thick-lipped fish occur much less frequently (c.a. 5-

339   20%) than thin-lipped fish in all the lakes where both ecotypes coexist, it would be

340   expected under random mating that most of the pairs including thick-lipped fish would

341   be disassortative; thus the probability of pairing up of two thick-lipped fish is expected

342   to be very low (e.g. 4%, assuming a 20% frequency of thick-lipped fish). However, we

343   found significantly more assortative thick-lipped pairs than expected. The total

344   number of thick-lipped fish included in this survey (11 out of 136, around 8%) was

345   similar to the frequency of this ecotype in the wild (around 20% for crater lake

346   Apoyeque (Elmer *et al.* 2010b)). Interestingly, many of the thin-lipped pairs (20 out of

347   a total of 30 pairs where the kind of substrate was recorded) were found in rocky

348   areas, although they supposedly have a feeding advantage in sandy areas. Thus,

349   both ecotypes tend to breed in rocky areas of the lake suggesting that there is no

350   habitat isolation during breeding season.

351   In order to validate the strong assortative mating between ecotypes found in

352   the wild, we also performed mate choice experiments in the laboratory in very large

353   tanks with *A. citrinellus* (thin-lipped) and *A. labiatus* (thick-lipped) from Lake

354   Nicaragua. All 25 pairs that formed in these experiments were assortative (Fig. 3B).

355   There was no correlation between lip size (corrected by body area) of the paired

356   individuals within *A. labiatus*, suggesting that the presence of lips alone, rather than

357   their size, might matter most for mate choice (Fig. 3C). Moreover, pair composition

358   was independent of size (weight) for *A. labiatus* and *A. citrinellus* pairs (r = -0.232 p =

359   0.446 and r = -0.207 p = 0.518, respectively).

360

361   *Recent colonization and weak genome-wide genetic differentiation between ecotypes*

362   *in the crater lakes*

363   Based on more than 16 740 polymorphic genetic markers, clear genome-wide

364   genetic differentiation was found with each lake population forming a distinct genetic

365   cluster (Figs 1A, and S3A, Supporting information). Using a much smaller genetic

366   data set of microsatellite markers, we had previously suggested that crater lake

367   Masaya was colonized from the nearby great lake Nicaragua (Barluenga & Meyer

368   2010). Yet, the analysis of this new substantial set of genome-wide markers shows

369   that both crater lakes share a more recent common ancestor with great lake

370   Managua (Fig. 4A). Moreover, individuals from crater lake Masaya tend to cluster

371   closer to great lake Managua (Fig. S3A, Supporting information). Individual-based

372   phylogenetic split networks also suggest that both crater lakes cluster together with

373   great lake Managua and not with Lake Nicaragua (Fig. S4, Supporting information).

374   Altogether, these findings support the notion that we have the interesting opportunity

375   to study the evolution of two independent radiations that were recently, but at

376   different times, founded from the same source population (Lake Managua).

377   Both crater lakes Apoyeque and Masaya are genetically clearly distinguishable

378   and distinct from Lake Managua (Table S2, Supporting information; Figs 1A, and S4,

379   Supporting information). The genetic differentiation between thick- (*A. labiatus*) and

380   thin-lipped (*A. citrinellus*) species in the source lakes Managua (F_{ST} = 0.050,

381   p<0.00001) and Nicaragua (F_{ST} = 0.019, p<0.00001) are subtle, but highly significant

382   (see Figs 4B, C, and S3B, C, Supporting information). These findings agree with the

383   strong assortative mating between *A. labiatus* and *A. citrinellus* from Lake Nicaragua

384   that we found in lab experiments (Fig. 3B) and support their status as two biological

385   species. Similarly, we found significant, yet even lower, genome-wide differentiation

386   between ecotypes in the about 6000 years old crater lake Masaya (F_{ST} = 0.016,

387   p<0.00001) (Fig. 4E). However, no genetic differentiation was found between

388   ecotypes in the less than 2000 year old crater lake Apoyeque (F_{ST} = 0.002, p = 0.827)

389   (Fig. 4D). This lack of genome-wide differentiation is not unexpected at this early

390   stage of speciation without geographical isolation.

391   The demographic history of the crater lake populations was inferred (Fig. 4F)

392   through coalescent simulations and comparing their fit against the empirical data

393   summarized in the two-dimensional site frequency spectra (SFS) (Excoffier *et al.*

394   2013). Assuming great lake Managua as the source population, 13 and 17 different

395   models were tested for crater lakes Apoyeque and Masaya, respectively (Table S3,

396   Supporting information).

397   According to the best model for crater lake Apoyeque the source population

398   (great lake Managua) experienced a population bottleneck about 1870 (95%

399   confidence interval: 1480-2520) generations ago and was reduced to only about

400   1490 (1040-2290) individuals and it has since been growing exponentially reaching a

401   population size of about 460,270 (0-898,670) individuals at present. The results for

402   the source lakes are similar to Kautt *et al.* (2016a) in which they were used as

403   outgroups for two other crater lake radiations. The confidence intervals around the

404   current population sizes are wide, probably because small differences in the

405   estimated growth rates will lead to a high variance since growth is exponential. Lake

406   Apoyeque was colonized by a small founder population of only about 110 (50-200)

407   individuals around 580 (430-770) generations ago, growing to a current size of 14

408   720 (1480-32 990) individuals. Continuous migration between the lakes is not

409   supported, but a single admixture event 380 (290-470) generations ago in which the

410 crater lake population received approximately 16% (8.3%-22.4%) of its gene pool

411 from the source population is strongly supported.

412   In the best model for crater lake Masaya the estimates for the population

413 dynamics in the source lake (Managua) are similar to above (Fig. 4F). Lake Masaya

414 was colonized only about 1560 (1400-1800) generations ago. Surprisingly, a

415 population size change in L. Masaya is not supported, and it seems thus to have

416 remained relatively stable at about 8610 (7800-9760) individuals throughout its

417 history. Continuous migration from the source into Masaya is not supported, yet Lake

418 Masaya received about 21.0% (14.5%-29.2%) of its gene pool from the source

419 population in an admixture event about 240 (120-400) generations ago. While we

420 consider the model above to be the best model, a different model in which L. Masaya

421 was colonized before the bottleneck in the source lake received a higher support.

422 Yet, considering the geographic isolation of this lake, this model assumes biologically

423 unrealistically high amounts of gene flow (continuous migration of 1.91 x 10^{-4} and

424 57.1% admixture event) and a colonization time (6390 generations ago) prior to the

425 formation of the crater lake itself.

426   In summary, L. Apoyeque was colonized much more recently than L. Masaya

427 and is characterized by a much smaller long-term effective population size.

428 Differences were also visually apparent in a plot of the two-dimensional site

429 frequency spectrum (2D-SFS). While the spectrum of the source population together

430 with L. Apoyeque is characterized by a considerable number of diverging sites, in the

431 case of L. Masaya many more sites fall along the diagonal, indicating only weak

432 signs of divergence (Fig. S6, Supporting information).

433

434 **Discussion**

435 We find that a conspicuous morphological trait, hypertrophied lips, is associated with

436 genetic differentiation in several repeated radiations of the Nicaraguan Midas cichlid

437 species complex. In several experiments in the field and the laboratory we document

438 a trade-off with respect to feeding behavior between thick- and thin-lipped ecotypes:

439 thick-lipped fish perform better on non-evasive prey attached to narrow angles

440 ("crevices"), but are significantly less successful in feeding on free swimming prey

441 than thin-lipped fish. These findings from laboratory experiments were supported by

442 enclosure studies in crater lake Apoyeque, in which thick-lipped fish performed better

443 in rocky than sandy habitats. That these two types of fish are ecologically distinct was

444 suggested before based on stomach content and stable isotope analyses (Elmer *et*

445 *al.* 2010b; Manousaki *et al.* 2013), but there was no information on feeding

446 performance. Thus, hypertrophied lips are likely subject to disruptive natural

447 selection. Moreover, our results suggest that this trait might also play a role in

448 assortative mating or is tightly linked to a trait resulting in assortative mating,

449 however, further manipulative experiments are needed to confirm this hypothesis.

450

451 *Disruptive selection acting on lips*

452 Here, we provide evidence that hypertrophied lips, and not other correlated traits, are

453 subject to divergent natural selection. A bimodal distribution of lip size was found in

454 all populations, except for crater lake Masaya where the sample size was small, with

455 more pronounced differences in the older lakes (great lakes Managua and

456 Nicaragua) than the much younger crater lakes (see Fig. 1B-E). A trade-off in feeding

457 performance was found between thick- and thin-lipped fish: the former are more

458 successful in feeding in small crevices (characteristic of rocky habitat) whereas the

459 latter feed more efficiently on free swimming fish (sandy/open-water habitat) (Fig. 2).

460 Hybrids between thick- and thin-lipped ecotypes exhibiting an intermediate

461 phenotype (Machado-Schiaffino *et al.* 2014) and are expected to perform worse than

462 either of the more extreme phenotypes on their preferred foods/habitats (Rueffler *et*

463 *al.* 2006). Therefore, hybrids would be expected to have a lower fitness (Rundle &

464 Whitlock 2001). Unfortunately, we could not perform these experiments on hybrids

465 with an intermediate phenotype due to logistical reasons. Nonetheless, it seems that

466 in both experiments the performance of thick-lipped fish with relatively small lip sizes

467 was intermediate (Fig. 2A-B).

468   One crucial advantage of hypertrophied lips is that they increase suction

469 power during foraging by sealing cracks and grooves in rocky substrate (Barlow &

470 Munsey 1976; Konings 1998; Oliver & Arnegard 2010; Baumgarten *et al.* 2015). We

471 show that the benefits of hypertrophied lips also engender costs when feeding in the

472 open-water (i.e. are subject to a trade-off; Fig. 2); explaining the pronounced diet

473 differences we previously found where thin-lipped fish tend to prey more on evasive

474 prey (e.g. fish) than thick-lipped individuals (e.g. hard-shelled invertebrates)

475 (Colombo *et al.* 2013; Manousaki *et al.* 2013). The kinetics of the buccal protrusion

476 during prey capture might differ between these ecotypes and would explain the

477 differences in predation success (see also (Holzman *et al.* 2012)). Hypertrophied lips

478 might also be constraining the size and shape of the mouth in thick-lipped fish and

479 create more turbulence in the water, thereby decreasing feeding success

480 (Skorczewski *et al.* 2012) on fish (Fig. 2B). Preliminary analyses furthermore suggest

481 that *A. labiatus* and *A. citrinellus* also differ strongly in mouth shape, due to a

482 difference in gape height, which is in turn caused by differences in the ascending and

483 descending process of the premaxillary bone (Fig. S5, Supporting information). It

484 would be expected that a more planar and circular mouth shape, as is found in *A.*

485 *citrinellus*, increases feeding performance by maximizing the capacity of suction

486 feeders to exert hydrodynamic forces on the prey (Skorczewski *et al.* 2012). Overall,

487   thick-lipped fish might increase suction power by sealing cracks and grooves when

488   their mouth is firmly attached to the rocky substrate, meanwhile, thin-lipped fish might

489   increase suction forces when the mouth is not attached to any substrate as is the

490   case when they feed in open water.

491   Transplant experiments in the field have been shown to be a successful way

492   to test the role of natural selection in divergent environments (Schluter 1995; Soria-

493   Carrasco *et al.* 2014). Using enclosures in the wild (crater lake Apoyeque) we found

494   a trend with thick-lipped fish performing better in the rocky than in the sandy habitat

495   (Fig. 2C). To our knowledge, this is one of the first studies using reciprocal transplant

496   enclosure experiments to test for a fitness trade-off associated with an ecologically

497   relevant trait in cichlid fish. Despite the fact that loss in weight was detected for

498   several enclosures, the weight change after seven weeks differed significantly in the

499   predicted direction between thick-lipped fish confined to the rocky and sandy habitat.

500   It is known that the density of prey is significantly reduced in enclosures (Schluter

501   1995). However, we decided not to provide supplementary diet in order to avoid any

502   potential bias in growth that might affect our results. Moreover, Midas cichlids growth

503   is relatively slow and, thus, the duration of the experiment might not have been long

504   enough to detect significant growth in all experimental individuals.

505

506   *Reproductive isolation: assortative mating*

507   The evolution of non-random mating is crucial in speciation driven by selection for

508   local adaptation in the presence of gene flow (Schluter 2000). Assortative mating

509   based on colour has been reported in South American cichlids (Ready *et al.* 2006).

510   Similarly, non-random mating based on a color polymorphism (gold-dark) that is

511   present in some, but not all, species of Midas cichlids has been known for some time

512   (Elmer *et al.* 2009). This suggests that active mate choice might be one of the

513   mechanisms driving divergence in this young system.

514   We observed that both ecotypes are present at a similar depth in sandy and

515   rocky habitat during the breeding season in Lake Apoyeque and several thin-lipped

516   pairs were found in both habitats. This implies that "like-pairs" are formed

517   independently of the respective preferred feeding habitats. Moreover, laboratory mate

518   choice experiments demonstrated complete assortative mating between *A. citrinellus*

519   and *A. labiatus* (Fig. 3B) suggesting that mate choice is due to direct preference,

520   rather than differences in breeding habitat choice. Not a single disassortative pair

521   was found in the laboratory experiments, suggesting that prezygotic isolating

522   mechanisms predominantly prevent hybridization. A panel of F_2 individuals from a

523   forced cross of thick- and thin-lipped fish suggests, however, that there might be also

524   some postzygotic barriers, since we found that a large number (around 10-20%) of F_2

525   individuals showed some degree of malformation, particularly in their mouths (data

526   not shown). Although thick- and thin-lipped species from great lake Nicaragua mate

527   completely assortatively in the lab, the significant but relatively low level of genetic

528   differentiation suggests a certain level of gene flow between these species. This is in

529   concordance with previously documented low levels of genetic differentiation

530   between *A. citrinellus* and *A. labiatus* based on different molecular markers

531   (Barluenga & Meyer 2010; Kautt *et al.* 2012). Although clear evidence for assortative

532   mating between thick- and thin-lipped ecotypes was found both in the wild and in lab

533   experiments, we cannot completely rule out that other visual, behavioural or chemical

534   cues are involved in pair formation. Thus, further manipulative experiments should be

535   performed to test whether hypertrophied lips act as the main mating cue.

536

537   *Genomic fingerprints of divergence in sympatry*

538   Based on population genomics approaches one can gauge how many genomic

539   regions might contribute to the evolution of reproductive isolation and hence

540   speciation (Ellegren *et al.* 2012; Jones *et al.* 2012; Feder *et al.* 2013). The patterns of

541   genome-wide genetic differentiation that we detected between both ecotypes support

542   a scenario of ecological speciation with gene flow under sympatric conditions (Feder

543   *et al.* 2012; Feder *et al.* 2013); only very few regions of the genome show signs of

544   differentiation between ecotypes in the very recently colonized crater lakes ("L-

545   shaped" frequency distribution of genetic differentiation across loci with most loci

546   having low Fst values) compared to the older Lakes Nicaragua and Managua (Fig.

547   S3B, Supporting information). Also the demographic analyses, assuming great lake

548   Managua as the source population, suggested that both crater lakes were colonized

549   very recently (580 and 1600 generations for L. Apoyeque and L. Masaya,

550   respectively). Despite the fact that only a small number of individuals were estimated

551   to have colonized the crater lakes (at least L. Apoyeque) and the much lower

552   frequency of thick-lipped fish in the great lakes (e.g. around 5%), it is likely that thick-

553   lipped fish colonized the crater lakes from the great lakes. Thus, we think that the

554   presence of the thick-lipped ecotype in the crater lakes is due to the retention of

555   ancestral standing genetic variation, rather than due to *de novo* mutations. However,

556   it seems evident that a strong homogenization event has occurred after the

557   colonization, erasing the genetic differentiation already present between the ecotypes

558   in the great lakes. Only by elucidating the genetic basis of this ecologically relevant

559   trait, would it be possible to test whether hypertrophied lips originated *de novo* in the

560   crater lakes or from standing genetic variation present in the great lakes.

561   Thick-lipped ecotype could also have been brought into the crater lake by a

562   secondary invasion from the source lakes. In this regard, our best demographic

563   models provide evidence for admixture events in both crater lakes. We note,

564   however, that we used the thin-lipped *A. citrinellus* from L. Managua as a source

565   population in these models. Our rationale for this was that the site frequency spectra

566   of both species (*A. citrinellus* and *A. labiatus*) are almost identical and our data thus

567   was not powerful enough to tease apart whether the colonization and subsequent

568   admixture happened by either one or a mix of both species (see also Kautt *et al*.

569   2016a). In any case, together with recent evidence for admixture events in

570   Nicaraguan crater lakes Apoyo, Xiloá (Kautt *et al.* 2016a), and Asososca Managua

571   (Kautt *et al.* 2016b) and Cameroonian crater lakes (Martin *et al.* 2015), this might

572   suggest that crater lakes are not as isolated as they were believed to be. In this

573   regard, fish must have somehow come into the crater lakes in the first place and

574   evidence for reasonably small admixture events indicating secondary waves of

575   colonization are not unreasonable. However, we note that distinguishing between

576   more ancient divergence events and high amounts of gene flow and very recent

577   divergence with little or no gene flow based on genetic data is challenging (Hey *et al.*

578   2015) and the validity of admixture events will have to be further tested in the future.

579   Even though some differentiation for a few markers was found, our results

580   show that there is almost no genome-wide genetic differentiation between the young

581   sympatric ecotypes in crater lake Apoyeque and only very weak genomic

582   differentiation in the relatively older crater lake Masaya. Thus, the lack of genomic

583   differentiation found in Lake Apoyeque could be explained by different factors. Firstly,

584   although it is expected that genome-wide differentiation increases as phenotypic

585   divergence increases (Roesti *et al.* 2012; Gagnaire *et al.* 2013; Seehausen *et al.*

586   2014), divergence could remain heterogeneous across the genome for a long time

587   due to repeated events of interspecific gene flow even after reproductive isolation

588   has become strong (Neafsey *et al.* 2010; Garrigan *et al.* 2012). In this regard, we

589   have found some disassortative pairs in crater lake Apoyeque (Fig. 3A), suggesting

590   that gene flow and recombination between ecotypes is plausibly generating a

591   homogenizing effect (Felsenstein 1981). Thus, ongoing gene flow between

592   populations might homogenize variation in genomic regions not affected by divergent

593   selection or reproductive isolation (Noor & Bennett 2009; Turner & Hahn 2010). Yet,

594   reduced diversity in genomic islands of speciation should also be considered as an

595   alternative explanation (Cruickshank & Hahn 2014). Secondly, due to the localized

596   genomic effect of disruptive selection, divergence during the early stages of

597   ecological speciation is expected to be greater in genomic regions that harbor key

598   quantitative loci (QTL) than it is in regions that have no effect on the phenotypic

599   divergence of the populations (Via 2009). Therefore, adaptation with gene flow tends

600   to result in concentrated genetic architectures with fewer, larger, and more tightly

601   linked divergent alleles (Yeaman & Whitlock 2011). Although the exact genetic

602   architecture of hypertrophied lips remains unknown, ongoing projects in our

603   laboratory provide evidence that only a few major QTL are involved in this trait in

604   Midas cichlids: we find intermediate phenotypes of F_1 hybrids (Machado-Schiaffino *et*

605   *al.* 2014) and a high proportion of F_2 individuals that almost completely recover both

606   parental phenotypes (Fig. S7, Supporting information). Thus, it is very probable that

607   due to the reduction of genome complexity associated with RAD-sequencing,

608   considering just a fraction of the whole genome, no markers closely linked to the QTL

609   were present in our data set.

610

611   *Progress toward sympatric speciation*

612   It has been estimated by Thibert-Plante & Gavrilets (2013), assuming a scenario of

613   relatively weak divergent selection, that 10 000 generations might be needed in order

614   to observe strong reproductive isolation (complete non-random mating). This seems

615   to be the case in thick- and thin-lipped fish in the young crater lakes, where we found

616   a clear trade-off between ecotypes, morphological bimodality, and assortative mating,

617   while our enclosure experiments suggest that the differences between ecotypes in

618   different habitats might not be very strong.

619   Recently, it has been shown for Ejagham tilapia that even in the presence of

620   strong assortative mating, the process of sympatric speciation might be stalled if

621   disruptive selection is weak (morphological unimodality and minimal ecological

622   divergence; (Martin 2013). Despite the fact that we were not able to quantify the

623   strength of disruptive selection (see (Martin 2012)) in the crater lakes, we found clear

624   morphological bimodality and ecological divergence (also see (Colombo *et al.* 2013;

625   Manousaki *et al.* 2013), suggesting that disruptive selection and assortative mating

626   acting in concert might be strong enough to lead to speciation. This process might be

627   relatively slow and yet incomplete due to the young age of the crater lakes, small

628   effective population sizes, or complexity of the genetic architectures of trophic

629   morphology. The absence of other sister species pairs in the crater lakes precluded

630   us from testing if complete divergence is possible at the same timescale, as was the

631   case in Lake Ejagham (Martin 2013).

632   Our results suggest that hypertrophied lips might be promoting incipient

633   sympatric speciation through divergent selection (ecological divergence in feeding

634   performance) and non-random mating (assortative mating) in the young Nicaraguan

635   crater lakes. All this opens the possibility of hypertrophied lips acting as a "magic

636   trait" driving speciation in these young crater lake cichlids, although further

637   manipulative experiments are necessary to confirm its role as the main cue in

638   assortative mating. Nicaraguan crater lakes are a unique natural laboratory for

639   examining, "in real time", how ecologically relevant traits under selection affect the

640   evolution of reproductive isolation during the earliest stages of speciation in

641   sympatry.

642

Acknowledgments

659

Data Accessibility

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

663

664

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.

Cruickshank 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.

714   Feder JL, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends in*
715   *Genetics* **28**, 342-350.
716   Feder JL, Flaxman SM, Egan SP, Comeault AA, Nosil P (2013) Geographic Mode of
717   Speciation and Genomic Divergence. In: *Annual Review of Ecology, Evolution, and*
718   *Systematics, Vol 44* (ed. Futuyma DJ), pp. 73-97. Annual Reviews, Palo Alto.
719   Felsenstein J (1981) Skepticism towards santa rosalia, or why are there so few kinds of
720   animals. *Evolution* **35**, 124-138.
721   Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) version 3.6a2. Department of
722   Genetics, University of Washington, Seattle.
723   Fitzpatrick BM, Fordyce JA, Gavrilets S (2008) What, if anything, is sympatric speciation?
724   *Journal of Evolutionary Biology* **21**, 1452-1459.
725   Gagnaire PA, Pavey SA, Normandeau E, Bernatchez L (2013) The genetic architecture of
726   reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs
727   assessed by RAD sequencing. *Evolution* **67**, 2483-2497.
728   Garrigan D, Kingan SB, Geneva AJ*, et al.* (2012) Genome sequencing reveals complex
729   speciation in the *Drosophila simulans* clade. *Genome Research* **22**, 1499-1511.
730   Gavrilets S (2004) *Fitness Landscapes and the Origin of Species* Princeton University Press,
731   Princeton.
732   Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the Joint
733   Demographic History of Multiple Populations from Multidimensional SNP Frequency
734   Data. *Plos Genetics* **5**, 11.
735   Hatfield T, Schluter D (1999) Ecological speciation in sticklebacks: Environment-dependent
736   hybrid fitness. *Evolution* **53**, 866-873.
737   Hendry AP, Taylor EB, McPhail JD (2002) Adaptive divergence and the balance between
738   selection and gene flow: Lake and stream stickleback in the misty system. *Evolution*
739   **56**, 1199-1216.
740   Henning F, Jones JC, Franchini P, Meyer A (2013) Transcriptomics of morphological color
741   change in polychromatic Midas cichlids. *Bmc Genomics* **14**, 13.
742   Henning F, Machado-Schiaffino G, Baumgarten L, Meyer A (2017) Genetic dissection of
743   adaptive form and function in rapidly-speciating cichlid fishes. *Evolution* **in press**.
744   Henning F, Meyer A (2014) The evolutionary genomics of cichlid fishes: explosive speciation
745   and adaptation in the postgenomic era. *Annual Reviews of Genomics and Human*
746   *Genetics* **15**, 417–441.
747   Hey J, Chung Y, Sethuraman A (2015) On the occurrence of false positives in tests of
748   migration under an isolation-with-migration model. *Molecular Ecology* **24**, 5078-
749   5083.
750   Holzman R, Collar DC, Mehta RS, Wainwright PC (2012) An integrative modeling approach
751   to elucidate suction-feeding performance. *Journal of Experimental Biology* **215**, 1-13.
752   Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies.
753   *Molecular Biology and Evolution* **23**, 254-267.
754   Jones FC, Grabherr MG, Chan YF*, et al.* (2012) The genomic basis of adaptive evolution in
755   threespine sticklebacks. *Nature* **484**, 55-61.
756   Kautt AF, Elmer KR, Meyer A (2012) Genomic signatures of divergent selection and
757   speciation patterns in a natural experiment', the young parallel radiations of
758   Nicaraguan crater lake cichlid fishes. *Molecular Ecology* **21**, 4770-4786.
759   Kautt AF, Machado-Schiaffino G, Meyer A (2016a) Multispecies Outcomes of Sympatric
760   Speciation after Admixture with the Source Population in Two Radiations of
761   Nicaraguan Crater Lake Cichlids. *Plos Genetics* **12**, e1006157.
762   Kautt AF, Machado-Schiaffino G, Torres-Dowdall J, Meyer A (2016b) Incipient sympatric
763   speciation in Midas cichlid fish from the youngest and one of the smallest crater lakes

764   in Nicaragua due to differential use of the benthic and limnetic habitats? *Ecology and*
765   *Evolution* **6**, 5342-5357.
766   Kohda M, Tanida K (1996) Overlapping territory of the benthophagous cichlid fish,
767   *Lobochilotes labiatus*, in Lake Tanganyika. *Environmental Biology of Fishes* **45**, 13-
768   20.
769   Konings A (1998) *Tanganyika Cichlids in their natural habitats* Cichlid Press, St Leon-Rot.
770   Kutterolf S, Freundt A, Perez W, Wehrmann H, Schmincke HU (2007) Late Pleistocene to
771   Holocene temporal succession and magnitudes of highly-explosive volcanic eruptions
772   in west-central Nicaragua. *Journal of Volcanology and Geothermal Research* **163**, 55-
773   82.
774   Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler
775   transform. *Bioinformatics* **25**, 1754-1760.
776   Losos JB (2011) Convergence, adaptation and constraint. *Evolution* **65**, 1827-1840.
777   Machado-Schiaffino G, Henning F, Meyer A (2014) Species-specific differences in adaptive
778   phenotypic plasticity in an ecologically relevant trophic trait: hypertrophic lips in
779   Midas cichlid fishes. *Evolution*, 2086-2091.
780   Malinsky M, Challis RJ, Tyers AM*, et al.* (2015) Genomic islands of speciation separate
781   cichlid ecomorphs in an East African crater lake. *Science* **350**, 1493-1498.
782   Manousaki T, Hull PM, Kusche H*, et al.* (2013) Parsing parallel evolution: ecological
783   divergence and differential gene expression in the adaptive radiations of thick-lipped
784   Midas cichlid fishes from Nicaragua. *Molecular Ecology* **22**, 650-669.
785   Martin CH (2012) Weak Disruptive Selection and Incomplete Phenotypic Divergence in Two
786   Classic Examples of Sympatric Speciation: Cameroon Crater Lake Cichlids. *American*
787   *Naturalist* **180**, E90-E109.
788   Martin CH (2013) Strong assortative mating by diet, color, size, and morphology but limited
789   progress toward sympatric speciation in a classic example: Cameroon crater lake
790   cichlids. *Evolution* **67**, 2114-2123.
791   Martin CH, Cutler JS, Friel JP*, et al.* (2015) Complex histories of repeated gene flow in
792   Cameroon crater lake cichlids cast doubt on one of the clearest examples of sympatric
793   speciation. *Evolution* **69**, 1406-1422.
794   Mayr E (1942) *Systematics and the origin of species from the viewpoint of a zoologist*
795   Harvard University Press, Cambridge, MA.
796   Meyer A, Kautt AF (2014) Sympatric speciation. In: *Oxford Bibliographies in Evolutionary*
797   *Biology* (ed. Losos JB). Oxford University Press, New York.
798   Neafsey DE, Lawniczak MKN, Park DJ*, et al.* (2010) SNP Genotyping Defines Complex
799   Gene-Flow Boundaries Among African Malaria Vector Mosquitoes. *Science* **330**, 514-
800   517.
801   Noor MAF, Bennett SM (2009) Islands of speciation or mirages in the desert? Examining the
802   role of restricted recombination in maintaining species. *Heredity* **103**, 439-444.
803   Nosil P (2012) *Ecological speciation* Oxford University Press, Oxford, UK.
804   Oliver MK, Arnegard ME (2010) A new genus for Melanochromis labrosus, a problematic
805   Lake Malawi cichlid with hypertrophied lips (Teleostei: Cichlidae). *Ichthyological*
806   *Exploration of Freshwaters* **21**, 209-232.
807   Papadopulos AST, Baker WJ, Crayn D*, et al.* (2011) Speciation with gene flow on Lord
808   Howe Island. *Proceedings of the National Academy of Sciences of the United States of*
809   *America* **108**, 13188-13193.
810   Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *Plos Genetics*
811   **2**, 2074-2093.
812   Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double Digest RADseq:
813   An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and
814   Non-Model Species. *Plos One* **7**.

815   R Development Core Team (2014) R: a Language and Environment for Statistical Computing.
816   Ready JS, Sampao I, Schneider H, *et al.* (2006) Colour forms of Amazonian cichlid fish
817   represent reproductively isolated species. *Journal of Evolutionary Biology* **19**, 1139-
818   1148.
819   Recknagel H, Elmer KR, Meyer A (2013) A hybrid genetic linkage map of two ecologically
820   and morphologically divergent Midas cichlid fishes (*Amphilophus* spp.) obtained by
821   massively parallel DNA sequencing (ddRADSeq). *G3-GENES GENOMES*
822   *GENETICS*.
823   Roesti M, Hendry AP, Salzburger W, Berner D (2012) Genome divergence during
824   evolutionary diversification as revealed in replicate lake-stream stickleback population
825   pairs. *Molecular Ecology* **21**, 2852-2862.
826   Rueffler C, Van Dooren TJM, Leimar O, Abrams PA (2006) Disruptive selection and then
827   what? *Trends in Ecology & Evolution* **21**, 238-245.
828   Rundle HD, Whitlock MC (2001) A genetic interpretation of ecologically dependent isolation.
829   *Evolution* **55**, 198-201.
830   Schliewen UK, Tautz D, Paabo S (1994) Sympatric speciation suggested by monophyly of
831   crater lake cichlids. *Nature* **368**, 629-632.
832   Schluter D (1995) Adaptive radiation in sticklebacks - Trade-offs in feeding performance and
833   growth. *Ecology* **76**, 82-90.
834   Schluter D (2000) *The Ecology of Adaptive Radiation* Oxford University Press, Oxford.
835   Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image
836   analysis. *Nature Methods* **9**, 671-675.
837   Seehausen O, Butlin RK, Keller I, *et al.* (2014) Genomics and the origin of species. *Nature*
838   *Reviews Genetics* **15**, 176-192.
839   Servedio MR, Van Doorn GS, Kopp M, Frame AM, Nosil P (2011) Magic traits in speciation:
840   'magic' but not rare? *Trends in Ecology & Evolution* **26**, 389-397.
841   Skorczewski T, Cheer A, Wainwright PC (2012) The benefits of planar circular mouths on
842   suction feeding performance. *Journal of the Royal Society Interface* **9**, 1767-1773.
843   Soria-Carrasco V, Gompert Z, Comeault AA, *et al.* (2014) Stick Insect Genomes Reveal
844   Natural Selection's Role in Parallel Speciation. *Science* **344**, 738-742.
845   Thibert-Plante X, Gavrilets S (2013) Evolution of mate choice and the so-called magic traits
846   in ecological speciation. *Ecology Letters* **16**, 1004-1013.
847   Turner TL, Hahn MW (2010) Genomic islands of speciation or genomic islands and
848   speciation? *Molecular Ecology* **19**, 848-850.
849   Via S (2009) Natural selection in action during speciation. *Proceedings of the National*
850   *Academy of Sciences of the United States of America* **106**, 9939-9946.
851   Yeaman S, Whitlock MC (2011) The genetic architecture of adaptation under migration-
852   selection balance. *Evolution* **65**, 1897-1911.
853

854

Author contributions

856   G.M.S and A.M designed the experiments. G.M.S analyzed the morphological data.

857   G.M.S and J.T.D performed and analyzed the experiments in the field. G.M.S, A.F.K

858   and J.T.D performed and analyzed assortative mating experiments in the lab. G.M.S,

859   L.B and F.H performed and analyzed the feeding performance experiments in the

860   lab. Population genomics data were generated and analyzed by A.F.K and G.M.S.

861   Demographic analyses were performed by A.F.K. G.M.S and F.H designed the

862   crosses and obtained the F2 populations. G.M.S wrote the manuscript with

863   contributions from all the authors.

864

865   **Figure Legends**

866   **Fig. 1** Thin- and thick-lipped populations in the Midas cichlid species complex. (A)

867   Map showing the four lakes where thin- and thick-lipped Midas cichlids coexist.

868   Admixture plot showing clear genetic clustering among the four lakes together with

869   cross-validation errors for admixture runs are shown below the map. The lowest

870   cross-validation error (i.e. highest support) was found for four genetic clusters. (B, C,

871   D, E) Histograms and Kernel density plots of normalized lip area (lip area corrected

872   by body area) in wild-caught individuals from the great lakes Managua (purple) and

873   Nicaragua (green) and the crater lakes Apoyeque (red) and Masaya (blue). Sample

874   sizes (N), Hartigans' Dip tests for unimodality (D) and p-values are shown for each

875   lake.

876

877   **Fig. 2** Functional trade-off between ecotypes. (A) Correlation between normalized lip

878   area and feeding performance (distance from the base of the acrylic device to the

879   furthermost removed mosquito larvae) for wild caught thick- and thin-lipped fish from

880   crater lake Apoyeque exposed to mosquito larvae attached to an acrylic structure

881   with continuously decreasing angle. (B) Correlation between normalized lip area and

882   feeding success ratio (number of eaten prey items divided by the number of

883   attempts) for wild caught thick- and thin-lipped fish from crater lake Apoyeque

884   exposed to free swimming fry. (C) Differential performance (mean change in weight

885   per day in grams and standard deviation) between thick- and thin-lipped ecotypes in

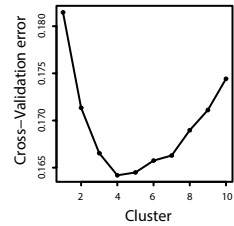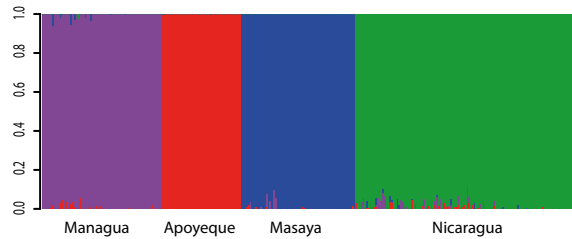886   rocky and sandy habitats areas in crater lake Apoyeque.
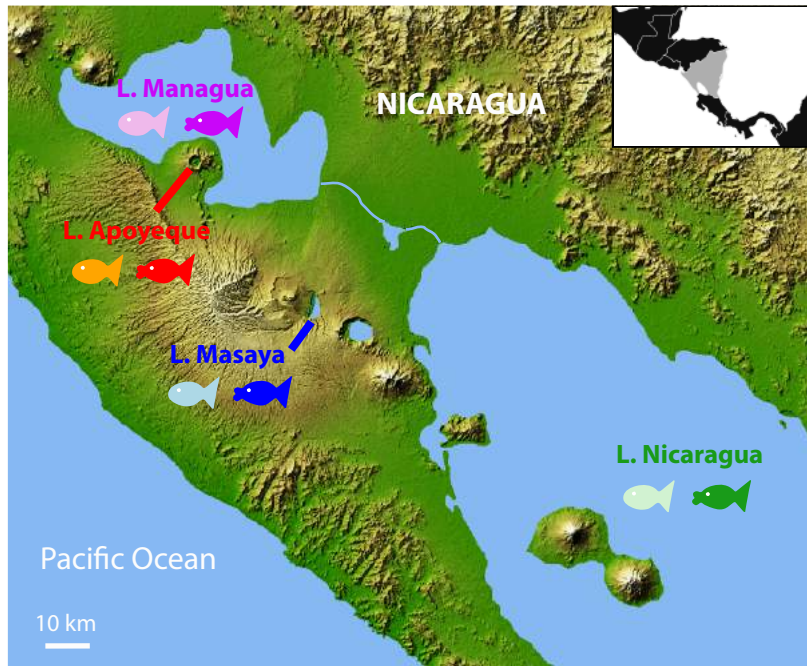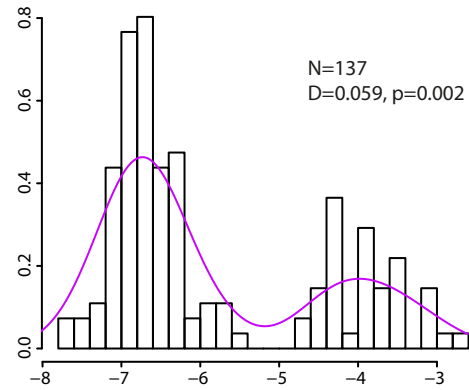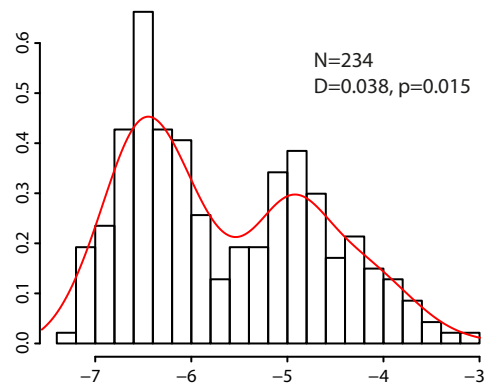
887

888   **Fig. 3** Assortative mating by ecotypes. (A) Assortative mating in crater lake

889   Apoyeque. Census of pairs during breeding season 2013-2014. Observed number of

890   pairs (red/orange) and expected under random mating (grey) are shown. (B) Strong

891   Assortative mating under laboratory conditions between *A. citrinellus* (thin-lipped) and

892   *A. labiatus* (thick-lipped) from great lake Nicaragua. Observed number of pairs (green

893   or light green) and expected under random mating (grey) are shown. (C) Correlation

894   between normalized lip area of mating females and males under laboratory
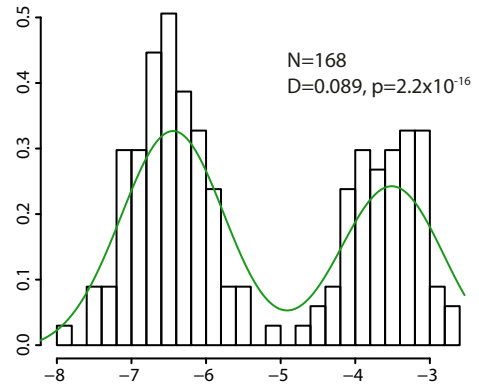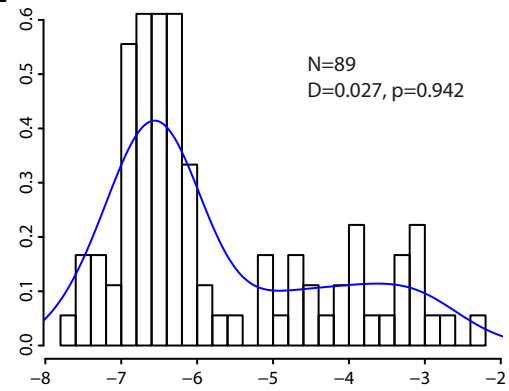
895   conditions.

896

897   **Fig. 4** Genetic relationship and differentiation between ecotypes. (A) Neighbor-joining

898   tree showing that fish from crater lakes Apoyeque (red/orange) and Masaya

899   (blue/light blue) share a more recent common ancestor with those from great lake

900   Managua than great lake Nicaragua (purple/light purple). Thick- and thin-lipped

901   ecotypes from great lake Nicaragua are shown in green and light green, respectively.

902   (B, C) First three main axes of genetic variation (Principal Component Analysis)

903   between ecotypes in the great lakes Managua (purple/light purple) and Nicaragua

904   (green/light green). The first principal component of the PCA was highly significant

905   (p-value ~ 0) for both great lakes. Analyses based on 16 286 and 21 465 polymorphic

906   loci for great lake Managua and Nicaragua, respectively. (D, E) First three main axes

907   of genetic variation (principal components) between ecotypes in the crater lakes

908   Apoyeque (red/orange) and Masaya (blue/light blue). The first principal component of

909   the PCA was highly significant for crater lake Masaya but only slightly significant for

910   L. Apoyeque (p = 9.34×10^{-29} and p = 0.011, respectively). Analyses based on 13 547

911　　and 6693 polymorphic loci for crater Masaya and Apoyeque, respectively. (F)

912　　Schematic illustrations of the most supported demographic models of crater lakes

913　　Apoyeque and Masaya assuming great lake Managua as the source population.
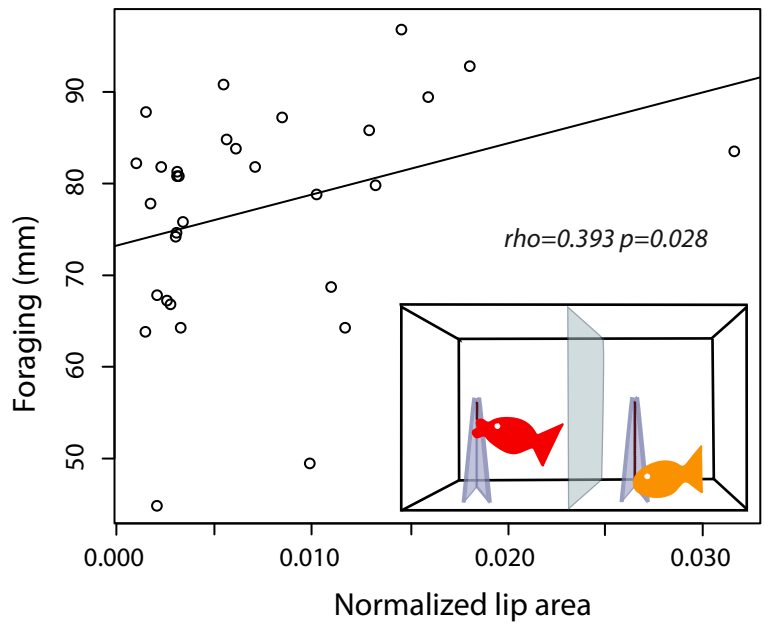
914

915

Log(normalized lip area)

Log(normalized lip area)

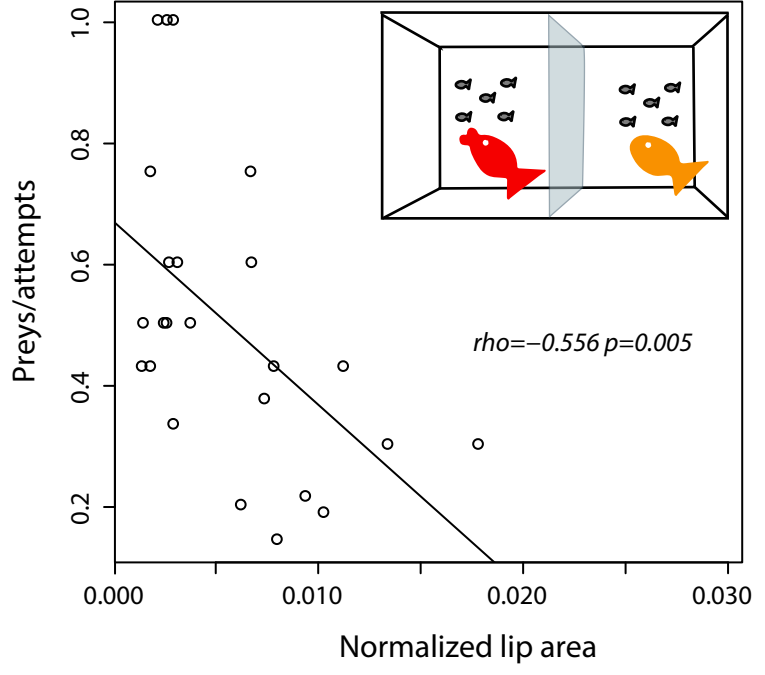**A**

rho=0.393 p=0.028

B

rho=−0.556 p=0.005

C

A

Crater Lake Apoyeque (wild)

N = 68
Yates' X^2 = 25.766, p=3.46x10^{-6}

- Observed
- Expected

Number of pairs

61   57.45   3   10.1   4   0.45

B

Great Lake Nicaragua (lab)

N = 25
Yates' X^2 = 22.18, p=1.53x10^{-5}

- Observed
- Expected

Number of pairs

12   6.25   0   12.5   13   6.25

C

rho=-0.071 p=0.821

Normalized lip area_males

Normalized lip area_females

A

1000

534

977

1000

528

977

1000

A. centrarchus

B Great lake Managua
16 286 Loci
PC3
PC2
PC1

C Great lake Nicaragua
21 465 Loci
PC3
PC2
PC1

D Crater lake Apoyeque
6693 Loci
PC3
PC2
PC1

E Crater lake Masaya
13 547 Loci
PC3
PC2
PC1

F

$N_0 = 19\ 757$

$N_{bottle} = 1484$

$T_{div} = 577; N = 110$

$T_{admix} = 376$
Admix = 16%

$N_1 = 460{,}267$

$N_1 = 14\ 716$

L. Managua   **L. Apoyeque**

past

coalescence

present

$N_0 = 18\ 380$

$N_{bottle} = 1696$

$T_{div} = 1560$

$T_{admix} = 189$
Admix = 20%

$N_1 = 683{,}670$

$N_1 = 8610$

L. Managua   **L. Masaya**