Phylogenomics of pike cichlids (Cichlidae: Crenicichla): the rapid ecological speciation of an incipient species flock

E. D. BURRESS*, F. ALDA†, A. DUARTE‡, M. LOUREIRO‡§, J. W. ARMBRUSTER* & P. CHAKRABARTY†

*Department of Biological Sciences and Auburn University Museum of Natural History, Auburn University, Auburn, AL, USA
†Museum of Natural Science, Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, USA
‡Sección Zoología Vertebrados, Departamento de Ecología y Evolución, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay
§Sección Ictología, Departamento de Zoología, Museo Nacional de Historia Natural, Montevideo, Uruguay

Keywords:
adaptive radiation; ecological speciation; phylogeny; ultraconserved elements.

Abstract

The rapid rise of phenotypic and ecological diversity in independent lake-dwelling groups of cichlids is emblematic of the East African Great Lakes. In this study, we show that similar ecologically based diversification has occurred in pike cichlids (Crenicichla) throughout the Uruguay River drainage of South America. We collected genomic data from nearly 500 ultraconserved element (UCEs) loci and >260,000 base pairs across 33 species, to obtain a phylogenetic hypothesis for the major species groups and to evaluate the relationships and genetic structure among five closely related, endemic, co-occurring species (the Uruguay River species flock; URSF). Additionally, we evaluated ecological divergence of the URSF based on body and lower pharyngeal jaw (LPJ) shape and gut contents. Across the genus, we recovered novel relationships among the species groups. We found strong support for the monophyly of the URSF; however, relationships among these species remain problematic, likely because of the rapid and recent evolution of this clade. Clustered co-ancestry analysis recovered most species as well-delimited genetic groups. The URSF species exhibit species-specific body and LPJ shapes associated with specialized trophic roles. Collectively, our results suggest that the URSF consists of incipient species that arose via ecological speciation associated with the exploration of novel trophic roles.

Introduction

The rapid accumulation of ecological roles and their associated adaptations is a hallmark of adaptive radiations (Simpson, 1953; Schluter, 2000). Such rapid diversification may occur in response to ecological opportunities afforded by the colonization of novel environments (e.g. islands or lakes) that provides competitive release (Yoder et al., 2010), the rise of a key innovation that permits exploration of novel regions of the adaptive landscape (Wainwright et al., 2012) and/or histories of hybridization that allow clades to overcome evolutionary constraints and permit the attainment of novel adaptive opportunities (Seehausen, 2004; Meier et al., 2017). Regardless, the role of ecological divergence and ultimately of ecological speciation is central to the progression of adaptive radiations. The environmental gradients along which groups diversify and specialize as well as the degree to which ecological and genetic divergence correspond, are also fundamental characteristics that provide insight into the catalysts that drive adaptive radiations.

Cichlid fishes are textbook examples of adaptive radiations (Seehausen, 2015). In particular, the species flocks of East Africa have adapted to varied environmental conditions and often subsequently diversified extensively in their morphologies and ecologies (Muschick et al., 2012). In addition to the large radiations of Lakes Tanganyika, Malawi and Victoria...
(Wagner et al., 2012), smaller lacustrine radiations also exhibit ecological diversification throughout small African and Middle American lakes (Schliwenn et al., 2001; Barluenga et al., 2006; Elmer et al., 2014; Martin et al., 2015; Ford et al., 2016).

Cichlid radiations exhibit common themes such as specialization along the benthic-to-pelagic habitat axis, soft-bodied-to-hard-shelled prey axis and rapid transitions to herbivory (reviewed in Burress, 2015). Rapid accumulation of ecological roles during cichlid adaptive radiations has conspicuously occurred frequently in lakes and more rarely rivers, which may be due to intrinsic differences in the opportunities afforded by those environments (Seehausen, 2015). For example, river assemblages are generally immigration-assembled such that co-occurring species often represent disparate lineages rather than monophyletic groups (i.e. speciation-assembled), which is common within lakes (Seehausen, 2015). The unstable and unpredictable environmental conditions provided by rivers may favour the evolution of omnivory rather than specialization (Jepsen & Winemiller, 2002), perhaps due to reduced need for the evolution of accommodative processes such as niche partitioning (Grossman et al., 1982), which may act as initial sources of diversifying selection (Shafer & Wolf, 2013). Additionally, many ecomorphs observed in lakes may be implausible evolutionary results in rivers due to niches that are uncommon or temporally unstable in fluctuating environments (Seehausen, 2015).

Cichlids either colonized South America via trans-Atlantic dispersal from Africa (Friedman et al., 2013; Matschiner et al., 2017) or are Gondwanan in origin (Chakrabarty, 2004; Sparks & Smith, 2005; Genner et al., 2007; McMahan et al., 2013). Diversification in locomotor and trophic-associated functional morphology occurred quickly after the origin of the Neotropical clade (López-Fernández et al., 2013; Arbour & López-Fernández, 2014; Burress, 2016; Feilich, 2016). Perhaps the most dramatic case involved the evolution of elongate tubular bodies specialized for feeding via high ram velocity in pike cichlids (Crenicichla; López-Fernández et al., 2013). Despite highly conserved body morphology, pike cichlids have diversified extensively in terms of craniofacial and pharyngeal jaw morphology (Burress et al., 2013a, 2015; Burress, 2016). Following colonization of the La Plata Basin via stream capture with southern tributaries of the Amazon (Reclus, 1893), pike cichlids have been particularly successful, where they exhibit high degrees of endemism and species diversity (de Lucena & Kullander, 1992; Piálek et al., 2012).

Our goals for this study are two-fold. First, we aim to produce a robust phylogenetic framework for all the major lineages of Crenicichla, including all recognized species groups and the type species (C. macrophthalmus). The monophyly of the species groups has been supported by previous analyses with largely mitochondrial loci; however, the relationships among them remain unresolved (Kullander et al., 2010; Piálek et al., 2012; Fig. S1). Second, we assess the diversification of the Uruguay River species flock (URSF). We infer the relationships among five closely related species that are endemic to the Uruguay River (i.e. the URSF): Crenicichla missioneira, C. minutua, C. hadrostigma, C. celidochilus and C. tendybaguassu, which co-occur throughout the drainage and often form mixed species aggregations (de Lucena & Kullander, 1992; de Lucena, 2007; Serra et al., 2011). These species were originally diagnosed based on combinations of morphology, colour pattern and meristics (de Lucena & Kullander, 1992; de Lucena, 2007). Subsequent molecular analyses with primarily mitochondrial loci have not resolved the relationships among these species nor supported their monophyly (Kullander et al., 2010; Piálek et al., 2012).

To address objective 1, we sequenced ultraconserved elements (UCEs) and used a combination of concatenated and multispecies coalescent methods of phylogenetic inference to assess relationships among major species groups and among species. We then discuss these findings in relation to those of previous studies based largely on mitochondrial loci. For objective 2, we assessed the relationships and co-ancestry within the URSF using UCEs. Additionally, to assess ecological diversification within the URSF, we employed landmark-based geometric morphometrics of the body and lower pharyngeal jaw as well as analyses of gut contents. Previous studies have hypothesized that the URSF represents a rapid trophic-based adaptive radiation (de Lucena & Kullander, 1992; Kullander et al., 2010; Piálek et al., 2012; Burress et al., 2013a). Here, we test this hypothesis with a combination of molecular and ecological data and discuss this clade in the context of other prevalent examples of cichlid adaptive radiations that frequently arose in lakes throughout Africa and Middle America.

Materials and methods

Study species

With more than 90 valid species, Crenicichla is the most species-rich genus of the Cichlinae (Neotropical cichlids; Piálek et al., 2015). Crenicichla is present in all major cis-Andean drainages, ranging across coastal Venezuela, the Guianas, the Amazon and La Plata Basin (Piálek et al., 2012). Crenicichla is traditionally divided into five species groups (i.e. clades): the C. lacustris, C. lugubris, C. wallacii, C. saxatilis and C. reticulata species groups (Piálek et al., 2012 and references therein). The species groups are mostly characterized by morphology (e.g. colour pattern and meristics) and geographic distribution. Most species groups occur in sympatry in the Amazon and Orinoco drainages. The C. lacustris species group, however, is distributed in the La Plata Basin (i.e.
Paraná and Uruguay rivers) and South Atlantic coastal drainages (de Lucena & Kullander, 1992; Kullander & Lucena, 2006; Piálek et al., 2012). Crenicichla semifasciata (reticulata group), C. britskii (saxatilis group) and C. leptodota (saxatilis group) are the only species known from other species groups that also have a sub-Amazonian distribution (Kullander et al., 2010; Piálek et al., 2012). Although the monophyly of the five traditional species groups is generally supported, their phylogenetic relationships are uncertain (Fig. S1), as is the relationship of the type species, C. macrophthalmus, which in previous analyses was recovered as a long branch not allied with any of the traditional species groups (Piálek et al., 2012). Additionally, the position and monophyly of Teleocichla, which appears to be nested within Crenicichla, is also unresolved (Kullander et al., 2010; Piálek et al., 2012).

**Phylogeny of Crenicichla**

**Taxon sampling**
The 33 taxa selected for this study (Table S1) represent all species groups established by previous studies (Kullander et al., 2010; Piálek et al., 2012). The nominal species for each of the traditional species groups was included if available. Exceptions include C. lugubris and C. lacustris, but in these cases other representatives were chosen based on previously reported close relationships with the nominal species (Table S1; Kullander et al., 2010; Piálek et al., 2012). Three additional taxa closely related to Crenicichla were chosen as outgroups (López-Fernández et al., 2010; McMahen et al., 2013): Acarichthys heckelii, Apistogramma orinomani and Gymnogeophagus tiraparae (Table S1). Most samples were acquired via museum loans, some tissues were collected during field expeditions to Uruguay in 2013, and additional samples were acquired from the aquarium trade (Table S1). For the URSF, which is part of the C. lacustris species group, we included two to six individuals from each species for the molecular analyses. Voucher specimens are accessioned in the Auburn University Museum of Natural History (AUM), the Academy of Natural Sciences of Drexel University (ANSP) and Universidade Federal do Rio Grande do Sul (UFRGS) (Table S1).

**Library preparation**
We used a sequence capture method to construct 49 DNA libraries enriched for ultraconserved elements (UCEs; complete species list, their species group membership and collection locations are provided in Table S1). UCEs are segments of the genome that are highly conserved (≥100 bp and ≥80% identity) between orthologous regions of evolutionarily divergent taxa (Bejerano et al., 2004; Faircloth et al., 2012). These properties, together with their abundance throughout the genome, little overlap with known paralogous genes, and increasing variability in sequence flanking regions, make UCEs desirable molecular markers that have been proven useful for reconstructing deep phylogenetic relationships, as well as for comparative phylogeography and population genetics at shallower timescales (Faircloth et al., 2012; Smith et al., 2014; Newman & Austin, 2016).

Genomic DNA was extracted from fin or muscle tissue preserved in ethanol or RNA-later using the Omegabiotek E.Z.N.A. animal tissue extraction kit (product #D3396-02) and verified for quality and quantity using agarose gel electrophoresis and a Qubit fluorometer (Life Technologies, Waltham, MA, USA). We randomly sheared 400–1000 ng DNA by sonication to a target size of 400–600 bp using an Agilent MultiPrep Kit (Agilent Technologies). We used a generic SPRI (strepavidin pull-down) protocol (Rohland & Reich, 2012; hereafter SPRI) for all cleanup steps involving magnetic beads. Following adapter ligation, we performed two cleanup steps of the ligation reaction using 0.8X SPRI, resuspended in 33 μL ddH₂O and quantified 2 μL of the resulting library using a Qubit fluorometer. We amplified 15 μL of the adapter-ligated library using a reaction mix of 25 μL 2X Kapa HiFi HotStart ReadyMix (Kapa Biosystems), 5 μL of Illumina TruSeq (San Diego, CA, USA) primer mix (5 μm each) and 5 μL of ddH₂O and the following thermal profile: 98 °C for 45 s, 10 cycles at 98 °C for 15 s, 60 °C for 30 s, 72 °C for 60 s; and a final extension of 72 °C for 5 min. We purified completed PCR using 1X SPRI and resuspended libraries in 23 μL ddH₂O, and we quantified 2 μL of each library using a Qubit fluorometer. We combined groups of eight libraries at equimolar concentrations, having a final concentration of each enrichment pool of 147 ng μL⁻¹ in 7 μL (500 ng in total).

**Target enrichment and sequence of UCEs**
We enriched libraries using a set of 2001 probes (Actino-nots-UCe-0.5Kv1) targeting 500 UCE loci across Actinopterygii (Faircloth et al., 2013). We followed library enrichment procedures for the MYcroarray MYBaits kit v.3.0 (Ann Arbor, MI, USA), except that we added 500 ng custom blocking oligos designed against our custom sequence tags, and using 10 inosines to block the 10-nucleotide index sequence. We ran the hybridization reaction for 24 h at 65 °C. Following hybridization, we bound all pools to streptavidin beads (MyOne CI, Life Technologies) and washed bound libraries to remove nonhybridized and nonspecifically hybridized molecules. We added 30 μL of ddH₂O to each sample and combined 15 μL of streptavidin bead-bound enriched libraries in ddH₂O with 25 μL HiFi HotStart Taq (Kapa Biosystems), 5 μL of Illumina
TruSeq primer mix (5 μM each) and 5 μL of ddH2O. We recovered each library by PCR using the following thermal profile: 98 °C for 2 min; 16 cycles of 98 °C for 20 s, 60 °C for 30 s, 72 °C for 60 s; and a final extension of 72 °C for 5 min. We placed the resulting PCR in a magnet stand and removed the supernatant to separate the PCR-recovered, enriched DNA (supernatant) from the streptavidin beads. We subsequently purified the enriched DNA for each pool using 1X SPRI, and we rehydrated enriched DNA in 33 μL of ddH2O. We quantified 2 μl of each enriched pool using a Qubit fluorometer, and we diluted enriched pools to 2.5 ng μL⁻¹ in 10 μL Tris-HCl. We combined five diluted enriched pools of eight samples and one pool of nine samples from this study with 12 diluted enriched pools from a separate study (146 samples in total) to create an equimolar pool-of-pooled libraries at 10 ng concentration. We sequenced 10 pmol of this mixture in one lane of PE100 sequencing on an Illumina HiSeq (University of Missouri-Columbia DNA Core Facility). Raw read data are archived in the NCBI Sequence Repository Archive (SRA; BioProject ID PRJNA396208), and concatenated and individual gene alignments are archived on Dryad (doi:10.5061/dryad.7qs13).

Analysis of captured sequence data
We preprocessed demultiplexed sequences and prepared them for analyses using programs in the PHYLUCE package (Faircloth, 2016) available at https://github.com/faircloth-lab/phyluce. We trimmed reads to remove adapter contamination and low-quality bases using a parallel wrapper (https://github.com/faircloth-lab/illumini processor) around trimmomatic (Bolger et al., 2014), and assembled cleaned reads using a parallel wrapper around trinity (trinityrnaseq-r2013-02-25; phyluce_assembly_get_match_counts.py; Grabherr et al., 2011; Marcais & Kingsford, 2011).

To identify assembled contigs representing enriched UCE loci, we aligned species-specific contig assemblies to a FASTA file of all enrichment baits using phyluce_assembly_match_contigs_to_probes.py. This program implements the matching process using LASTZ and ensures that matches are 80% identical over 80% of their length. This program also screens and removes potential duplicate contigs or contigs that are hit by baits targeting more than one UCE locus. After screening and removing non-target and duplicated or misassembled contigs, the program creates a relational database containing several tables that map the contig names generated by the assembler to the names of each corresponding UCE locus across all taxa. We used the program phyluce_assembly_get_match_counts.py to query the relational database and generate two lists: one containing those loci having data for all taxa (a complete data matrix) and another containing all loci having data for any taxon (an incomplete data matrix). We input these lists of loci to an additional program (phyluce_assembly_get_fastas_from_match_counts.py) to create separate monolithic FASTA files matching the locus lists for complete and incomplete UCE data matrices. We exploded each monolithic FASTA by locus, and we aligned sequence data for loci containing more than four taxa using phyluce_align_sequence_align.py and MAFFT (Katoh & Standley, 2013). Following alignment, we removed the locus names from all alignments using phyluce_align_remove_locus_name_from_nexus_lines.py.

For the complete UCE data matrix, we computed alignment statistics and the number of informative sites across all alignments using phyluce_align_get_align_summary_data.py and phyluce_align_get_informative_sites.py; we concatenated the resulting alignments into a PHYLIP-formatted supermatrix (phyluce_align_format_nexus_files_for_raxml.py). For the incomplete matrix, we filtered the entire set of aligned loci to create three different incomplete matrices: a 95% complete matrix (alignments contained ≥46 of 49 individuals), an 80% complete matrix (alignments contained ≥39 of 49 individuals) and a 50% complete matrix (alignments contained ≥24 of 49 individuals). Following alignment filtering, we computed alignment statistics and the number of informative sites across all alignments, and we concatenated the resulting alignments into a PHYLIP supermatrix.

Analysis of concatenated UCE data
For all data matrices, we estimated the best-fitting locus-specific site rate substitution models using Cladoforest (Crawford & Faircloth, 2014) and partitioned the UCEs by their best-fitting substitution models. We conducted 20 maximum-likelihood (ML) searches for the phylogenetic tree that best fit the data using the best-fitting partitioning scheme using RAxML v. 8.0.19 (Stamatakis, 2006) and the GTRGAMMA model. Nodal support was assessed by 500 nonparametric bootstrap replicates. We used Bayesian analyses for phylogenetic inference as implemented in MrBayes 3.1 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003), and ran four independent Markov chains of 5 × 10⁶ generations. We sampled trees every 500 iterations to yield 10 000 trees and discarded 25% as burnin. Convergence was confirmed by checking effective sampling size values >200 in TRACER (Rambaut et al., 2014), and by ensuring the average standard deviation of split frequencies was <1%, and the potential scale reduction factor (PSRF) for estimated parameters was 1.0.

In addition to the analysis and representation of evolutionary relationships using phylogenetic trees, network methods provide a useful tool for phylogenetic analysis and visualization of reticulate relationships among taxa and possible mixed ancestry or hybrid individuals. We constructed a neighbour-net analysis using ML distances inferred from concatenated UCE sequence data in Splitstree v.4.12.3 (Huson & Bryant, 2006), and performed 1000 bootstrap replicates to assess support. We also tested for recombination using the Φ statistic.
Species tree analysis
We inferred a species tree of *Crenicichla* using SVDquartets (Chifman & Kubatko, 2014, 2015) as implemented in *paup* version 4.0a150 (Swofford, 2003). The SVDquartet method does not rely on prior inference of individual gene trees; rather, it uses single-site patterns to estimate the species tree in a way that is statistically consistent with the multispecies coalescent. The algorithm takes multilocus SNP data to infer quartet trees for subsets of four species in a coalescent framework, and then combines the set of quartet trees into a species tree using a supertree method (Chifman & Kubatko, 2014). We evaluated 100,000 random quartets and performed 1000 bootstrap replicates of the data to assess support, and then assembled the species tree using the quartet max-cut method (Snir & Rao, 2012).

Genetic structure and ecological diversification of the Uruguay River species flock

Genetic structure analysis
We investigated the genetic structure among species within the URSF using the Markov chain Monte Carlo (MCMC) clustering algorithm implemented in the program fineRADstructure (Malinsky et al., 2016). This program is a modified version of the fineSTRUCTURE package (Lawson et al., 2012), which uses haplotype linkage information of the sequence of all single nucleotide polymorphisms (SNPs) to calculate a co-ancestry matrix based on the most recent coalescence (i.e. the closest relatives for each allele) among sampled individuals, but as opposed to fineSTRUCTURE it does not require information on the chromosomal location of the markers or phased haplotypes. In our case, we assumed perfect linkage among SNPs within each UCE locus and frequent recombination between loci.

To identify SNPs and indels from our UCE data, we created a reference dictionary for one of our samples using Picard (http://broadinstitute.github.io/picard/), and indexed the reference using SAMtools (Li et al., 2009). We then aligned all our read data to the reference using bowtie (Li & Durbin, 2009), and the output SAM files were converted into BAM format using SAMtools. We removed duplicate reads using Picard, to ensure that all our fragments have been independently targeted. We merged all the individual BAM files and realigned them around the indels using IndelRealigner target. We merged all the individual BAM files and ensure that all our fragments have been independently targeted. We then aligned all our read data to the reference using bwa (Li & Durbin, 2009), and the output SAM files were converted into BAM format using SAMtools. We used the Tag Haplotype Matrix as input for the program RADpainter, included in the fineRADstructure package, to calculate the co-ancestry matrix. We assigned individuals to populations using fineSTRUCTURE by running the MCMC for 100,000 generations following an initial 100,000 generations that were discarded as burnin. We visualized and plotted the results using R scripts fineRADstructurePlot.R and FinestructureLibrary.R (available at http://cichlid.gurdon.cam.ac.uk/fineRADstructure.html).

Body and pharyngeal jaw shape analysis
We quantified body and lower pharyngeal jaw (LPJ) shape of the five species within the URSF using museum collections at the Auburn University Museum of Natural History and Universidad Federal do Rio Grande do Sul (Table S2): *Crenicichla celidochilus* \((N = 12)\), *C. hadrostigma* \((N = 6)\), *C. minuano* \((N = 19)\), *C. missioneira* \((N = 24)\) and *C. tendybaguassu* \((N = 7)\). Only specimens that represented adult size classes were included to avoid confounding effects of ontogeny (i.e. Burress et al., 2013b). We photographed specimens in lateral view using a mounted Nikon D5100 digital camera (Nikon Corporation, Tokyo, Japan). Additionally, we dissected, cleaned and photographed the LPJ in dorsal aspect. To quantify body shape, we used 15 homologous and 10 sliding landmarks that describe ecologically meaningful shape variation (Fig. S2a). The placements of sliding landmarks emphasize the craniofacial region because of the hypothesized importance of trophic-based diversification within the URSF (i.e. Pálek et al., 2012; Burress et al., 2013a). We used four homologous and 18 sliding landmarks that describe the shape of the LPJ (Fig. S2b). Sliding landmarks are not associated with a homologous structure, but quantify the curvature between two homologous landmarks. All analyses were performed with the tps program suite. Photographs were consolidated and landmarked using tpsUtil (Rohlf, 2004) and tpsDIG2 (Rohlf, 2006), respectively. Landmarks were superimposed and aligned and relative warps were generated using tpsrelw (Rohlf, 2007). Relative warps are principal components of shape variation and describe the major axes of shape variation among individuals. Scale, rotation and translation were removed from the analysis during superimposition and generation of the Procrustes fit.

To estimate the direction and magnitude of body and LPJ shape change along branches of the URSF phylogeny, we reconstructed a population-level phylomorphospace (Sidlauskas, 2008) by overlaying the ML phylogeny onto the PC biplots of body and LPJ shape. For this procedure, we pruned the set of PC scores to include only the UCE voucher specimens (Table S1). We mapped the phylogeny onto body, and LPJ PC scores using Mesquite v2.75 (Maddison & Maddison, 2011). The values of internal nodes were calculated using weighted squared-change parsimony (Maddison, 1991; Revell et al., 2007).

© 2017 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY. J. EVOL. BIOL. doi: 10.1111/jeb.13196
JOURNAL OF EVOLUTIONARY BIOLOGY © 2017 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY
Gut content analysis

We summarized the diets of four URSF species: C. celi-
dochilus (N = 30), C. tendybaguassu (N = 26), C. minuano
(N = 37) and C. missioneira (N = 44) based on previous
studies (Burress, 2012; Burress et al., 2013a, 2015).
Using the same methods, we also analysed the diet of
C. hadrostigma (N = 15). We quantified the relative pro-
portions of prey items as described by Winemiller
(1989). We removed, sorted, identified and measured
the contents from the anterior half of the digestive tract
in appropriately sized graduated cylinders. For small
items (e.g. insect fragments), we spread sorted material
onto slides and compared its area to a substance of a
known volume. Prey items were pooled into five gener-
alized categories: fishes, insects, molluscs, crustaceans
and periphyton. The periphyton category includes fila-
mentous algae, diatoms and amorphic vegetation.

We calculated niche overlap based on the relative vol-
umetric proportions of these five prey items using Scho-
ener’s (Schoener, 1970) and Pianka’s (Pianka, 1973)
indices. Both indices represent the degree of niche over-
lap between species pairs, where 0 represents total sepa-
ration and 1 represents total overlap. Generally, high
overlap between species pairs, where 0 represents total sepa-
rare, moderate support (BS ≥ 78, PP = 0.77). The second lineage
was formed by the C. saxatilis and C. lugubris species
groups, which always constituted sister clades, and the
C. wallacii group, which was the sister group to them
(Figs S3 and S4), except in the ML and BI 100% com-
plete data set and the ML 95% complete data set
(Figs S3 and S4). Teleocichla was recovered as the sister
group to the clade formed by the C. saxatilis, C. lugubris
and C. wallacii species groups (Figs S3 and S4), except in
the ML 100% and 95% complete data sets in which
Teleocichla was recovered as the sister group to the
C. wallacii species group with poor support, and the BI
100% complete data set that recovered Teleocichla as
the sister group to C. wallacii with high support (PP = 0.94).

The species trees inferred using SVDquartets were
highly consistent with the concatenated phylogenetic
methods (Fig. 1) and across data sets, except for the
100% complete data set, which did not recover the
C. lugubris and C. saxatilis species groups as reciprocally
monophyletic, or any of the relationships within species
groups with high support (Fig. S5). Also, Teleocichla was
recovered as the sister group to all Crenicichla sensu
stricto using the 100% complete matrix, but was recov-
ered as the sister group to the clade containing the
C. wallacii, C. saxatilis and C. lugubris species groups in
the 95%, 80% and 50% matrices (Fig. S5). All the
remaining relationships between species groups, as well
as species-level relationships within groups, were identi-
cal between ML, BI and species tree analyses, and
across 95%, 80% and 50% complete data matrices
(Figs S3, S4 and S5).

The evolution and ecological diversification of the
Uruguay River species flock

Phylogeny and genetic structure

The URSF was monophyletic in all the concatenated
ML and BI analyses (BS = 100, PP = 1.0). The relation-
ships among species within the URSF were mostly
unresolved and differed between methods and data

In general, higher levels of support and monophyly were recovered with increasing number of loci (i.e. decreasing levels of matrix completeness) and were also higher for the BI than for the ML analyses (Fig. 2). For example, using the 100% complete data matrix, only *C. celidochilus* and *C. tendybaguassu* were recovered as monophyletic in the BI analysis (Fig. 2b). With the 80% and 50% complete data set, *C. hadrostigma* was also recovered as monophyletic, with strong support, in the ML and BI analyses (Fig. 2a,b). Using the 80% complete data, *C. missioneira* was also recovered as monophyletic in the

### Table 1 Summary information, including number of reads and contigs, contig lengths and number of loci recovered for all the samples analysed.

<table>
<thead>
<tr>
<th>Species ID</th>
<th>Reads</th>
<th>Contigs</th>
<th>Mean contig length</th>
<th>SD</th>
<th>Min. contig length</th>
<th>Max. contig length</th>
<th>Loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRAC</td>
<td>1,359,076</td>
<td>3,602</td>
<td>1,859.5</td>
<td>588.1</td>
<td>224</td>
<td>74,128</td>
<td>438</td>
</tr>
<tr>
<td>CRALB4786</td>
<td>943,452</td>
<td>3,221</td>
<td>540.9</td>
<td>601.7</td>
<td>224</td>
<td>16,615</td>
<td>426</td>
</tr>
<tr>
<td>CRALT1408</td>
<td>1,812,649</td>
<td>6,217</td>
<td>445.8</td>
<td>460.8</td>
<td>224</td>
<td>16,576</td>
<td>450</td>
</tr>
<tr>
<td>CRBC</td>
<td>1,246,632</td>
<td>3,457</td>
<td>529.8</td>
<td>576.7</td>
<td>224</td>
<td>10,796</td>
<td>431</td>
</tr>
<tr>
<td>CRCE2237</td>
<td>2,583,739</td>
<td>8,832</td>
<td>452.4</td>
<td>449.9</td>
<td>224</td>
<td>15,450</td>
<td>434</td>
</tr>
<tr>
<td>CRCE2238</td>
<td>675,831</td>
<td>2,299</td>
<td>459.8</td>
<td>334.8</td>
<td>224</td>
<td>4975</td>
<td>452</td>
</tr>
<tr>
<td>CRCE2239</td>
<td>682,324</td>
<td>2,323</td>
<td>479.1</td>
<td>410.9</td>
<td>224</td>
<td>5018</td>
<td>455</td>
</tr>
<tr>
<td>CRCI</td>
<td>2,514,513</td>
<td>8,276</td>
<td>459.3</td>
<td>465.3</td>
<td>224</td>
<td>16,096</td>
<td>424</td>
</tr>
<tr>
<td>CROO</td>
<td>1,179,436</td>
<td>3,099</td>
<td>508.2</td>
<td>423.8</td>
<td>224</td>
<td>6168</td>
<td>394</td>
</tr>
<tr>
<td>CRHA1597A</td>
<td>1,198,057</td>
<td>3,481</td>
<td>530.7</td>
<td>505.1</td>
<td>224</td>
<td>13,162</td>
<td>423</td>
</tr>
<tr>
<td>CRHA1597B</td>
<td>702,898</td>
<td>2,172</td>
<td>501.5</td>
<td>479.2</td>
<td>224</td>
<td>15,477</td>
<td>433</td>
</tr>
<tr>
<td>CRHA1597C</td>
<td>1,832,045</td>
<td>5,003</td>
<td>485.2</td>
<td>487.4</td>
<td>224</td>
<td>16,216</td>
<td>422</td>
</tr>
<tr>
<td>CRHA1597D</td>
<td>866,306</td>
<td>2,299</td>
<td>529.1</td>
<td>512.4</td>
<td>224</td>
<td>16,424</td>
<td>435</td>
</tr>
<tr>
<td>CRG4703</td>
<td>1,329,632</td>
<td>3,447</td>
<td>483.5</td>
<td>460.2</td>
<td>224</td>
<td>16,468</td>
<td>430</td>
</tr>
<tr>
<td>CRJE</td>
<td>736,894</td>
<td>2,396</td>
<td>504.2</td>
<td>551.7</td>
<td>224</td>
<td>9346</td>
<td>448</td>
</tr>
<tr>
<td>CRLE2244</td>
<td>2,514,513</td>
<td>8,276</td>
<td>459.3</td>
<td>465.3</td>
<td>224</td>
<td>16,096</td>
<td>424</td>
</tr>
<tr>
<td>CRMACU3462</td>
<td>947,048</td>
<td>2,458</td>
<td>494.9</td>
<td>560.8</td>
<td>224</td>
<td>16,662</td>
<td>439</td>
</tr>
<tr>
<td>CRMAP40588</td>
<td>894,303</td>
<td>2,748</td>
<td>471.1</td>
<td>488.5</td>
<td>224</td>
<td>14,122</td>
<td>453</td>
</tr>
<tr>
<td>CRMN2209</td>
<td>272,073</td>
<td>1,078</td>
<td>465.2</td>
<td>302.9</td>
<td>224</td>
<td>3751</td>
<td>428</td>
</tr>
<tr>
<td>CRMN2240</td>
<td>1,322,722</td>
<td>3,979</td>
<td>507.4</td>
<td>438.7</td>
<td>224</td>
<td>4837</td>
<td>418</td>
</tr>
<tr>
<td>CRMS2201</td>
<td>986,827</td>
<td>3,126</td>
<td>484.7</td>
<td>413.9</td>
<td>224</td>
<td>5743</td>
<td>428</td>
</tr>
<tr>
<td>CRMS2211</td>
<td>765,869</td>
<td>2,238</td>
<td>534.8</td>
<td>483.9</td>
<td>224</td>
<td>4880</td>
<td>437</td>
</tr>
<tr>
<td>CRMS2221</td>
<td>845,254</td>
<td>2,439</td>
<td>523.9</td>
<td>414.3</td>
<td>224</td>
<td>5476</td>
<td>430</td>
</tr>
<tr>
<td>CRMS2226</td>
<td>371,889</td>
<td>1,489</td>
<td>439</td>
<td>594</td>
<td>224</td>
<td>4493</td>
<td>438</td>
</tr>
<tr>
<td>CRMS2249</td>
<td>1,047,233</td>
<td>2,989</td>
<td>496.7</td>
<td>455.8</td>
<td>224</td>
<td>5503</td>
<td>444</td>
</tr>
<tr>
<td>CRMS2256</td>
<td>453,737</td>
<td>1,476</td>
<td>488.1</td>
<td>421.3</td>
<td>224</td>
<td>4504</td>
<td>436</td>
</tr>
<tr>
<td>CRMU189593</td>
<td>585,006</td>
<td>1,633</td>
<td>542.8</td>
<td>503.8</td>
<td>224</td>
<td>5424</td>
<td>449</td>
</tr>
<tr>
<td>CRPE193085</td>
<td>480,653</td>
<td>1,672</td>
<td>492</td>
<td>360</td>
<td>224</td>
<td>4130</td>
<td>427</td>
</tr>
<tr>
<td>CRPU2266</td>
<td>2,006,012</td>
<td>6,326</td>
<td>471</td>
<td>478.1</td>
<td>224</td>
<td>12,936</td>
<td>426</td>
</tr>
<tr>
<td>CRREG</td>
<td>2,772,521</td>
<td>10,280</td>
<td>470.1</td>
<td>503.8</td>
<td>224</td>
<td>15,765</td>
<td>411</td>
</tr>
<tr>
<td>CRSA2274</td>
<td>1,373,904</td>
<td>3,967</td>
<td>556.2</td>
<td>610</td>
<td>224</td>
<td>14,068</td>
<td>422</td>
</tr>
<tr>
<td>CRCB2242</td>
<td>1,876,474</td>
<td>5,892</td>
<td>458.1</td>
<td>434.5</td>
<td>224</td>
<td>13,075</td>
<td>425</td>
</tr>
<tr>
<td>CRSC</td>
<td>755,514</td>
<td>2,561</td>
<td>479.2</td>
<td>383.2</td>
<td>224</td>
<td>4311</td>
<td>447</td>
</tr>
<tr>
<td>CRTE2</td>
<td>1,619,091</td>
<td>4,958</td>
<td>499.1</td>
<td>492.6</td>
<td>224</td>
<td>15,660</td>
<td>416</td>
</tr>
<tr>
<td>CRTE5</td>
<td>1,588,618</td>
<td>4,799</td>
<td>499.1</td>
<td>586.7</td>
<td>224</td>
<td>16,700</td>
<td>438</td>
</tr>
<tr>
<td>CRV2225</td>
<td>1,987,172</td>
<td>6,678</td>
<td>452.4</td>
<td>427.6</td>
<td>224</td>
<td>16,416</td>
<td>420</td>
</tr>
<tr>
<td>CRWA3446</td>
<td>1,489,718</td>
<td>4,481</td>
<td>563.3</td>
<td>595.7</td>
<td>224</td>
<td>15,970</td>
<td>347</td>
</tr>
<tr>
<td>CRX2196400</td>
<td>189,640</td>
<td>873</td>
<td>468.9</td>
<td>433</td>
<td>224</td>
<td>3479</td>
<td>295</td>
</tr>
<tr>
<td>CRX3193066</td>
<td>693,241</td>
<td>2,134</td>
<td>513.7</td>
<td>587.2</td>
<td>224</td>
<td>14,190</td>
<td>447</td>
</tr>
<tr>
<td>CRZE</td>
<td>755,317</td>
<td>2,226</td>
<td>525.1</td>
<td>473.2</td>
<td>224</td>
<td>6029</td>
<td>451</td>
</tr>
<tr>
<td>CRX2196400</td>
<td>349,138</td>
<td>1,187</td>
<td>4304</td>
<td>289</td>
<td>224</td>
<td>4329</td>
<td>429</td>
</tr>
<tr>
<td>CRX3193066</td>
<td>689,165</td>
<td>2,687</td>
<td>471.9</td>
<td>339.2</td>
<td>224</td>
<td>5929</td>
<td>448</td>
</tr>
</tbody>
</table>

The data matrices (Fig. 2). In general, higher levels of support and monophyly were recovered with increasing number of loci (i.e. decreasing levels of matrix completeness) and were also higher for the BI than for the ML analyses (Fig. 2a,b). For example, using the 100% complete data matrix, only *C. celidochilus* and *C. tendybaguassu* were recovered as monophyletic in the BI analysis (Fig. 2b). With the 80% and 50% complete data set, *C. hadrostigma* was also recovered as monophyletic, with strong support, in the ML and BI analyses (Fig. 2a,b). Using the 80% complete data, *C. missioneira* was also recovered as monophyletic in the
BI analysis (PP = 0.95; Fig. 2b). Relationships among species were also unresolved in the species trees across all data matrices (Fig. 2c).

The network constructed from the concatenated UCE data resulted in highly supported monophyletic groups (BS ≥ 92) for C. tendybaguassu, C. celidochilus and C. hadrostigma. On the other hand, the network showed a lack of differentiation with respect to C. missioneira and C. minuano and evidenced the reticulated nature of the relationships between these two species (Fig. 2d). No statistical support for recombination was recovered (all loci P ≥ 0.26).

These results are consistent with the clustered coancestry analysis in fineRADstructure, where the most distinct clusters and largest amounts of co-ancestry were obtained for C. tendybaguassu and C. celidochilus (Fig. 3). Overall, all the species of the URSF were assigned to distinct genetic clusters that share larger co-ancestry levels within clusters, than between them. Between species, the largest amount of shared co-ancestry was found between C. missioneira and C. minuano, whereas the lowest level of co-ancestry sharing was found between C. hadrostigma and C. celidochilus. Additionally, and despite the low sampling size, there

<table>
<thead>
<tr>
<th>Data set completeness</th>
<th>Individuals</th>
<th>Loci</th>
<th>Locus length</th>
<th>Polymorphic sites</th>
<th>Parsimony informative sites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Range</td>
<td>Mean</td>
</tr>
<tr>
<td>100%</td>
<td>49</td>
<td>31</td>
<td>611</td>
<td>74.47</td>
<td>18 964</td>
</tr>
<tr>
<td>95%</td>
<td>46</td>
<td>248</td>
<td>607.47</td>
<td>84.06</td>
<td>150 654</td>
</tr>
<tr>
<td>80%</td>
<td>39</td>
<td>427</td>
<td>573.97</td>
<td>104.77</td>
<td>245 084</td>
</tr>
<tr>
<td>50%</td>
<td>24</td>
<td>485</td>
<td>561.6</td>
<td>113.35</td>
<td>261 143</td>
</tr>
</tbody>
</table>

Table 2 Summary information including number of individuals and loci, locus length, polymorphic and parsimony informative sites for each of the data set.

Fig. 1 Crenicichla phylogenetic hypothesis based on maximum-likelihood analysis of the concatenated data set (a) and SVDquartets species tree analysis (b). All nodes have 100 bootstrap supports unless otherwise noted. Analyses are based on the 80% complete data matrix (see Table 2). For relationships within the Uruguay River species flock (URSF), see Fig. 2. Inset summarizes relationships among species groups. The scale bar is in units of substitutions per site.
appears to exist some degree of population substructure in *C. tendybaguassu* and *C. hadrostigma* (Fig. 3).

**Body and lower pharyngeal jaw shape**

Analysis of body shape resulted in the first two PCs collectively explaining 49.5% of the body shape variation among individuals. Principal component 1 explained 29.3% of the body shape variation and described variation in the relative length and depth of the head, body depth and orientation of the snout and mouth (Fig. 4a). Principal component 2 explained 20.2% of the body shape variation and mainly described variation in relative body depth (Fig. 4a). Species were primarily discriminated along PC1 (Fig. 4a). On one extreme, *C. celidochilus* and *C. missioneira* had elongated heads and terminal mouths, whereas on the other extreme, *C. hadrostigma* had a short head and benthic-oriented snout and mouth. *Crenicichla minuano* also had a short head, but a terminal mouth, and *C. tendybaguassu* had an intermediate head length and shallow body (Fig. 4a). Analysis of LPJ shape resulted in the first two PCs collectively explaining 95.2% of the LPJ shape variation among individuals. Principal component 1 explained 83.5% of the shape variation and described variation in the relative length of the medial and lateral processes (Fig. 4b). Principal component 2 explained 11.7% of the shape variation and described variation in the orientation of the lateral processes (Fig. 4b). Species exhibited species-specific LPJ shapes that were particularly discrete along PC1. *Crenicichla celidochilus* as well as *C. missioneira* had relatively atrophied LPJs with long and widely spaced lateral processes. In contrast, *C. minuano* had hypertrophied LPJs with short robust medial and lateral processes (Fig. 4b). *Crenicichla hadrostigma* had a somewhat robust LPJ with short widely spaced lateral processes and a short medial process, whereas *C. tendybaguassu* had intermediately long lateral processes and a long medial process (Fig. 4b).
Population-level phylomorphospace indicates that some species, such as *Crenicichla hadrostigma* and *C. tendybaguassu*, exhibit little body and LPJ shape variation among populations (Fig. 4c,d). *Crenicichla celidochilus* exhibits more variation in LPJ shape among populations than in body shape (Fig. 4c,d). The *C. minuano* population from the Cuareim River Basin were deeper-bodied than those from the Negro River Basin (Fig. 3c). The *C. missioneira* populations from the Uruguay and Queguay rivers were more elongated than those from the other populations (Fig. 4c), whereas populations from the Cuareim, Uruguay, Queguay and Negro rivers had LPJs with more widely spaced lateral processes (Fig. 4d).

**Gut contents**

Gut contents were different among species, particularly in the relative consumption of fishes, molluscs and periphyton (Fig. 4e). *Crenicichla missioneira* consumed primarily fishes, but also insects and crustaceans (Fig. 4e). The remaining four species specialized on specific types of prey. *Crenicichla celidochilus* consumed almost exclusively fishes (i.e. 91% by volume). In contrast, *C. minuano* consumed primarily molluscs (i.e. bivalves and snails; 73% by volume; Fig. 4e). *Crenicichla tendybaguassu* consumed almost exclusively insects (i.e. 90% by volume; Fig. 4e), and *C. hadrostigma* also consumed primarily insects, but was the only species to consume a large fraction of periphyton (i.e. 26% by volume; Fig. 4e). Schoener’s and Pianka’s measures of niche overlap indicated that most pairwise dietary overlaps were low, with the exception of high overlap between *C. missioneira* and *C. celidochilus* and between *C. hadrostigma* and *C. tendybaguassu* (Table 3).
Discussion

Phylogenomics of Crenicichla

We inferred a phylogenomic hypothesis that is in part consistent with previous morphological and molecular studies. Our phylogeny recovers all Crenicichla species groups as monophyletic, but includes novel relationships among species groups. These relationships were well supported and consistent across phylogenetic methods and varying levels of missing data.

The five species groups within Crenicichla were recognized based on morphology, colour pattern and meristics (Piälek et al., 2012 and references therein). Previous molecular studies support the monophyly of these species groups; however, they exhibit disagreement about the relationships among them (Fig. S1). We recovered novel relationships among the major species groups; it is most noteworthy that the type species, C. macrophthalmalma, has a sister relationship with the C. reticulata species group. This result was consistent between concatenated (ML and BI) and species tree analyses, and across matrices of various degrees of missing data (Figs S3, S4 and S5). In previous studies, the position of C. macrophthalmalma was either completely unresolved (i.e. Kullander et al., 2010) or recovered as the sister lineage to the C. lacustris species group (i.e. Piälek et al., 2012). Crenicichla macrophthalmalma is an Amazonian species; thus, geographically, being allied with a species group consisting of almost exclusively Amazonian species (i.e. the C. reticulata species group) is more likely than being allied with species group consisting of exclusively subtropical species (i.e. the C. lacustris species group).

The positions of the C. wallacii species group and Teleocichla clade have been particularly problematic in previous studies based largely on mitochondrial loci (Kullander et al., 2010; Piälek et al., 2012). Teleocichla was originally distinguished from Crenicichla based on morphology, such as their curved snouts and small mouths (Kullander, 1988), although Ploeg (1991) considered Teleocichla as part of the C. wallacii species group.
Table 3  
Pairwise Schoener’s (above diagonal) and Pianka’s (below diagonal) niche overlap indices within the Uruguay River species flock based on gut content analysis.

<table>
<thead>
<tr>
<th></th>
<th>celidochilus</th>
<th>missioneira</th>
<th>minuano</th>
<th>tendybaguassu</th>
<th>hadrostigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>celidochilus</td>
<td>0.80</td>
<td>0.09</td>
<td>0.38</td>
<td>0.95</td>
<td>0.09</td>
</tr>
<tr>
<td>missioneira</td>
<td>0.97</td>
<td></td>
<td>0.12</td>
<td>0.12</td>
<td>0.14</td>
</tr>
<tr>
<td>minuano</td>
<td>0.03</td>
<td>0.06</td>
<td></td>
<td>0.34</td>
<td>0.31</td>
</tr>
<tr>
<td>tendybaguassu</td>
<td>0.10</td>
<td>0.16</td>
<td>0.42</td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>hadrostigma</td>
<td>0.09</td>
<td>0.16</td>
<td>0.38</td>
<td>0.95</td>
<td></td>
</tr>
</tbody>
</table>

also based on morphology, focusing on their shared diminutive body sizes. We recovered the C. wallaci species group as the sister group to the clade containing the C. lugubris and C. saxatilis species groups, and the Teleocichla clade as the sister group to the clade containing those three species groups (Fig. 1) as in Pialek et al. (2012). These relationships were consistent between concatenated (ML and BI) and species tree analyses and across matrices of various degrees of missing data (Figs S3, S4 and S5). Members of the C. wallaci species group often inhabit litter within small creeks (Montaňa & Winemiller, 2009), whereas Teleocichla inhabit rocky rapids of large rivers (Kullander, 1988; Varella et al., 2016). Therefore, these two groups likely evolved small body sizes in parallel; however, their positions in the phylogeny also suggest that small body sizes may have evolved once followed by the evolution of large-bodied species.

The rapid rise of species flocks

The rise of species flocks in lakes is a hallmark of cichlid evolution, including extensive radiations throughout East Africa (Wagner et al., 2012, 2013; Seehausen, 2015) as well as small radiations in Africa and Middle America (Schliewen et al., 2001; Elmer et al., 2014; Ford et al., 2016); however, such modes of evolution are conspicuously absent from rivers, which tend to be immigration-assembled rather than speciation-assembled (Seehausen, 2015). Species flocks did arise in the lower Congo River (Schwarzer et al., 2011, 2012; Stiassny & Alter, 2015), with the species likely diversifying in allopatry via isolation among rapids and other hydrological and topographical barriers. However, these clades do not exhibit dramatic patterns of morphological and ecological diversification associated with lake-dwelling adaptive radiations (Schwarzer et al., 2011, 2012; Seehausen, 2015; Stiassny & Alter, 2015; Alter et al., 2017).

In previous analyses, based largely on mitochondrial loci, the relationships among species as well as the monophyly of species within the URSF were unresolved (Kullander et al., 2010; Pialek et al., 2012), and this poor resolution was largely attributed to their shallow divergence times (~1.2 My; Pialek et al., 2012). Using hundreds of UCE loci, we show that most species are monophyletic with high support; however, the relationships among species remain elusive (Fig. 2). For example, C. missioneira and C. minuano were the only species not recovered as monophyletic by the Splitstree network (Fig. 2d), and the clustered co-ancestry analysis based on SNPs extracted from UCE loci revealed a higher genetic affinity between them, as indicated by the relatively large shared co-ancestry between individuals from these two species. Other species (i.e. C. tendybaguassu and C. celidochilus), on the other hand, showed long and well-supported branches in the Splitstree network analysis coupled with large amounts of shared co-ancestry among all their individuals. These variable levels of intrapopulation co-ancestry are likely a reflection of differences in evolutionary and demographic histories, such as different degrees of isolation, local selection regimes or effective population sizes (Malinsky et al., 2016; Egger et al., 2017).

In combination with species-specific head and pharyngeal jaw morphologies and associated trophic roles, our molecular results are consistent with ecologically based speciation among incipient species. In particular, the lack of accumulation of significant genetic differentiation among species, and considerable recent shared ancestry between some individuals, suggests that there is likely gene flow within the URSF. Furthermore, the reticulation observed in the phylogenetic network between C. missioneira and C. minuano could be due to events such as hybridization or recombination, among others (Huson & Bryant, 2006). Hybridization often occurs after colonization events and may predispose populations to adaptive radiation by increasing their responses to disruptive and divergent selection (Seehausen, 2004). In these situations, hybridization likely generates (or maintains) sufficient genomic variation, including at functional loci and, by extension, ecologically relevant traits that would permit populations to explore available adaptive peaks. For example, hybrids are often ecologically different than either parent species (Parsons et al., 2011) and therefore have different potential to utilize opportunities afforded by the environment that can ultimately lead to the colonization of different adaptive peaks (Seehausen, 2004; Genner & Turner, 2012). Therefore, hybridization has not only played an important role in the adaptive radiations in Lakes Malawi (Joyce et al., 2011; Malinsky et al., 2017) and Tanganyika (Salzburger et al., 2002; Koblmüller et al., 2010; Meyer et al., 2015, 2017), but also it has...
been hypothesized that it may serve as major substrate from which adaptive radiations arise (Seehausen, 2005).

Rapid ecological diversification

After the colonization of lakes, cichlids have repeatedly diversified into species-rich assemblages that exhibit dramatic morphological and ecological diversity, and in many cases, similar suites of ecomorphs have arisen in each lake (Wagner et al., 2012; Seehausen, 2015). We show that a small radiation of five riverine species of *Crenicichla* have rapidly diversified into five discrete ecomorphs associated with different trophic roles (Fig. 3). This diversity arose rapidly, such that these species are more ecologically differentiated than they are genetically differentiated, likely representing incipient ecological speciation. Furthermore, this clade has diversified along familiar environmental gradients.

Adaptive radiations among lake-dwelling cichlids have used a few ecological axes extensively and repeatedly during their proliferation, namely the benthic-to-pelagic habitat axis and the soft-bodied-to-hard-shelled prey axis (Kidd et al., 2006; Cooper et al., 2010; Muschick et al., 2012; Hulsey et al., 2013; Kusche et al., 2014; Seehausen & Wagner, 2014). The URSF has utilized both of these axes during their radiation. Although most species are associated with benthic habitat and benthic-oriented prey items such as insects, molluscs and periphyton, *C. celidochilus* specializes upon small schooling fishes that occupy open waters (i.e. Characidae; Burress et al., 2013a). Likewise, although most species consume either soft-bodied prey items or fishes, *C. minutano* specializes upon molluscs, including both bivalves and snails (Burress et al., 2013a). *Crenicichla missioneira* is intermediate along both axes. For example, they consume a mixture of benthic- and pelagic-oriented prey items as well as moderately hard-bodied macrocrustaceans (Burress et al., 2013a). Although *C. minutano* and *C. missioneira* are each other’s closest relatives (Fig. 2), they have dramatically different head and LPJ morphologies associated with their highly divergent diets (Fig. 3).

The URSF exhibits highly specialized trophic adaptations that permit the exploitation of resources that would otherwise be inaccessible. For example, molluscs impose functional demands such as the necessity to generate sufficient biting force and the ability to withstand the associated structural stress incurred during the shell crushing process (Hulsey et al., 2008). These adaptations usually involve reinforcement of the pharyngeal jaw bones and the development of molariform teeth (Burress, 2016). *Crenicichla minutano* has one of the most robust LPJs among cichlids (Burress, 2015), characterized by short processes and relatively few, robust, molariform teeth (Fig. 3f), which are specialized for crushing the shells of bivalves and snails (Fig. 3e).

On the opposite end of the spectrum, piscivores often exhibit atrophied LPJs with long processes (Burress et al., 2015), which is likely in response to gape limitations imposed by pharyngeal jaws that can constrain feeding efficiency in piscivores (McGee et al., 2015). This morphology is exhibited by *Crenicichla celidochilus*, which is a specialist piscivore (Fig. 3). Lastly, hypertrophied lips, such as those exhibited by *C. tendybaguassu*, are a rare phenotype among cichlids; however, they have conspicuously evolved in all major clades of cichlids, including in Lakes Tanganyika, Malawi and Victoria (reviewed in Burress, 2015). Hypertrophied lips are also a polymorphism frequently associated with incipient species pairs (Elmer et al., 2010; Colombo et al., 2013; Manousaki et al., 2013; Machado-Schiavino et al., 2014). Their trophic function is associated with foraging from rocky crevices (Baumgarten et al., 2015). Indeed, *C. tendybaguassu* consumes primarily rock-associated insects (Burress et al., 2013a; Fig. 3e).

Rapid transitions to herbivory are rare among fishes (Seehausen & Wagner, 2014), likely due to the physiological constraints associated with digesting nutrient-poor resources (Burress et al., 2016 and references therein); however, in Lakes Tanganyika and Malawi, cichlids have extensively diversified into primary producer-associated trophic roles such as algae grazers (Danley & Kocher, 2001; Wagner et al., 2009; Muschick et al., 2012), which, among other trophic roles, has led to convergence in trophic-associated morphologies in these two lakes (Kocher et al., 1993). The evolution of algae grazing is often associated with small, compact jaws (Winemiller et al., 1995) and sometimes jats that are conspicuously benthic-oriented (Ford et al., 2016). Herbivory has also rapidly evolved within the URSF and in association with predictable changes in head, jaw and mouth morphology. *Crenicichla hadrostigma* grazes algae, probably directly from rock surfaces because their guts lacked detritus and sediment that might indicate a nongrazing mode of feeding (Fig. 3e) that characterizes many other geophagines (i.e. substratesifting; Burress, 2015, 2016). Algae grazing has likely evolved in concert with their compact, benthic-oriented jaws (Fig. 3a,f) as in lake-dwelling clades in which herbivory has rapidly evolved (Ford et al., 2016).

Species flocks replicated in lakes and rivers

We demonstrate that lake-like adaptive radiations have occurred among lake cichlids in the Uruguay River, where five species have evolved rapidly in syntopy in response to trophic-associated mechanisms. Seehausen (2015) discussed several factors that might explain the paucity of such adaptive radiations in rivers, including physical differences between lakes and rivers that might influence the ecological opportunities and therefore evolutionary results in each ecosystem. The obvious spatial disparity between rivers and lakes is the depth...
dimension offered by lakes, which is hypothesized to play a major role during diversification of lake-dwelling cichlids (Seehausen & Magalhaes, 2010) and fishes in general (Seehausen & Wagner, 2014). Secondly, rocky shorelines and reefs associated with many of the East African Great Lakes may provide sufficient conditions for primary producer-associated trophic roles (e.g. algae grazers). Lastly, heterogeneous and unstable conditions associated with rivers are hypothesized to favour the evolution of omnivory (Jepsen & Winemiller, 2002 and references therein), and thus, there may not be suitable conditions for in situ ecological speciation and the subsequent evolution of specialization (Nosil, 2012). It is unclear how the URSF overcame these barriers, but several possibilities exist: (1) the paucity of other cichlid (and noncichlid) lineages associated with elongate tubular bodies (López-Fernández et al., 2013) may have opened new evolutionary opportunities, (2) the shallow nature of the Uruguay River that may have permitted the evolution of primary producer-associated trophic roles and (4) gene flow that may maintain a diverse genetic substrate for selection (Seehausen, 2005) may have allowed this clade to overcome ecological and evolutionary constraints and thereby facilitated their attainment of novel adaptive opportunities. Whatever the case, the URSF lineage is an excellent example of a rapid and diverse cichlid radiation in a river system that in many ways parallels other cichlid radiations that are typically restricted to lacustrine environments.

Data archiving

Raw read data are archived in the NCBI Sequence Repository Archive (SRA: BioProject ID PRJNA396208), and concatenated and individual gene alignments are archived on Dryad (doi:10.5061/dryad.7qs13). Accession numbers for voucher specimens used in the molecular and morphological aspects of this study are available in the Supporting Information.

Acknowledgments

We are grateful to Mark Sabaj and Luiz Malabarba for kindly sharing tissues and Wilson S. Serra and Daniel Hernandez for assistance with fieldwork in Uruguay. Mark Sabaj kindly contributed the photograph of T. monogramma. Feedback from two anonymous reviewers and an Editor improved this paper. Fishes were collected in Uruguay in accordance with Dirección Nacional de Recursos Acuáticos (DINARA) Permit #13/2014. This research was supported by the Jim Smith Endowment Fund from the Ohio Cichlid Association (to E.D.B) and National Science Foundation DEB #1354149 (to P.C.). This is Contribution No. 867 of the Auburn University Museum of Natural History.

References


Burress, E.D. 2012. Food web structure of a subtropical South American stream with particular focus on the co-evolution of form and function in an endemic species flock (M.S. Thesis, Appalachian State University).


