

The ecological and genetic basis of convergent thick-lipped phenotypes in cichlid fishes

MARCO COLOMBO,*¹ EVELINE T. DIEPEVEEN,*¹ MORITZ MUSCHICK,*[‡] M. EMILIA SANTOS,* ADRIAN INDERMAUR,* NICOLAS BOILEAU,* MARTA BARLUENGA[†] and WALTER SALZBURGER*

*Zoological Institute, University of Basel, Vesalgasse 1, 4051, Basel, Switzerland, [†]Museo Nacional de Ciencias Naturales, CSIC, José Gutierrez Abascal 2, 28006, Madrid, Spain

Abstract

The evolution of convergent phenotypes is one of the most interesting outcomes of replicate adaptive radiations. Remarkable cases of convergence involve the thick-lipped phenotype found across cichlid species flocks in the East African Great Lakes. Unlike most other convergent forms in cichlids, which are restricted to East Africa, the thick-lipped phenotype also occurs elsewhere, for example in the Central American Midas Cichlid assemblage. Here, we use an ecological genomic approach to study the function, the evolution and the genetic basis of this phenotype in two independent cichlid adaptive radiations on two continents. We applied phylogenetic, demographic, geometric morphometric and stomach content analyses to an African (*Lobochilotes labiatus*) and a Central American (*Amphilophus labiatus*) thick-lipped species. We found that similar morphological adaptations occur in both thick-lipped species and that the 'fleshy' lips are associated with hard-shelled prey in the form of molluscs and invertebrates. We then used comparative Illumina RNA sequencing of thick vs. normal lip tissue in East African cichlids and identified a set of 141 candidate genes that appear to be involved in the morphogenesis of this trait. A more detailed analysis of six of these genes led to three strong candidates: *Actb*, *Cldn7* and *Copb*. The function of these genes can be linked to the loose connective tissue constituting the fleshy lips. Similar trends in gene expression between African and Central American thick-lipped species appear to indicate that an overlapping set of genes was independently recruited to build this particular phenotype in both lineages.

Keywords: adaptive radiation, cichlid species flocks, convergent evolution, East Africa, ecological genomics, RNAseq

Received 9 March 2012; revision received 4 July 2012; accepted 15 July 2012

Introduction

Adaptive radiation is the rapid evolution of an array of species from a common ancestor as a consequence of the emerging species' adaptations to distinct ecological niches (Simpson 1953; Schluter 2000; Gavrillets & Losos 2009). It is typically triggered by ecological opportunity

in form of underutilized resources—just as being provided after the colonization of a new habitat, the extinction of antagonists and/or the evolution of a novel trait, which is then termed an evolutionary 'key innovation' (Gavrillets & Vose 2005; Gavrillets & Losos 2009; Losos & Ricklefs 2009; Losos 2010; Yoder *et al.* 2010; Matschiner *et al.* 2011). Whatever the circumstances were that initiated an adaptive radiation, there is always a strong link between adaptively relevant traits and the habitat and/or foraging niche (a 'phenotype–environment correlation'; Schluter 2000). In the most illustrative examples of adaptive radiation, the Darwin's finches on the Galapagos archipelago, the *Anolis* lizards on the

Correspondence: Walter Salzburger, Fax: +41 61 267 0301;

E-mail: walter.salzburger@unibas.ch

[‡] Present address: Department of Animal and Plant Sciences, The University of Sheffield, Sheffield, S10 2TN, UK.

¹These authors contributed equally to this work.

Caribbean islands and the cichlid fishes of the East African Great Lakes, this correlation exists between beak-shape and food source (finches), limb morphology and twig diameter (anoles), and the architecture of the mouth and jaw apparatus and foraging mode (cichlids) (Schluter 2000; Butler *et al.* 2007; Grant & Grant 2008; Losos 2009; Salzburger 2009).

An interesting aspect of many adaptive radiations is the frequent occurrence of convergent (or parallel) evolution (Schluter & Nagel 1995; Harmon *et al.* 2005; Arendt & Reznick 2008; Losos 2011; Wake *et al.* 2011). For example, similar ecotype morphs of anoles lizards have evolved independently on different Caribbean islands (Losos *et al.* 1998; Harmon *et al.* 2005; Losos & Ricklefs 2009), benthic–limnetic and lake–stream species pairs of threespine sticklebacks emerged repeatedly in and around postglacial lakes (Rundle *et al.* 2000; Berner *et al.* 2010; Roesti *et al.* 2012), and a whole array of convergent forms of cichlid fish emerged between the lakes of East Africa (Kocher *et al.* 1993; Salzburger 2009). Such instances of convergent evolution are generally interpreted as the result of the action of similar selection regimes in isolated settings (Schluter & Nagel 1995; Rundle *et al.* 2000; Nosil *et al.* 2002; Harmon *et al.* 2005; Losos 2011). It has further been suggested that if radiations are truly replicated (i.e. driven by adaptive processes), convergence in morphology should tightly be associated with convergence in ecology and behaviour (Johnson *et al.* 2009).

The species flocks of cichlid fishes in the East African Great Lakes Victoria, Malawi and Tanganyika represent the most species-rich extant adaptive radiations in vertebrates (Kocher 2004; Seehausen 2006; Salzburger 2009). Several hundreds of endemic cichlid species have emerged in each lake within a period of several millions of years (as is the case for Lake Tanganyika; Salzburger *et al.* 2002; Genner *et al.* 2007) to <150 000 years (as in Lake Victoria; Verheyen *et al.* 2003). The various endemic cichlid species differ greatly in the morphology of the trophic apparatus (mouth form and shape, jaw structure and dentition) as well as in coloration and pigmentation, suggesting that both natural and sexual selection are jointly responsible for adaptive radiation and explosive speciation in cichlids (Salzburger 2009). Interestingly, convergent forms that emerged in independent cichlid adaptive radiations often show very similar coloration patterns in addition to matching body shapes and mouth morphologies (Kocher *et al.* 1993; Stiassny & Meyer 1999; Salzburger 2009). This has led to speculations whether selection alone is sufficient to explain convergence, or whether genetic or developmental constraints have contributed to the morphogenesis of these matching phenotypes (Brakefield 2006).

The present study focuses on the morphology, ecology and the genetic basis of a peculiar mouth trait in cichlid fishes, which has evolved multiple times: hypertrophied ('fleshy') lips (see Box 1 in Salzburger 2009). The exact function of the thick lips in cichlids is unknown, although this feature is generally implicated in a specific foraging mode (Fryer 1959; Fryer & Iles 1972; Arnegard *et al.* 2001). Fleshy lips are often interpreted as an adaptation for feeding on invertebrates and crustaceans hidden in crannies, with the lips being used to seal cracks and grooves to facilitate the sucking of prey (Barlow & Munsey 1976; Ribbink *et al.* 1983; Seehausen 1996; Konings 1998). Alternatively, it has been suggested that hypertrophied lips protect from mechanical shocks (Greenwood 1974; Yamaoka 1997), and that they function as taste receptors (Arnegard *et al.* 2001) or as mechanoreceptors (Fryer 1959; Fryer & Iles 1972). [Note, however, that there is no increase in sensory cells in lip tissue (Greenwood 1974).]

It is remarkable that thick-lipped species appear to be a common outcome of cichlid adaptive radiations. For example, the large cichlid assemblages in East Africa all contain at least one such taxon (Lake Victoria: *Haplochromis chilotes*; Lake Malawi: *Chilotilapia euchilus*, *Abactochromis labrosus*, *Otopharynx pachycheilus*, *Placidochromis milomo*, *Protomelas ornatus*; Lake Tanganyika: *Lobochilotes labiatus*). In addition, cichlids featuring hypertrophied lips are known from, for example, the Midas Cichlid (*Amphilophus* spp.) assemblage in the large lakes of Nicaragua, where a thick-lipped species (*A. labiatus*) is common in rocky habitats (Fig. 1). Occasionally, hypertrophied lips are also observed in other related cichlids in Nicaragua, such as in the riverine species *Tomacichla tuba* (Villa 1982) or in *Astatheros rostratus* (pers. obs.). Additional riverine representatives with hypertrophied lips are also found in South America (*Crenicichla tendybaguassu*) and Western Africa (*Thoracochromis albolabris*). Hypertrophied lips are not unique to cichlids, though. For example, the adaptive radiation of the sailfin silver-side fish (Telmatherinidae) in the Malili lakes of Sulawesi (Herder *et al.* 2006) and the barbs of Lake Tana in Ethiopia (Sibbing *et al.* 1998; de Graaf *et al.* 2008) also produced thick-lipped species.

Members of the family Cichlidae are distributed in the Southern hemisphere, with a few ancestral lineages in India, Sri Lanka and Madagascar and two exceptionally species-rich clades, one in Central and South America and one in Africa (Salzburger & Meyer 2004). This biogeographical pattern is consistent with a Gondwanan origin of the Cichlidae, dating the split between American and African representatives to ~100 Ma (Salzburger & Meyer 2004; Sereno *et al.* 2004; Genner *et al.* 2007). This set-up opens the possibility to study the ecological and genetic basis of a convergent trait across one of the

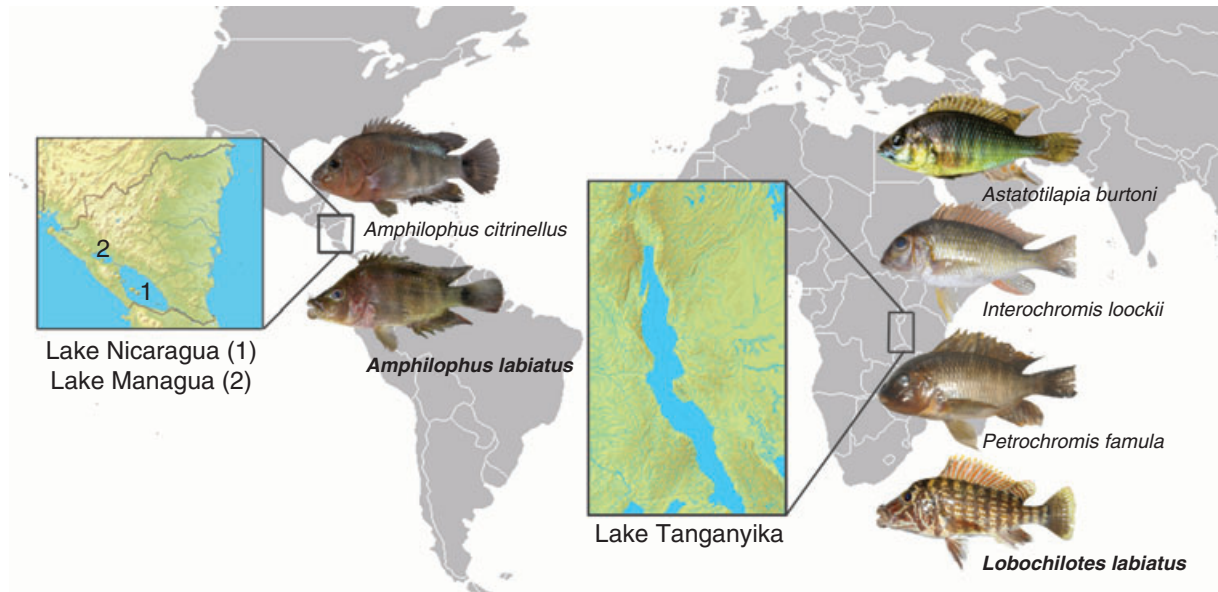


Fig. 1 Map of the Southern hemisphere showing the two study systems, the Midas Cichlid (*Amphilophus* sp.) species complex in Nicaragua, Central America, and the Tropheini in Lake Tanganyika, East Africa.

largest possible phylogenetic and geographical distances in cichlids and, hence, in the complete absence of gene flow and outside the influence of ancestral polymorphism and/or standing genetic variation.

Here, we applied an integrative approach in two cichlid fish radiations, the one of the Tropheini in East African Lake Tanganyika and the Midas Cichlid assemblage in Nicaragua, to uncover the ecological and genetic basis of the thick-lipped phenotype. More specifically, we compared the two '*labiatus*' species to one another and to their sister species by means of geometric morphometric and stomach content analyses; we placed them in their respective radiations by phylogenetic and demographic analyses; and we provide field observations on foraging strategies for one of them (*L. labiatus*). To study the genetic basis of hypertrophied lips, we first applied comparative transcriptome analyses (RNA-seq) on the basis of Illumina next-generation sequencing of juvenile and adult individuals of the African species *L. labiatus* (in comparison with a closely related species for which a genome sequence is available). In a second step, we tested candidate genes identified by RNAseq in representatives of both radiations in a quantitative real-time PCR environment.

Materials and methods

Study species

This study focuses on two thick-lipped species, *Loboichilotes labiatus* from East African Lake Tanganyika and *Amphilophus labiatus* from Nicaragua. *Loboichilotes labiatus* is

a member of the rock-dwelling Tanganyikan cichlid tribe Tropheini and therefore part of the most species-rich group of cichlids, the haplochromines, which include the Tanganyikan Tropheini, many riverine species and the species flocks of Lakes Victoria and Malawi (Salzburger *et al.* 2002, 2005). The Tropheini themselves underwent a subradiation within Lake Tanganyika (see e.g. Sturmbauer *et al.* 2003). *Amphilophus labiatus* is part of the Midas Cichlid assemblage in Nicaragua and occurs in the large Central American lakes Managua and Nicaragua, where it co-occurs with the most common species in the area, *A. citrinellus* (Barlow 1976; Barluenga & Meyer 2010). For this study, we sampled a total of 84 and 74 specimens of the Central American species *Amphilophus citrinellus* and *A. labiatus*, respectively, and 143 specimens of *L. labiatus* plus 14 additional Haplochromini/Tropheini specimens from Lake Tanganyika. Exact sampling locations and dates for specimens used for the genetic analysis and GenBank accession numbers are provided in Appendix S1.

Sampling, DNA and RNA extraction

Sampling of *L. labiatus* and other Tropheini species was performed between 2007 and 2011 in the Southern part of Lake Tanganyika, East Africa; *A. labiatus* and its congeners were collected in September 2009 in the two large Nicaraguan lakes Managua and Nicaragua (see Appendix S1 for details). Fishes were processed in the field following our standard operating procedure: fishes were individually labelled, measured (total and standard length) and weighted and a photograph was taken from the left side

of each specimen using a Nikon P5000 or a Nikon D5000 digital camera (fins were spread out using clips); then, a piece of muscle tissue and a fin-clip were taken as DNA sample and preserved in ethanol; fishes were then dissected and RNA samples from lip and other tissues were preserved in RNAlater (Ambion); the whole intestinal tract was removed and stored in ethanol.

For DNA extraction, we either applied a high-salt extraction method (Bruford *et al.* 1998) or used a Mag-naPure extraction robot (Roche, Switzerland) following the manufacturer's protocol. RNA was extracted according to the Trizol method with either Trizol (Invitrogen) or TRI reagent (Sigma). Lip tissue was homogenized with a PRO200 Homogenizer (PRO Scientific Inc.) or with a BeadBeater (FastPrep-24; MP Biomedicals). DNase treatment following the DNA Free protocol (Ambion) was performed to remove any genomic DNA from the samples. Subsequent reverse transcription was achieved by using the High Capacity RNA-to-cDNA kit (Applied Biosystems). For the *A. burtoni* samples, up to two individuals (adults) or up to eight individuals (juveniles) were used per sample, due to a diminutive amount of lip tissue extracted from these fishes. All other samples were taken from a single specimen.

Phylogenetic and demographic analyses

We first wanted to phylogenetically place the thick-lipped species into the respective clade of East African and Nicaraguan cichlids. We thus performed a phylogenetic analysis of the Tanganyikan cichlid tribe Tropheini (see also Sturmbauer *et al.* 2003) and used haplotype genealogies to reconstruct the evolutionary history in the much younger *Amphilophus* species assemblage in Nicaragua, where phylogenetic analyses are not expedient due to the lack of phylogenetic signal (see also Barluenga *et al.* 2006; Barluenga & Meyer 2010). We also performed mismatch analyses within *A. citrinellus*, *A. labiatus* and *L. labiatus* to compare their demographic histories.

We amplified three gene segments for each of the three focal species and additional Tropheini/Haplochromini species: the first segment of the noncoding mtDNA control region and two nuclear loci containing coding and noncoding DNA (a segment each of the *endothelin receptor 1*, *ednrb1* and the *phosphatidyl phosphatase 1*, *phpt1*). We used previously published primers L-Pro-F (Meyer *et al.* 1994) and TDK-D (Lee *et al.* 1995) for the control region and *ednrb1F* and *ednrb1R* (Lang *et al.* 2006) for *ednrb1*, and so far unpublished primers 38a_F (5'-AGC AGG GTT GAC CTT CTC AA-3') and 38a_R (5'-TGG CTA AAA TCC CCG ATG TA-3') for *phpt1*. Polymerase chain reaction (PCR) amplification, purification and cycle sequencing were performed as described elsewhere (Diepeveen & Salzburger 2011); an

ABI 3130xl capillary genetic analyzer (Applied Biosystems) was used for DNA sequencing.

The resulting sequences were complemented with already available sequences. In the case of the Tropheini, we also included available sequences of the mitochondrial NADH dehydrogenase subunit 2 gene (ND2) (see Appendix S1 for GenBank accession numbers). Sequences were aligned with MAFFT (Katoh & Toh 2008) resulting in a total length of 2345 bp for the Tropheini (control region: 371 bp; ND2: 1047 bp; *ednrb1*: 538 bp; *phpt1*: 389 bp) and 1620 bp for *Amphilophus* (control region: 371 bp; *ednrb1*: 743 bp; *phpt1*: 469 bp). Maximum-likelihood and Bayesian inference phylogenetic analyses of the Tropheini were performed for each gene segment separately (not shown) and for a concatenated alignment with PAUP* (Swofford 2003) and MRBAYES (Ronquist & Huelsenbeck 2003), respectively. The appropriate model of sequence evolution was detected with jMODELTEST (Posada 2008) applying the Akaike Information Criterion (AIC). A maximum-likelihood bootstrap analysis with 100 pseudoreplicates was performed in PAUP*, and MR. BAYES was run for eight million generations with a sample frequency of 100 and a burn-in of 10%. We then used MESQUITE (www.mesquiteproject.org) to map feeding specializations on the resulting maximum-likelihood topology and to reconstruct ancestral character states with parsimony. Data on feeding mode from the Haplochromini/Tropheini species other than *L. labiatus* are based on Brichard (1989), Nori (1997), Yamaoka (1997) and Konings (1998).

Haplotype genealogies for the *Amphilophus* data set were constructed following the method described in the study by Salzburger *et al.* (2011) on the basis of a maximum-likelihood tree and sequences of the mitochondrial control region and the nuclear *ednrb1* gene (*phpt1* was not used here due to the limited number of haplotypes found). Mismatch analyses were performed on the basis of mtDNA sequences with ARLEQUIN 3.0 (Excoffier *et al.* 2005).

Geometric morphometric analyses

In order to test for similarities in overall body shape between the thick-lipped forms from Central America and East Africa, we performed geometric morphometric analyses on the basis of digital images. Body shape was quantified in a set of 58 *A. citrinellus*, 27 *A. labiatus* and 27 *L. labiatus* using 17 homologous landmarks (see Appendix S2; note that lip shape was not assessed to prevent a bias). Data acquisition was carried out using TPSDIG (Rohlf 2006), and data were analysed with MORPHOJ (Klingenberg 2011). For all shape comparisons, we used the residuals of a within-species regression of shape on centroid size to reduce allometric effects within species, in

order to retain shape differences between differently sized species. For the same reason, we only included *L. labiatus* individuals with a body size larger than 12 cm total length. We then performed a discriminant function analysis between all pairs of species and a principal component analysis (PCA). To identify morphological changes associated with the enlarged lip phenotype, we compared *A. labiatus* to its closest relative, *A. citrinellus*. In the case of *L. labiatus*, we made use of our new phylogeny of the Tropheini (Fig. 2a) and body shape data of *L. labiatus* and its nine closest relatives [*Petrochromis macrognathus*, *P. polyodon*, *P. ephippium*, *Lobochilotes labiatus*, *Simochromis diagramma*, *S. babaulti*, *Gnathochromis pfefferi*, *Pseudosimochromis curvifrons*, *Limnotilapia dardenni* and *Ctenochromis horei* (M. Muschick, A. Indermaur & W. Salzburger, unpublished data)] to reconstruct the landmark configuration of the direct ancestor to *L. labiatus*. This was carried out in MORPHOJ using branch length-weighted squared-change parsimony. The changes in landmark configurations along a discriminant function (Nicaraguan species) or along the shape-change vector from the estimated ancestral shape to *L. labiatus* were increased threefold to produce Fig. 3. The shape differences between species shown in Fig. 3 accurately reflect the shape-change vectors for landmark positions. Outlines were interpolated and added to Fig. 3 to help the reader envision these shape differences in the context of fish body shape.

Stomach and gut content analyses

To assess trophic specialization of the thick-lipped cichlid species, we performed comparative stomach and gut content analyses. To this end, stomachs and guts were opened step-by-step. First, the stomach was opened and emptied under a binocular followed by the remaining parts of the intestine. All items were grouped into seven food categories: hard-shelled (crustaceans, snails, mussels), small arthropods (insects and zooplankton), fish scales, fish remains, plant seeds and plant material other than seeds. For each specimen, the wet weight of each food category was measured on a Kern ALS 120-4 scale (Kern, Germany) and was then used to calculate Schoener's index of proportional diet overlap (Schoener 1970). We analysed stomach and gut contents in a total of 159 specimens: *A. citrinellus* ($N = 58$; of which 25 had contents), *A. labiatus* ($N = 62$; 34) and *L. labiatus* ($N = 39$; 29). We note that such an analysis has the drawback that it only covers food uptake in the last few hours or days before sampling.

Field observations in *Lobochilotes labiatus*

The feeding behaviour of *L. labiatus* was observed at our field site near Mpulungu, Zambia, in concrete ponds

($1.5 \times 1.5 \times 1$ m). The purpose of these observations under semi-natural conditions and with wild specimens was to document if and how the lips are used in processing the main prey item identified in the stomach content analyses. The ponds were equipped with stones of ~20–30 cm diameters that covered the ground and formed caves as they occur naturally in the habitat of *L. labiatus*. Each pond was stocked with five to six freshly caught and unharmed adult individuals of *L. labiatus*. After an acclimatization period of at least 4 days, fish were offered snails of different sizes and their feeding behaviour was recorded with two underwater cameras (Canon Ixus 65 with WP-DC3 underwater case; Olympus μ tough-6000) for a period of 1 h each.

Comparative gene expression assays using RNAseq

For the identification of differentially expressed genes in thick-lipped species, we performed RNA sequencing (RNAseq) comparing lip tissue from a thick-lipped species to lip tissue from a reference species. We decided to perform these experiments in the African species *L. labiatus* and to use the closely related species *Astatotilapia burtoni* as reference taxon for several reasons such as the availability of laboratory strains and of sufficient RNA samples from adult and juvenile individuals. Most importantly, we chose this set-up because of the availability of various genomic resources for *A. burtoni*, such as a whole-genome sequence and a set of ~50 000 partly annotated expressed sequence tags (ESTs) (Salzburger *et al.* 2008; Baldo *et al.* 2011), which is crucial for the analysis and interpretation for RNAseq data. Such resources are currently not publicly available for *Amphilophus*.

In a first step, RNA was extracted from adult and juvenile individuals of *L. labiatus* and *A. burtoni* (see above for the RNA extraction protocol). RNA quality and quantity were determined on a NanoDrop 1000 spectrophotometer (Thermo Scientific) and by gel electrophoresis. RNA samples were pooled to create four samples subjected to RNA sequencing (RNAseq): (i) *A. burtoni* adult ($N = 3$); (ii) *A. burtoni* juvenile ($N = 1$); (iii) *L. labiatus* adult ($N = 2$); and (iv) *L. labiatus* juvenile ($N = 3$). Five micrograms of RNA per RNAseq sample was sent for Illumina sequencing at the Department of Biosystems Science and Engineering (D-BSSE), University of Basel and ETH Zurich. For library construction and sequencing, standard protocols were applied. Poly-A mRNA was selected using poly-T oligo-attached magnetic beads. The recovered mRNA was fragmented into smaller pieces using divalent cations under increased temperature. cDNA was produced using reverse transcriptase and random primers, followed by second-strand cDNA synthesis using DNA polymerase

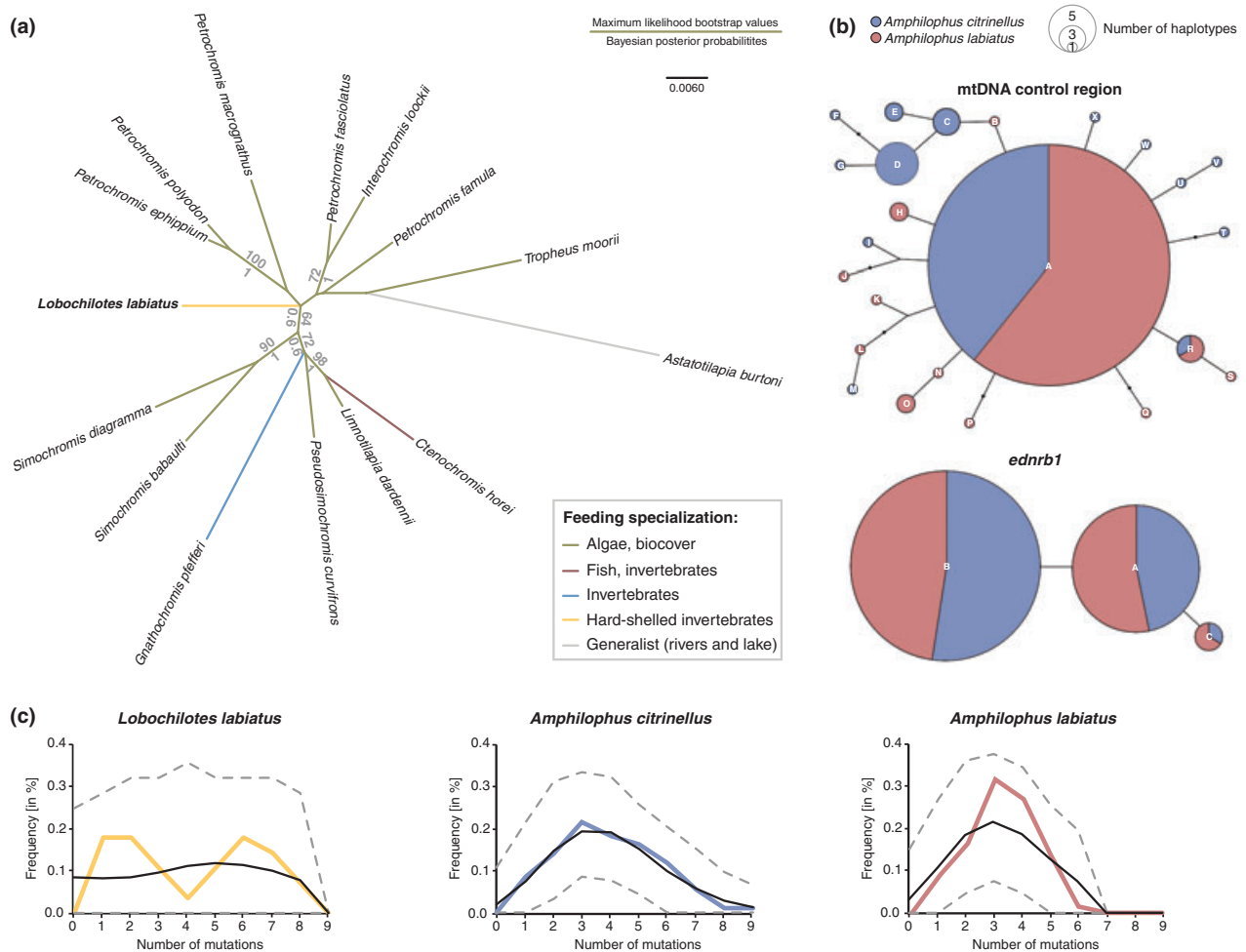


Fig. 2 Evolutionary origin of the thick-lipped species in East African Lake Tanganyika and in the Great Lakes of Nicaragua. (a) Maximum-likelihood tree of the Tropheini from Lake Tanganyika based on two mitochondrial (control region and ND2) and two nuclear (*ednrb1* and *phpt1*) gene segments (2345 bp in total) and the GTR+G+I model of molecular evolution. Numbers above the branches refer to maximum-likelihood bootstrap values, and numbers below are Bayesian posterior probabilities (note that support values are only shown for branches with bootstrap values >60). Branches are colour-coded according to feeding specializations; the trait values for internal branches have been reconstructed with MESQUITE. (b) Haplotype genealogies of the two *Amphilophus* species based on the mitochondrial control region and the nuclear *ednrb1* gene. A large fraction of the haplotypes is shared between *A. citrinellus* and *A. labiatus*. (c) Results from the mismatch analysis on the basis of the mitochondrial control region showing the inferred demographic histories for *L. labiatus*, *A. citrinellus* and *A. labiatus*. Coloured lines represent the observed data, the black line indicates the best-fit model, and the dashed lines in grey indicate the upper and lower boundaries from the simulations in ARLEQUIN.

I and RNaseH. cDNA went through an end-repair process, the addition of a single 'A' base and ligation of the adapters. It was then purified and enriched with PCR to create the final cDNA library. Each library was sequenced in one lane on an Illumina Genome Analyzer Ix (read length was 76 bp). Illumina reads are available from the Sequence Read Archive (SRA) at NCBI under the accession number SRA052992.

The Illumina reads were assembled into three different data sets for further analyses: (i) a quality-filtered data set (Data set 1), where the quality of the reads was assessed with the FASTX toolkit tools implemented in GALAXY [version September/October 2011; available at <http://main.g2.bx.psu.edu/> (Giardine *et al.* 2005; Blankenberg *et al.* 2010; Goecks *et al.* 2010)]; low-quality reads were discarded applying quality filter cut-off values of 22–33.

(ii) a quality-filtered plus trimmed data set (Data set 2), in which all the reads were trimmed to a length of 42 bp to evaluate the effects of read length (iii) as a control for the effect of trimming and filtering, a nonquality-filtered, nontrimmed data set (Data set 3).

The reads of the three data sets were then aligned to a reference cichlid assembly (Baldo *et al.* 2011) with NOVOALIGN 2.07.06 (<http://www.novocraft.com/>) after indexing the reference sequences with NOVOINDEX (<http://www.novocraft.com/>) using default parame-

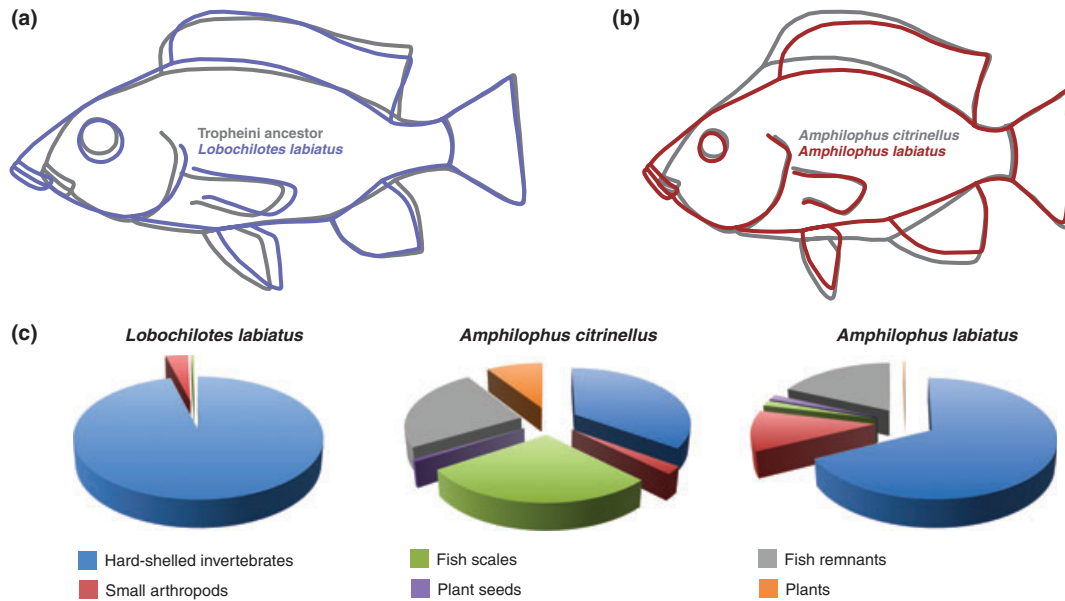


Fig. 3 Ecomorphology of the thick-lipped cichlid species in Central America and in Lake Tanganyika. (a) Body shape of *L. labiatus* in comparison with a reconstruction of the ancestor of *L. labiatus* and nine closely related Tropheini species. (b) Differences in body shape between *A. citrinellus* and *A. labiatus* along a discriminate function. In both plots, changes in landmark positions were increased threefold and interpolated outlines added for illustration purposes. Landmark locations are indicated in black on the reconstructed outlines in plot (a). (c) Analysis of stomach and gut content in the focal species. The fraction of each food category is shown.

ters. The alignment was performed using default settings with a maximum alignment score (t) of 180 and a maximum number of alignments for a single read (e) of 100; reads with multiple alignment locations were discarded. Next SAMTOOLS version 0.1.18 (Li *et al.* 2009) was used to sort and index the files and to generate count files, which were subsequently transformed into count tables and analysed in the R package DESEQ version 1.0.5 (Anders & Huber 2010). Differentially expressed genes between the four experimental groups were detected using a model based on a negative binomial distribution implemented in DESEQ. Differentially expressed genes with *P*-values (adjusted for multiple testing) >0.05 and/or a quotient of variance >1.00 were discarded to reduce the number of false positives. The remaining differentially expressed genes of all pairwise comparisons were tested for multiple hits. Next the hits of the three data sets were compared with each other to create a candidate gene list, consisting of genes that were found in multiple analyses in all three data sets. Lastly, these hits were compared to the annotated *A. burtoni* ESTs of Baldo *et al.* (2011).

Comparative gene expression assays using quantitative real-time PCR

Based on their function according to gene ontology terms (GO terms; <http://www.geneontology.org/>) and their putative involvement in lip formation and/or

hypertrophy in other organisms, six candidate genes were selected out of the list of differentially expressed genes for further characterization by means of quantitative real-time PCR (qPCR). These candidate genes are the *Bcl2 adenovirus e1b 19-kda protein-interacting protein 3* (*BNIP3*), *long-chain-fatty-acid(CoA)-ligase 4* (*ACSL4*), *histone 3.3* (*His3*), *beta actin* (*Actb*), *coatamer subunit beta* (*Copb*) and *claudin 7* (*Cldn7*; see Table 1 for primer details). qPCR experiments were performed in total of 36 cichlid specimens: *L. labiatus* (six adults, six juveniles), *A. burtoni* (six adults, six juveniles), *A. labiatus* (six adults) and *A. citrinellus* (six adults). By performing two pairwise comparisons between a thick-lipped and a normal-lipped species (a species pair each from Africa and Nicaragua), we effectively control for species-specific expression differences, as genes specific to thick-lip tissue should be upregulated in both comparisons.

The experiments were conducted on a StepOnePlus Real-Time PCR system (Applied Biosystems) as described elsewhere (Diepeveen & Salzburger 2011) using the *elongation factor 1* (*EF1*) and the *ribosomal protein SA3* (*RpSA3*) as endogenous controls. Average relative quantifications (RQ) were calculated for the six experimental groups and subsequently analysed with a two-tailed unpaired t-test using GRAPHPAD PRISM version 5.0a for Mac OS X (www.graphpad.com). We compared the expression levels between the two thick-lipped species and a closely related normally lipped species (i.e. *L. labiatus* vs. *A. burtoni* and *A. labiatus* vs. *A. citrinellus*). We also compared adults vs.

Locus	Forward (5'–3')	Reverse (5'–3')
<i>Actb</i>	CAGGCATCAGGGTGTAAATGGTT	CAGGCATCAGGGTGTAAATGGTT
<i>Copb</i>	GAGGCTACCTGGCTGTCAAAG	GTGCTGGATGGTTTGAGGGTAA
<i>His3</i>	CATCTACTGGTGGAGTGAAGAAACC	GGATCTCACGCAGAGCAACA
<i>ACSL4</i>	TGGTTCTGCACCGGAGATG	TCTTGGCGTCAACAATTTGTAGA
<i>BNIP3</i>	AACAGTCCACAAAGGAGTTCCT	CCTGATGCTGAGAGAGGTTGTG
<i>Cldn7</i>	GACATCATCCGGCCTTCT	CACCGAACTCATACTAGTGTGACA
<i>EF1</i>	GCCCTGCAGGACGTCTA	CGCCGACGGGTACAGT
<i>RpSA3</i>	AGACCAATGACCTGAAGGAAGTG	TCTCGATGCCTTGCCAACA

Table 1 Primers used for the quantitative real-time PCR experiments

juveniles in the African species, as hypertrophy in lips is much less pronounced at juvenile stages, so that this experiment also captures ontogenetic changes in lip formation. As primer efficiency was lower in the Nicaraguan samples, no direct comparisons between African and Nicaraguan tissues were possible.

Results

Phylogenetic and demographic analyses

Our phylogenetic analysis of members of the Tanganyikan cichlid tribe Tropheini based on two mitochondrial and two nuclear DNA gene segments reveals only limited phylogenetic resolution between the main lineages of the tribe (Fig. 2a). This confirms an earlier analysis based on mitochondrial DNA only, which attributed the star-like phylogeny of the Tropheini to the rapidity of lineage formation in the early stages of the adaptive radiation of this clade (Sturmbauer *et al.* 2003). Just as in the previous study, the thick-lipped species *L. labiatus* represents a separate lineage (without a closely related sister-taxon) that branches off relatively early in the phylogeny, but shows affinities to the algae-eating genera *Petrochromis* and *Simochromis*.

The haplotype genealogies of the *Amphilophus* samples based on the mitochondrial control region and the nuclear *ednrb1* gene (Fig. 2b) revealed haplotype sharing between *A. citrinellus* and *A. labiatus* (see also Barluenga & Meyer 2010). While all *Amphilophus* sequences were identical in *phpt1*, we detected three shared haplotypes in *ednrb1* and 24 haplotypes in the mitochondrial control region (two shared, ten unique to *A. labiatus* and twelve unique to *A. citrinellus*).

The mismatch analyses based on the mitochondrial control region sequences revealed unimodal distributions for the two sympatrically occurring *Amphilophus* species and a bimodal distribution for *L. labiatus* (Fig. 2c). According to this analysis, the demographic expansion of the two *Amphilophus* species happened at similar times, with the one of *A. citrinellus* being slightly older than that of *A. labiatus* (mean number of differences: 3.9 vs. 3.2; τ : 3.9 vs. 3.5; see also Barluenga &

Meyer 2010, who provide a relative time frame for the evolution of the Midas Cichlid species complex); the mean number of differences in *L. labiatus* was 6.4 (τ : 6.5).

Geometric morphometric analyses

The PCA of overall body shape revealed substantial overlap between the two Nicaraguan species *A. citrinellus* and *A. labiatus* (Appendix S3). The African thick-lipped species *L. labiatus* is separated from these mainly by principal component 1 (accounting for 20.2% of the variance), whereas principal component 2 (covering 16.0% of the variance) did not discriminate much between species. The discriminant function analysis, in which we compared species in a pairwise manner, revealed the main morphological differences between species. Of the two Nicaraguan species, *A. labiatus* had a more acute head, less deep body and a larger mouth than *A. citrinellus* (Fig. 3) (see also Klingenberg *et al.* 2003). These characters were even more pronounced in *L. labiatus*, when compared to either of the *Amphilophus* species. However, the distance in morphospace between the two species with fleshy lips was somewhat smaller than between *A. citrinellus* and *L. labiatus* (procrustes distance 0.08 and 0.1, respectively). We also estimated the body shape of the ancestor of *L. labiatus* and the 9 most closely related Tropheini species. A comparison of this reconstructed shape and the mean shape of our *L. labiatus* samples highlighted similar morphological differences as the comparison of the Nicaraguan species (Fig. 3), especially in the mouth region.

Stomach and gut content analyses

The fractions of food categories in guts and stomachs differed between *A. citrinellus*, *A. labiatus* and *L. labiatus* (Fig. 3c). While the diet of *A. citrinellus* did not overlap with that of *A. labiatus* (Schoener's index: 0.58) or *L. labiatus* (Schoener's index: 0.38), we found significant overlap between the two thick-lipped species *A. labiatus* and *L. labiatus* (Schoener's index: 0.71) (note that any value >0.6 is considered 'biologically significant'; see Wallace 1981). The stomach and gut contents of both

thick-lipped species consisted of a substantial fraction of hard-shelled prey (*Lobochilotes labiatus* 96%, *Amphilophus labiatus* 67.6%, *Amphilophus citrinellus* 35%).

Field observations in *Lobochilotes labiatus*

A careful inspection of the video material confirmed the findings from the stomach and gut content analyses that *L. labiatus* regularly feeds on snails (more than 90% of the stomach and gut content of *L. labiatus* consisted of snail shells). Small snails were engulfed using suction feeding without the lips touching the prey item or the surface (rocks) on which the items were placed. When feeding on larger snails, however, *L. labiatus* exhibited a different feeding strategy and snails were no longer taken up using suction feeding. Instead, *L. labiatus* used their lips to snatch the snails and they turned the snails a few times before they either swallowed the snails or spat them out (see Appendix S4).

Comparative gene expression assays using RNAseq

On average, ca. 42 million total reads were retrieved for each of the four RNAseq samples (*A. burtoni* adult, *A. burtoni* juvenile, *L. labiatus* adult and *L. labiatus* juvenile). Quality filtering and trimming reduced this number so that on average 21.9 (Data set 1), 24.6 (Data set 2) and 23.5 (Data set 3) million reads were aligned to the reference cichlid assembly. Five different pairwise comparisons were made to obtain genes that are differentially expressed between thick lips and normal lips (see Table 2 for the three comparisons with the highest number of genes being different). The largest number of differentially expressed genes between *L. labiatus* and *A. burtoni* was detected in adult lip tissue, with the majority of the genes being upregulated in *L. labiatus*. The total number of differentially expressed genes ranged from 9050 (Data set 3; three pairwise comparisons) to 15230 (Data set 2; five pairwise comparisons). A substantial fraction of these differentially expressed genes appeared in at least two comparisons in each data set (Data set 1: 2085 [22.1% of all hits]; Data set 2: 8078 [53.0%]; Data set 3: 1693 [18.7%]). Of these 'multiple

hits', 1463 were detected in all three data sets and 560 of those could be unequivocally annotated.

A more stringent analysis, in which only loci that appeared in at least three of five comparisons were included, resulted in 231 differentially expressed genes. A functional annotation of these 231 hits with Blast2GO resulted in a total of 141 annotations (122 upregulated and 19 downregulated in *L. labiatus*; see Appendix S3). Based on their annotations, known functions and/or exceptional fold change (>1000) between *A. burtoni* and *L. labiatus*, thirteen genes were identified as good candidates for being involved in the morphogenesis of fleshy lips (Table 3).

Comparative gene expression assays using quantitative real-time PCR

The results of the comparative gene expression assays between the thick-lipped species and the normal-lipped species are depicted in Fig. 4 and Appendix S5. Overall, the qPCR experiments largely validate differential gene expression in normal and hypertrophied lip tissue as indicated by RNAseq. In the African species pair *L. labiatus* and *A. burtoni*, which were the two species used for RNAseq, differences were highly significant in four of the six genes tested: *Actb* ($P = 0.0099$), *Cldn7* ($P = 0.004$), *ACSL4* ($P = 0.0005$) and *His3* ($P = 0.0003$). However, we would like to point out one inconsistency between RNAseq and qPCR. *Actb* was actually found to be downregulated in hypertrophied lips by RNAseq, while it shows significantly higher expression levels in lip tissue in the qPCR experiments (Fig. 4).

The comparison between lip tissue in adult and juvenile *L. labiatus* and *A. burtoni* further revealed a trend towards higher expression in lip tissue of adult *L. labiatus* in *Actb*, *BNIP3*, *Cldn7* and *Copb* (Appendix S5), whereas, generally, an opposite trend is observed in *A. burtoni*, although statistical support was only found in two cases [*Cldn7* ($P = 0.0063$) and *ACSL4* ($P = 0.0328$)]. This again suggests that these genes are involved in the formation of fleshy lips. In the Nicaraguan species pair, a similar trend was observed as in the African species pair, with four of the five genes tested appearing to be upregulated in lip tissue

Comparison	Data set 1	Data set 2	Data set 3
AB vs. LL	7120 (4606; 2514)	7080 (4689; 2391)	7285 (4665; 2620)
AB vs. LLjuv	3611 (3395; 216)	13747 (10683; 3064)	2618 (2514; 104)
ABjuv vs. LLjuv	1116 (792; 324)	3971 (2710; 1261)	986 (687; 298)
Total	9407	15225	9050

Table 2 Pairwise comparisons of differentially expressed genes and total number of unique differentially expressed genes in the three data sets compiled in this study

AB, *Astatotilapia burtoni*; LL, *Lobochilotes labiatus*; juv, juvenile; numbers in brackets denote the number of upregulated and downregulated genes in *L. labiatus*.

Table 3 Thirteen candidate loci for the genetic basis of lip development in the East African cichlid *Lobochilotes labiatus*, based on RNAseq and qPCR in comparison with *Astatotilapia burtoni*, in combination with information on gene functions (in alphabetical order)

Locus	Abbreviation
<i>ATPase mitochondrial precursor</i>	<i>ATPmp</i>
<i>Bcl2 adenovirus e1b 19-kda protein-interacting protein 3</i>	<i>BNIP3</i>
<i>Beta actin</i>	<i>Actb</i>
<i>Caspase-8</i>	<i>Casp8</i>
<i>Claudin 7</i>	<i>Cldn7</i>
<i>Coatmer subunit beta</i>	<i>Copb</i>
<i>Grainyhead-like protein 1 homolog</i>	<i>Grhl1</i>
<i>Heat-shock 70-kda protein 12a-like</i>	<i>Hspa12al</i>
<i>Histone 3.3</i>	<i>His3</i>
<i>Laminin subunit gamma-2</i>	<i>Lamc2</i>
<i>Long-chain-fatty-acid(CoA)-ligase 4</i>	<i>ACSL4</i>
<i>Sodium-dependent phosphate transporter 1</i>	<i>Slc17a1</i>
<i>Transcription factor ap-2 gamma</i>	<i>Tfap2</i>

of *A. labiatus* as compared to *A. citrinellus* (Fig. 4; we could not amplify *BNIP3* here). We would like to note, however, that qPCR efficiency was less good in the *Amphilophus* samples, most likely because we used primers designed for the African species pair based on the

available genomic resources, which also explains the limited statistical support for these comparisons. Interestingly, it seems that several loci (i.e. *Actb*, *Cldn7*, *Copb*, *His3*) are upregulated in both thick-lipped species when compared to their normally lipped relatives.

Discussion

The species flocks of cichlid fishes in the East African Great Lakes Victoria, Malawi and Tanganyika, counting hundreds of endemic species each, are prime examples of adaptive radiation and explosive speciation (see e.g. Kocher 2004; Seehausen 2006; Salzburger 2009). Interestingly, the cichlid adaptive radiations in East Africa have independently produced ecomorphs with highly similar colour patterns and (mouth) morphologies (Kocher *et al.* 1993). Here, we explore the ecological and genetic basis of one of the particular trophic structures of cichlids, which has evolved convergently in various cichlid assemblages: fleshy lips. Instead of focusing on species with hypertrophied lips between the radiations in the East African lakes, we compare the thick-lipped phenotype between a cichlid assemblage in East African (Lake Tanganyika) and in Central American (the lake Nicaragua/Managua system), where thick-lipped species have evolved in parallel (see Fig. 1).

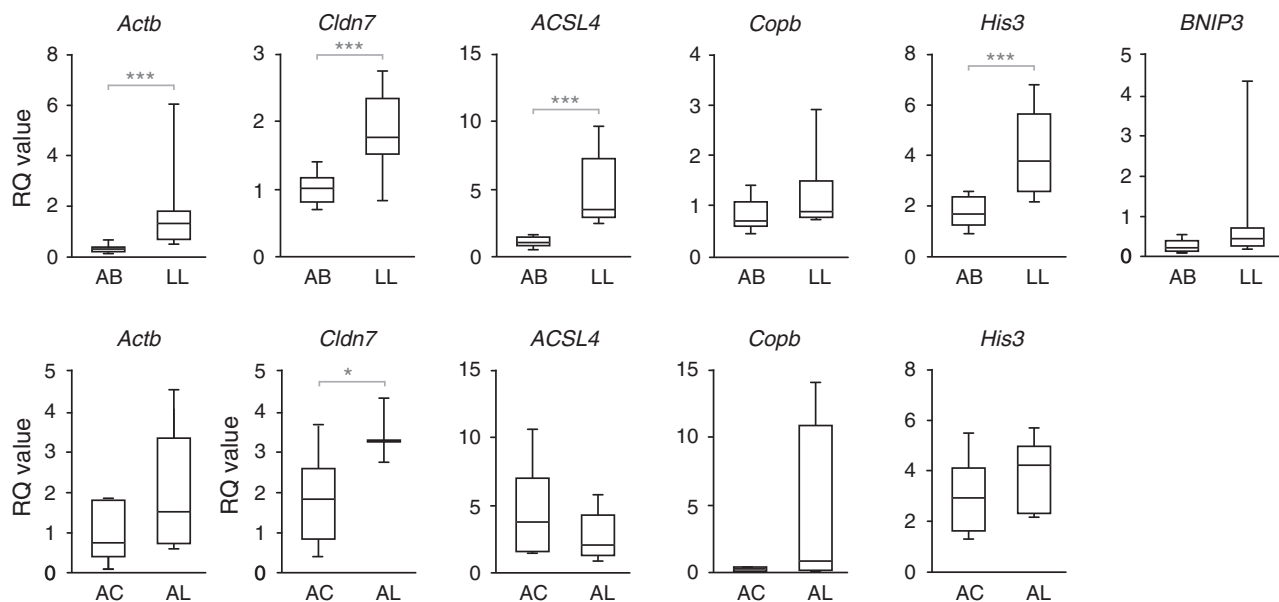


Fig. 4 Results from the comparative gene expression experiments via quantitative real-time PCR. The six genes tested in this experiment were selected on the basis of comparative RNA sequencing. All genes tested show a higher expression level in lip tissue of the Tanganyikan thick-lipped species *L. labiatus* as compared to *A. burtoni* (top panel; note that we used both juvenile and adult samples in these analyses to increase statistical power). A similar trend was found when comparing the Nicaraguan thick-lipped species *A. labiatus* to its sister species *A. citrinellus* (with the exception of *ACSL4*; lower panel). Note that *BNIP3* could not be amplified in the *Amphilophus* species. *Astatotilapia burtoni* (AB); *Lobochilotes labiatus* (LL); *Amphilophus citrinellus* (AC); *Amphilophus labiatus* (AL); * $P < 0.05$; *** $P < 0.01$.

The evolution of hypertrophied lips in cichlid adaptive radiations

Our phylogenetic and demographic analyses in the Tanganyikan Tropheini and the Nicaragua Midas Cichlid species complex reveal that the thick-lipped species are nested within their respective clade. The molecular phylogeny of 14 Tropheini species (Fig. 2a) shows a footprint characteristic for adaptive radiations: a 'bottom heavy' topology with only limited phylogenetic resolution at the deeper nodes due to rapid lineage formation (Gavrilets & Vose 2005). Our new analysis thus confirms previous results based on mtDNA only (Sturmbauer *et al.* 2003) or a combination of mtDNA and AFLPs (Koblmüller *et al.* 2010). In all analyses thus far, the thick-lipped species *L. labiatus* forms an independent evolutionary lineage that branches off deep in the Tropheini. Its exact position remains unclear, though. In the AFLP phylogeny of Koblmüller *et al.* (2010), *L. labiatus* appears as sister group to all Tropheini except for the genus *Tropheus*, which is sister to all other representatives of that clade (the topology has very little support, though). In our new phylogeny and the previous mtDNA trees of Sturmbauer *et al.* (2003), *L. labiatus* shows affinities to *Simochromis* and *Petrochromis* (with moderate support). In all phylogenies, however, *L. labiatus* is nested within a clade formed by various species that feed on algae and biocover (see our character state reconstruction in Fig. 2a).

In the Midas Cichlid species complex from Central America, a phylogenetic approach is not applicable with the available molecular markers. There is simply too little genetic variation, even in the rapidly evolving mitochondrial control region, as a consequence of the young age of the assemblage (see Barluenga & Meyer 2004, 2010; Barluenga *et al.* 2006). The structures of our haplotype genealogies, which now also include the analysis of a nuclear gene (Fig. 2b), confirm this scenario. In combination with the mismatch analyses (Fig. 2c), these data suggest that *A. labiatus* underwent its main demographic expansion soon after the expansion of the sympatric *A. citrinellus* populations (see Barluenga & Meyer 2010 for a large-scale analysis of the Midas Cichlid species complex).

In both species assemblages, the evolution of the thick-lipped phenotype was associated with similar modifications of overall body shape (Fig. 3a,b). Reduced body depth, a more acute head shape and a larger mouth, along with the prominently enlarged lips, can be hypothesized to be adaptations to the species' microhabitat and trophic niche. If individuals search for food in narrow rock crevices, these modifications appear advantageous. Klingenberg *et al.* (2003) already sug-

gested that the elongation of the head, as observed in both '*labiatus*' species, increases suction power. Other morphological differences between the two thick-lipped species, such as eye size or the length of anal fin insertion, might be either due to adaptations to the specific environments or due to phylogenetic effects. Inclusion of other thick-lipped species in future studies focusing on the ecology and morphological evolution of this trait might answer this question.

The function of hypertrophied lips in cichlids

Hypertrophied lips in cichlids have been implicated in several functions. For example, it has been suggested that fleshy lips are used to seal cracks and grooves to facilitate sucking of invertebrates (Barlow & Munsey 1976; Ribbink *et al.* 1983; Seehausen 1996; Konings 1998), that they act as bumpers to protect from mechanical shock (Greenwood 1974; Yamaoka 1997) or that they function as taste (Arnegard *et al.* 2001) or mechanoreceptors (Fryer 1959; Fryer & Iles 1972). Previous food web analyses on *L. labiatus* identified this species as mollusc eater (Nori 1997).

Our ecomorphological analysis of the thick-lipped species *L. labiatus* from Lake Tanganyika and *A. labiatus* from the large lakes in Nicaragua suggests that this phenotype is indeed associated with feeding on hard-shelled prey such as snails, mussels and crustaceans in rocky habitats (Fig. 3c). We cannot, however, conclusively answer the question whether the lips are used to seal rock crevices or whether they serve as bumpers or receptors. In the underwater observations at our field site at Lake Tanganyika, small snails were usually engulfed by *L. labiatus* via suction feeding, whereas larger snails were turned around several times before being swallowed or spit out (see Appendix S4). This would classify the lips as instrument to handle hard-shelled invertebrate food (mostly molluscs). Note, however, that our observations were made in semi-natural conditions only, in the form of concrete ponds equipped with stones from the lake and filled with lake water.

Our experimental set-up could not address the possibility that phenotypic plasticity plays a role in the formation of fleshy lips, as has previously been shown in certain foraging traits in cichlid fishes (oral jaws: Meyer 1987; pharyngeal jaws: e.g. Greenwood 1965; Huisseune 1995; Muschick *et al.* 2011). Interestingly, it has been reported that thick-lipped cichlid species lose their fleshy lips under unnatural conditions in captivity (when fed with standard food; Barlow & Munsey 1976; Barlow 1976; Loiselle 1998). So far, there is no evidence for the opposite process, the plastic development of fleshy lips due to environmental or feeding properties. In the common garden experiment of Muschick *et al.*

(2011), one group of normally lipped *A. citrinellus* individuals was fed with whole snails over a period of several months, and—although not formally assessed—no increase in lip size was apparent (compared to the other two treatment groups peeled snails and crushed snails). Another study on a snail crusher (Huysseune 1995) did not report such changes either, which seems to suggest that phenotypic plasticity in the lips, if at all present, is specific to thick-lipped species only. Future common garden and feeding experiments should thus expand on this question. Such experiments, combined with molecular analyses, should focus on the plastic component of this trait and its genomic basis.

Insights into the genetic basis of hypertrophied lips in cichlids

Our comparative gene expression assays with RNA sequencing between tissue from thick and normal lips identified a set of 141 candidate genes that might be responsible for the morphogenesis or the maintenance of fleshy lips in (East African) cichlid fish (Appendix S3). Six genes were tested further by means of quantitative real-time PCR, and these experiments largely confirm the results obtained from RNAseq (Fig. 4). While there is no obvious functional connection to fleshy lips for three of these differentially expressed genes (*ACSL4*, *His3* and *BNIP3*), the observed upregulation of the remaining three (*Actb*, *Cldn7* and *Copb*) makes sense in the light of the structure of hypertrophied lips. These three genes (together with *BNIP3*) also show a higher expression in lip tissue from adult vs. juvenile *L. labiatus* (Appendix S5).

It has previously been shown that the 'fleshy' lips of the Lake Malawi cichlid *Otopharynx pachycheilus* mainly consist of loose connective tissue covered by dermis and a layer of epithelial cells (Arnegard *et al.* 2001). Interestingly, the known functions of *Actb*, *Cldn7* and *Copb* can be directly implicated in cell and/or intercell or membrane structure. The cytoplasmic *Actb* is found in high abundance in nonmuscle cells, where it promotes cell surface and cell thickness (Schevzov *et al.* 1992), which is also consistent with its upregulation in the more massive adult compared to juvenile *L. labiatus* lips (Appendix S5). The integral membrane protein *Cldn7* (among other *claudin* gene family members) constitutes the backbone of tight junctions between epithelial cells (Tsukita *et al.* 2001). The coatomer coat proteins (such as *Copb*) are involved in protein and membrane trafficking via vesicle secreting between the endoplasmic reticulum and the Golgi apparatus, plus the intra-Golgi transport (Duden 2003). In addition, they mediate lipid homeostasis and lipid storage for energy use and membrane assembly (Soni *et al.* 2009). *Copb*

might thus be involved in cellular (membrane) development but possibly also in the formation of fat cells that compose adipose tissue, a specific subtype of connective tissue. Clearly, much more work will be necessary to unravel the development and genetic basis of hypertrophied lips in cichlids, for which we herewith established a valuable starting ground.

Our results, especially the comparison of gene expression levels between the thick-lipped species in East Africa and Central America (Fig. 4), allow us to touch on ongoing discussions related to the genetic basis of convergent morphologies (reviewed in Brakefield 2006; Arendt & Reznick 2008; Elmer & Meyer 2011). Although our qPCR results in Midas Cichlid (*Amphilophus* spp.) species must be taken with caution (efficiency was lower as a consequence of using molecular tools developed for the African species leading to a lack of statistical power), we find rather similar trends in gene expression. Our results seem to indicate that a largely overlapping set of genes was recruited to develop the hypertrophied lips in Nicaraguan and African species, which are—according to most authors—separated by ~ 100 million years of evolution. This important question about the basis of convergent phenotypes should be addressed in future studies, and thick-lipped fish species, including those outside the family Cichlidae, appear as an excellent model system.

Conclusion

Our integrative evolutionary, ecological, morphological, observational and genomic analysis of thick-lipped species in East Africa and in Nicaragua reveals stunning similarities between these convergent morphs. Both thick-lipped species appear to have evolved early in the respective clade, they seem to have adapted to the same habitat (rocks) and food source (hard-shelled prey), and their evolution was associated with comparable morphological trajectories, especially in the mouth and head region. Importantly, we also show that the expression patterns of at least some genes are similar, too. We thus provide valuable resource for future studies focusing on the development of this trait and genetic basis of convergence.

Acknowledgements

We would like to thank our helpers in the field, V. Campos, C. Heule, B. Meyer, M. Roesti; E. P. van den Berghe and G. Tembo and his crew for their logistic support in Nicaragua and Africa, respectively; the Ministerio del Ambiente y los Recursos Naturales Nicaragua (MARENA) and the Lake Tanganyika Research Unit, Department of Fisheries, Republic of Zambia,

for research permits; and three anonymous referees and the Subject Editor, C. Eizaguirre, for valuable comments. This study was supported by grants from the Fundacao para Ciencia e a Tecnologia (FCT, Portugal) to M. E. S., the Swiss Academy of Sciences to A. I., the Spanish Ministerio de Economía y Competitividad to M. B., and the European Research Council (ERC, Starting Grant 'INTERGENADAPT'), the University of Basel and the Swiss National Science Foundation (SNF, grants 3100A0_122458 and CRSII3_136293) to W. S.

References

- Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biology*, **11**, R106.
- Arendt J, Reznick D (2008) Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in Ecology and Evolution*, **23**, 26–32.
- Arnegard ME, Snoeks J, Schaefer SA (2001) New three-spotted cichlid species with hypertrophied lips (Teleostei: Cichlidae) from the deep waters of Lake Malaŵi/Nyasa, Africa. *Copeia*, **2001**, 705–717.
- Baldo L, Santos ME, Salzburger W (2011) Comparative transcriptomics of Eastern African cichlid fishes shows signs of positive selection and a large contribution of untranslated regions to genetic diversity. *Genome Biology and Evolution*, **3**, 443–455.
- Barlow GW (1976) The Midas cichlid in Nicaragua. In: *Investigations of the Ichthyofauna of Nicaraguan lakes* (ed. Thorson TB), pp. 333–358. School of Life Sciences, University of Nebraska, Lincoln, Nebraska.
- Barlow GW, Munsey JW (1976) The Red Devil Midas Cichlid species complex in Nicaragua. In: *Investigations of the Ichthyofauna of Nicaraguan Lakes* (ed. Thorson TB), pp. 359–370. School of Life Sciences, University of Nebraska, Lincoln, Nebraska.
- Barluenga M, Meyer A (2004) The Midas cichlid species complex: incipient sympatric speciation in Nicaraguan cichlid fishes? *Molecular Ecology*, **13**, 2061–2076.
- 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, Stoltz KN, Salzburger W, Muschick M, Meyer A (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, **439**, 719–723.
- Berner D, Roesti M, Hendry AP, Salzburger W (2010) Constraints on speciation suggested by comparing lake-stream stickleback divergence across two continents. *Molecular Ecology*, **19**, 4963–4978.
- Blankenberg D, Von Kuster G, Coraor N (2010) Galaxy: a web-based genome analysis tool for experimentalists. *Current protocols in molecular biology*/edited by Frederick M. Ausubel et al. Chapter 19, Unit 19 10 11–21.
- Brakefield PM (2006) Evo-devo and constraints on selection. *Trends in Ecology and Evolution*, **21**, 362–368.
- Brichard P (1989) *Cichlids and all Other Fishes of Lake Tanganyika*, T.H.F. Publications, Neptune City, New Jersey.
- Bruford MW, Hanotte O, Brookfield JFY, Burke T (1998) Multilocus and single-locus DNA fingerprinting. In: *Molecular Analysis of Populations* (ed. Hoelzel AR), pp. 283–336. Oxford University Press, New York.
- Butler MA, Sawyer SA, Losos JB (2007) Sexual dimorphism and adaptive radiation in Anolis lizards. *Nature*, **447**, 202–205.
- Diepveven ET, Salzburger W (2011) Molecular characterization of two endothelin pathways in East African cichlid fishes. *Journal of Molecular Evolution*, **73**, 355–368.
- Duden R (2003) ER-to-Golgi transport: COP I and COP II function (Review). *Molecular Membrane Biology*, **20**, 197–207.
- Elmer KR, Meyer A (2011) Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends in Ecology and Evolution*, **26**, 298–306.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.
- Fryer G (1959) The trophic interrelationships and ecology of some littoral communities of Lake Nyasa with a special reference to the fishes, and a discussion on the evolution of a group of rock-frequenting Cichlidae. *Proceedings of the Zoological Society of London*, **132**, 153–281.
- Fryer G, Iles TD (1972) *The Cichlid Fishes of the Great Lakes of Africa: Their Biology and Evolution*. Oliver & Boyd, Edinburgh.
- Gavrilets S, Losos JB (2009) Adaptive radiation: contrasting theory with data. *Science*, **323**, 732–737.
- Gavrilets S, Vose A (2005) Dynamic patterns of adaptive radiation. *Proceeding of the National Academy of Sciences U S A*, **102**, 18040–18045.
- Genner MJ, Seehausen O, Lunt DH et al. (2007) Age of cichlids: new dates for ancient lake fish radiations. *Molecular Biology and Evolution*, **24**, 1269–1282.
- Giardine B, Riemer C, Hardison RC et al. (2005) Galaxy: a platform for interactive large-scale genome analysis. *Genome Research*, **15**, 1451–1455.
- Goecks J, Nekrutenko A, Taylor J (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology*, **11**, R86.
- de Graaf M, Dejen E, Osse JWM, Sibbing FA (2008) Adaptive radiation of Lake Tana's (Ethiopia) *Labeobarbus* species flock (Pisces, Cyprinidae). *Marine and Freshwater Research*, **59**, 391–407.
- Grant PR, Grant BR (2008) *How and Why Species Multiply: The Radiations of Darwin's Finches*. Princeton University Press, Princeton, USA.
- Greenwood PH (1965) Environmental effects on the pharyngeal mill of a cichlid fish, *Astatoreochromis alluaudi*, and their taxonomic implications. *Proceedings of the Linnean Society London*, **176**, 1–10.
- Greenwood PH (1974) The cichlid fishes of Lake Victoria East Africa: the biology and evolution of a species flock. *Bulletin of the British Museum for Natural History (Zool.) Supplementary*, **6**, 1–134.
- Harmon LJ, Kolbe JJ, Cheverud JM, Losos JB (2005) Convergence and the multidimensional niche. *Evolution*, **59**, 409–421.
- Herder F, Schwarzer J, Pfaender J, Hadiaty RK, Schlieven UK (2006) Preliminary checklist of sailfin silversides (Teleostei: Telmatherinidae) in the Malili Lakes of Sulawesi (Indonesia), with a synopsis of systematics and threats. *Verhandlungen der Gesellschaft für Ichthyologie*, **5**, 139–163.
- Huyseune A (1995) Phenotypic plasticity in the lower pharyngeal jaw dentition of *Astatoreochromis alluaudi* (Teleostei: Cichlidae). *Archives in Oral Biology*, **40**, 1005–1014.

- Johnson MA, Revell LJ, Losos JB (2009) Behavioral convergence and adaptive radiation: effects of habitat use on territorial behavior in Anolis lizards. *Evolution*, **64**, 1151–1159.
- Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, **9**, 286–298.
- Klingenberg CP (2011) MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, **11**, 353–357.
- Klingenberg CP, Barluenga M, Meyer A (2003) Body shape variation in cichlid fishes of the *Amphilophus citrinellus* species complex. *Biological Journal of the Linnean Society*, **80**, 397–408.
- Koblmüller S, Egger B, Sturmbauer C, Sevc KM (2010) Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. *Molecular Phylogenetics and Evolution*, **55**, 318–334.
- Kocher TD (2004) Adaptive evolution and explosive speciation: the cichlid fish model. *Nature Reviews Genetics*, **5**, 288–298.
- Kocher TD, Conroy JA, McKaye KR, Stauffer JR (1993) Similar morphologies of cichlid fish in lakes Tanganyika and Malawi are due to convergence. *Molecular Phylogenetics and Evolution*, **2**, 158–165.
- Konings A (1998) Tanganyikan Cichlids in their Natural Habitat. Cichlid Press, El Paso.
- Lang M, Miyake T, Braasch I *et al.* (2006) A BAC library of the East African haplochromine cichlid fish *Astatotilapia burtoni*. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, **306B**, 35–44.
- Lee WJ, Conroy J, Howell WH, Kocher TD (1995) Structure and evolution of teleost mitochondrial control regions. *Journal of Molecular Evolution*, **41**, 54–66.
- Li H, Handsaker B, Wysoker A *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
- Loiselle PV (1998) The *Amphilophus labiatus* Species Complex. The Cichlid Room Companion. Retrieved on June 05, 2012, from: <http://www.cichlidae.com/article.php?id=106>.
- Losos JB (2009) Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles. University of California Press, Berkeley.
- Losos JB (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. American Society of Naturalists E. O. Wilson award address. *American Naturalist*, **175**, 623–639.
- Losos JB (2011) Convergence, adaptation and constraint. *Evolution*, **65**, 1827–1840.
- Losos JB, Ricklefs RE (2009) Adaptation and diversification on islands. *Nature*, **457**, 830–836.
- Losos JB, Jackmann TR, Larson A, De Queiroz K, Rodrigues-Schettino L (1998) Contingency and determinism in replicated adaptive radiations of island lizards. *Science*, **279**, 2115–2118.
- Matschiner M, Hanel R, Salzburger W (2011) On the origin and trigger of the notothenioid adaptive radiation. *PLoS One*, **6**, e18911.
- Meyer A (1987) Phenotypic plasticity and heterochrony in *Cichlasoma managuense* (Pisces, Cichlidae) and their implications for speciation in cichlid fishes. *Evolution*, **41**, 1357–1369.
- Meyer A, Morrissey JM, Scharl M (1994) Recurrent origin of a sexually selected trait in *Xiphophorus* fishes inferred from a molecular phylogeny. *Nature*, **368**, 539–542.
- Muschick M, Barluenga M, Salzburger W, Meyer A (2011) Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. *BMC Evolutionary Biology*, **11**, 116.
- Nori M (1997) Structure of littoral fish communities. In: *Fish Communities in Lake Tanganyika* (eds Kawanabe H, Hori M, Nagoshi M), pp. 277–298. Kyoto University Press, Kyoto.
- Nosil P, Crespi BJ, Sandoval DP (2002) Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature*, **417**, 440–443.
- Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.
- Ribbink AJ, Marsch BA, Marsch AC, Ribbink AC, Sharp BJ (1983) A preliminary survey of the cichlid fishes of rocky habitats in Lake Malawi. *South African Journal of Zoology*, **18**, 149–310.
- 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.
- Rohlf FJ (2006) DIG, Version 2.10.0. Department of Ecology and Evolution. State University of New York, Stony Brook, NY.
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.
- Rundle HD, Nagel L, Wenrick Boughman J, Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science*, **287**, 306–308.
- Salzburger W (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Molecular Ecology*, **18**, 169–185.
- Salzburger W, Meyer A (2004) The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. *Naturwissenschaften*, **91**, 277–290.
- Salzburger W, Meyer A, Baric S, Verheyen E, Sturmbauer C (2002) Phylogeny of the Lake Tanganyika cichlid species flock and its relationship to the Central and East African haplochromine cichlid fish faunas. *Systematic Biology*, **51**, 113–135.
- Salzburger W, Mack T, Verheyen E, Meyer A (2005) Out of Tanganyika: Genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evolutionary Biology*, **5**, 17.
- Salzburger W, Renn SC, Steinke D, Braasch I, Hofmann HA, Meyer A (2008) Annotation of expressed sequence tags for the East African cichlid fish *Astatotilapia burtoni* and evolutionary analyses of cichlid ORFs. *BMC Genomics*, **9**, 96.
- Salzburger W, Ewing GB, Von Haeseler A (2011) The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Molecular Ecology*, **20**, 1952–1963.
- Shevzov G, Lloyd C, Gunning P (1992) High level expression of transfected beta- and gamma-actin genes differentially impacts on myoblast cytoarchitecture. *The Journal of Cell Biology*, **117**, 775–785.
- Schluter D (2000) The Ecology of Adaptive Radiation. Oxford University Press, New York.
- Schluter D, Nagel LM (1995) Parallel speciation by natural selection. *American Naturalist*, **146**, 292–301.
- Schoener TW (1970) Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology*, **51**, 408–418.

- Seehausen O (1996) Lake Victoria Rock Cichlids, Verdujin Cichlids, Zevenhuizen, The Netherlands.
- Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. *Proceeding of the Royal Society London B*, **273**, 1987–1998.
- Sereno PC, Wilson JA, Conrad JL (2004) New dinosaurs link southern landmasses in the Mid-Cretaceous. *Proceeding of the Royal Society London B*, **271**, 1325–1330.
- Sibbing FA, Nagelkerke LAJ, Stet RJM, Osse JWM (1998) Speciation of endemic Lake Tana barbs (*Cyprinidae*, Ethiopia) driven by trophic resource partitioning: a molecular and ecomorphological approach. *Aquatic Ecology*, **32**, 217–227.
- Simpson GG (1953) *The Major Features of Evolution*, Columbia University Press, New York.
- Soni KG, Mardones GA, Sougrat R, Smirnova E, Jackson CL, Bonifacino JS (2009) Coatmer-dependent protein delivery to lipid droplets. *Journal of Cell Science*, **122**, 1834–1841.
- Stiassny MLJ, Meyer A (1999) Cichlids of the Rift Lakes. *Scientific American*, **280**, 64–69.
- Sturmbauer C, Hainz U, Baric S, Verheyen E, Salzburger W (2003) Evolution of the tribe Tropheini from Lake Tanganyika: synchronized explosive speciation producing multiple evolutionary parallelism. *Hydrobiologia*, **500**, 51–64.
- Swofford DL (2003) PAUP*—Phylogenetic Analyses Using Parsimony and Other Methods, Version 4.0, Sinauer, Sunderland, Massachusetts.
- Tsukita S, Furuse M, Itoh M (2001) Multifunctional strands in tight junctions. *Nature Reviews in Molecular Cell Biology*, **2**, 285–293.
- Verheyen E, Salzburger W, Snoeks J, Meyer A (2003) Origin of the superflock of cichlid fishes from Lake Victoria, East Africa. *Science*, **300**, 325–329.
- Villa J (1982) *Peces Nicaragüenses de Agua Dulce*, Fondo de Promoción Cultural, Banco de América, Managua.
- Wake DB, Wake MH, Specht CD (2011) Homoplasy: from detecting pattern to determining process and mechanism of evolution. *Science*, **331**, 1032–1035.
- Wallace RK (1981) An assessment of diet-overlap indexes. *Transactions of the American Fisheries Society*, **110**, 72–76.
- Yamaoka K (1997) Trophic ecomorphology of Tanganyikan cichlids. In: *Fish Communities in Lake Tanganyika* (eds Kawanabe H, Hori M, Nagoshi M), pp. 27–56. Kyoto University Press, Kyoto.
- Yoder JB, Clancey E, Des Roches S *et al.* (2010) Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, **23**, 1581–1596.

M.C., E.T.D., M.E.S. and A.I. are PhD students in the group of W.S. M.C. is interested in parallel evolution events as natural replicates to test hypotheses about trait evolution and the different (or similar) genetic bases that underlie these phenotypes. E.T.D. is interested in the genetic basis of adaptive traits and the selective forces acting upon these genes. M.E.S. is interested in the ecological and developmental mechanisms underlying the emergence and diversification of novel adaptive traits. A.I. is interested in ecomorphological adaptations, phylogeography and taxonomy in cichlid fishes. M.M. recently finished his PhD in the group of W.S. and is now postdoctoral fellow with Patrik Nosil in Sheffield. His research is concerned

with morphological and genomic evolution in adaptive radiations. N.B. is a technical assistant who is involved in several projects of the SalzburgerLab. M.B. is a group leader at the Natural History Museum in Madrid. Her research focuses on understanding incipient stages of speciation and the sequence of adaptations and specializations that organisms undergo after the colonization of new habitats. W.S. is Professor of Zoology and Evolutionary Biology at the University of Basel. The research of his team focuses on the genetic basis of adaptation, evolutionary innovation and animal diversification. The main model systems in the laboratory are threespine stickleback fish, Antarctic notothenioids and the exceptionally diverse assemblages of cichlid fishes. The laboratory's homepage at <http://www.evolution.unibas.ch/salzburger/> provides further details on the group's (research) activities.

Data accessibility

Newly generated DNA sequences for phylogenetic and haplotype analyses have been deposited in GenBank under accession numbers JX402217–JX402407 (see Appendix S1 for details). Illumina reads from the RNA-seq experiments are available from the Sequence Read Archive (SRA) at NCBI under the accession number SRA052992. Data from the stomach and gut content analyses, the MORPHOJ input files and the quantitative real-time PCR experiments have been deposited at Dryad (doi:10.5061/dryad.vf1ms).

Supporting information

Additional Supporting Information may be found in the online version of this article.

Appendix S1 List of specimens used in this study including sampling date and location and GenBank accession numbers.

Appendix S2 PCA of overall body shape of the African cichlid *Lobochilotes labiatus* and the Nicaraguan species *Amphilophus labiatus* and *A. citrinellus* (a) and distribution of landmarks for morphometric analyses (b).

Appendix S3 Blast2GO annotations of genes with differential expression between lip tissue from thick-lipped and normal-lipped cichlid species.

Appendix S4 Underwater video showing snail feeding in *Lobochilotes labiatus*.

Appendix S5 Results of the quantitative real-time PCR experiments comparing adult and juvenile lip tissue of the African cichlid species *Lobochilotes labiatus* and *Astatotilapia burtoni*.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.