

# Constraints to speciation despite divergence in an old haplochromine cichlid lineage

Claudio Müller<sup>1,2,†</sup>, Florian N. Moser<sup>1,2,†</sup>, David Frei<sup>1,2</sup>, Ole Seehausen<sup>1,2</sup>

<sup>1</sup>Department of Aquatic Ecology, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland

<sup>2</sup>Department of Fish Ecology & Evolution, EAWAG, Centre for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland

<sup>†</sup>shared first author

**Ethics.** This research was done under research permits no. 2013-251-ER-2014-177 (FM) and 2013-251-NA-2014-177 (OS) from the Tanzania Commission for Science and Technology (COSTECH). Behavioural experiments were done under research permits no. BE18/15 and BE65/18 from the Cantonal Veterinary Office of Bern.

**Competing interests.** We have no competing interests.

**Data Accessibility:** RADseq reads are available at the NCBI Sequence Read Archive (SAMN23036539-SAMN32036607). Phenotypic data are available in the Figshare Data Repository DOI: 10.6084/m9.figshare.21075574. Ecological data is provided in the supplementary materials. Behavioural data is provided in the manuscript.

**Authors' contributions.** CM designed and performed the mate choice experiments, dissected and analysed the stomachs, performed the parentage analyses, generated the geometric morphometric data, analysed the data and wrote the manuscript together with FM. FM carried out fieldwork, participated in the design of the study, bred the common-garden fish, generated the genomic data, helped analysing the data, wrote the manuscript together with CM. DF generated the linear morphometric data. OS designed and coordinated the study, made the sampling design, and helped drafting the manuscript.

**Acknowledgements.** We are thankful to the Tanzanian Fisheries Research Institute (TAFIRI) for hosting and supporting our fieldwork and to COSTECH for granting a research permit. We thank Jacco Van Rijssel, Salome Mwaiko, Mhoja Kayeba, Mohamed Haluna, Shane Wright, Oliver Selz, Joana Meier and Jonathan Makoye for their help in the field. We want to thank Andreas Taverna and Erwin Schaefer (animal care), Salome Mwaiko, Corinne Schmid and Susanne Tellenbach (help in the lab), and Beat Kienholz (handiwork for the experimental set-up). Additionally, we want to thank the Genetic Diversity Center (GDC) of the ETH Zurich for providing bioinformatics facilities and the Center for Integrative Genomics at the University of Lausanne for the Sequencing. This research was funded by Swiss National Science Foundation grant 31003A\_163338 to OS.

**Author for correspondence:**

Ole Seehausen (ole.seehausen@eawag.ch)

© The Author(s) 2023. Published by Oxford University Press on behalf of The Society for the Study of Evolution (SSE).

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

## Abstract

Most of the 500+ cichlid species of Lake Victoria evolved very rapidly in the wake of an adaptive radiation within the last 15'000 years. All 500 species have evolved from just one out of five old cichlid lineages that colonized the lake. Endemic to the Lake Victoria region, *Astatoreochromis alluaudi* is a member of an old haplochromine lineage that never speciated in the region. Even though the species occurs in a wide range of habitats, there were no indications of evolutionary diversification. Here, we tested predictions of several hypothetical mechanisms that might constrain speciation, including high dispersal rates, a generalist life style and the lack of behavioral assortative mating. Genomic analyses of individuals from thirteen populations revealed several genomically distinct groups, associated with major habitat classes, indicating the existence of two distinct ecotypes. We found significant phenotypic differences between these ecotypes in the wild, which were retained under common-garden conditions, potentially indicating heritable phenotypic adaptations. Female mate choice experiments revealed the absence of behavioural assortative mating despite genetic and phenotypic differentiation between ecotypes. We suggest that the lack of coupling between behavioral mating preferences and phenotypic and genetic divergence constrains speciation in this cichlid.

### Keywords:

Adaptive radiation, cichlid fish, mate choice, population genomics, reproductive isolation, speciation

# 1. Introduction

There is a large variation in species richness and speciation rates between different taxonomic lineages (Jablonski 2008, Schluter & Pennell 2017, Weir & Schluter 2007). While some of the factors associated with differences in species richness and speciation rates have been identified (Hugall & Stuart-Fox 2012, Rabosky 2016, Rolland et al. 2014), a lot of the observed variation remains unexplained. Many of the species-rich clades show evidence of adaptive radiations; instances where few founder species rapidly diversify into numerous species adapted to different ecological niches. Interestingly, many members of well-studied adaptive radiations co-occur with related lineages that did not radiate and remained species poor. For instance, numerous species of Darwin finches co-occur with a single species of mockingbird on each of the Galapagos Islands (Arbogast et al 2006), and the Caribbean Anoles co-occur with several other lizard clades that did not diversify (Crother & Guyer 1996, Losos 2009). Such settings where lineages with high and lineages with low speciation rates co-occur in a narrow geographic range provide ideal opportunities to study the mechanisms associated with speciation and species persistence.

African cichlids are a famous study system for investigating adaptive radiation (Kocher 2004, Seehausen 2006, Turner 2007, Ronco et al. 2020) because they radiated in multiple lakes into high numbers of different species (Genner et al 2004, Turner et al 2001, Wagner et al 2014). The evolutionary young age of some of these radiations and the large variety of ecological niches occupied by the different cichlid species in these lakes, make them an excellent system to study speciation and adaptation (Seehausen 2006).

The youngest large radiation in the East African Lakes can be found in Lake Victoria, where more than 500 haplochromine cichlid species evolved in less than 15'000 years (Johnson et al 1996), showing an exceptionally high speciation-rate for vertebrates (Rabosky et al 2013, Seehausen 2002, Seehausen 2015). However, not all cichlid lineages that

colonized Lake Victoria diversified equally quickly. In fact, only one of the five cichlid lineages that colonized Lake Victoria account for most of the species richness within the lake (Seehausen 2002, Seehausen 2006, Wagner et al 2012).

The haplochromine cichlid genus *Astatoreochromis* is widespread across several rivers and large lakes and in many little lakes in East Africa, occupying a large variety of different habitats. However, there are only two species described in the genus (Banyankimbona et al 2013): *Astatoreochromis straeleni*, which is restricted to Lake Tanganyika and surrounding rivers, and *A. alluaudi*, which occurs in the basins of Lake Victoria and Lake Edward. The latter is a generalist (Binning & Chapman 2008, Witte 1981) and can be found in many habitats, some of which have facilitated very high rates of speciation in other haplochromine cichlids. In Lake Victoria it is mainly molluscivorous (Greenwood 1965, Hoogerhoud 1986, Mbabazi 2004, Witte 1981), but its diet can vary drastically depending on the environment and food availability. *Astatoreochromis alluaudi* is well known for its phenotypic plasticity, especially in the lower pharyngeal jaw, expressed in response to different food sources (Huysseune 1995, Smits et al 1996, Witte & Van Oijen 1990, Young 2013). Phenotypic plasticity can hamper or facilitate speciation. On one hand, it can hamper evolutionary divergence by acting as a buffer against divergent selection (e.g. Agrawal 2001). On the other hand, it can be key to facilitate speciation and adaptive radiation for instance by allowing rapid ecological range expansion or by enhancing phenotypic differentiation and adaptive habitat choice (Nonaka et al 2015, Pfennig et al 2006, Pfennig et al 2010, Schneider & Meyer 2017).

In this study we investigated genetic and ecological divergence among thirteen populations of *Astatoreochromis alluaudi*, most of them from Lake Victoria but also from Lake Burungi and Lake Rwakajunju in the Kagera River region upstream of Lake Victoria and Lake Saka in the Lake Edward catchment, to the West of Lake Victoria (figure 1). We

chose populations to represent fish from two contrasting habitats: rocky mainland and island reefs in the relatively clear waters of the open lake on the one hand, and on the other hand vegetated swamps, mainland shores and inshore islands in Lake Victoria and satellite lakes (figure 2). Further, we investigated the phenotypic divergence between two populations from ecologically distinct habitats, performed a common garden rearing experiment and experimentally measured behavioural reproductive isolation between these two focal populations. One of these populations is from a swampy, vegetation-rich, slow-flowing inlet stream entering the Mwanza Gulf at the village Sweya in Nyegezi Bay. The other population is from Makobe Island, an offshore rocky reef composed of bare rocks and hardly any aquatic vegetation (figure 2). Makobe Island hosts a very species-rich cichlid community with around 30 distinct species of the haplochromine radiation as well as some non-radiated *Oreochromis* species (Seehausen & Bouton 1997). At Makobe Island many radiation genera occur in two or more closely related species that often occupy slightly different water depth ranges, or differ in the degree to which they live inside of crevices or outside in the open, and are often ecologically distinct in their feeding as well (Bouton et al 1997, Seehausen & Bouton 1997, Seehausen et al 2008). *Astatoreochromis alluaudi* occurs at the exact same location where it occupies the entire available depth range, and occurs in crevices as well as outside, overlapping in habitat utilization with all these other species. Yet, it shows no evidence for morphological or genetic divergence between the deep- and the shallow-living individuals (Moser et al unpublished).

The goal of this study is to better understand the factors that constrain speciation in *A. alluaudi*. In this study we investigated three potential factors. First, we investigated whether divergence within *A. alluaudi* may be hampered due to high dispersal rates, which may prevent genetic divergence between populations and parapatric or allopatric speciation. Second, we tested whether *A. alluaudi* is a habitat generalist that does not require heritable

adaptation to different environments since the species would be equally well suited to most habitats it occupies. This would mean that either phenotypes are similarly well equipped to thrive in different habitats, or individuals adapt to the environment by means of adaptive phenotypic plasticity. In either case no strong divergent selection and therefore no heritable phenotypic divergence would be expected. Third, we investigated whether differences in ecologically, genetically and phenotypically distinct allopatric populations may not have any effects on behavioural mating preferences and hence reproductive isolation, which could otherwise maintain differentiation in secondary contact. Reproductive isolation that is maintained in contact is a crucial step in the speciation process in cichlids (Seehausen 2009, Feller et al. 2020).

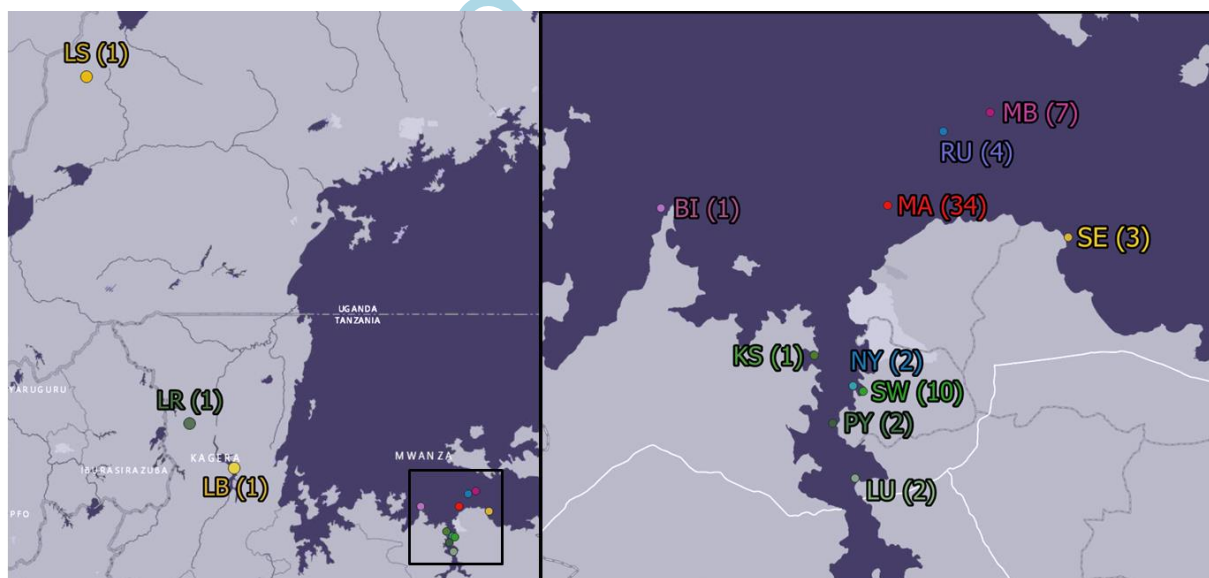
If distinct populations of *A. alluaudi* are genetically and phenotypically differentiated and can retain their distinctiveness in sympatry, the premise of a lack of speciation in this lineage might have to be reevaluated.

## 2. Material and Methods

### a) Fish sampling

*Astatoreochromis alluaudi* were caught from thirteen locations in the south of Lake Victoria, in Lake Burigi, Lake Rwakajunju and Lake Saka (figure 1) by angling, gill netting and minnow traps in several field seasons between 2000 and 2014. From all fish, standardized live pictures were taken before the fish were euthanized with an overdose of Phenoxyethanol. Of every fish, the pelvic fin of the right side was clipped and stored in 100% analytical ethanol for DNA extraction. Voucher specimens were fixed in 4% formalin, rinsed in water and transferred to ethanol through a 3-step dilution series before they were stored in 75% ethanol at the EAWAG Center for Ecology, Evolution and Biogeochemistry (CEEB) in Kastanienbaum, Switzerland.

Additionally, 15 adult individuals from Sweya and 18 adult individuals from Makobe were collected and transported alive to the aquarium facilities at the CEEB in 2014 for common garden rearing and mate choice experiments. To produce lab-bred individuals, fish from the wild were kept in population-specific tanks, where they could interact and mate freely. We obtained six F1 clutches from the Makobe population and ten F1 clutches from the Sweya-population. These were raised separately (one clutch per tank) in identical 20cm/40cm/20cm (length/depth/height) tanks. When they reached a standard length of more than 15 mm, the clutches were consecutively transferred into larger tanks. As soon as male offspring started to express sex-specific coloration, they were separated from their siblings and raised further in clutch-specific male-only tanks. This was done to prevent any experience with mate choice and courtship prior to our experiments. All tanks had a gravel floor, a 12h light-dark cycle, were connected to the same flow-through system and fish were fed flake food *ad libitum* on a daily basis and a mix of peas and shrimps once a week.



**Figure 1:** Map of sampling sites for *A. alluaudi* used in this study. We sampled from three more offshore rocky islands in Lake Victoria, Makobe Island (MA), Ruti Island (RU) and Mabibi Island (MB), two mainland sites or nearshore islands in the open lake, Bihiru (BI) and Senga (SE) and from

five populations within the Mwanza Gulf: Kissenda (KS) in the NW of the Gulf, Sweya (SW), Nyegezi (NY), Python (PY) and Luanso (LU) in the East. Additionally, we included two individuals from lakes west (and upstream) of Lake Victoria (Lake Burigi and Lake Rwakajunju) and the published sequence of an individual from Lake Saka (LS) in the Lake Edward drainage. The numbers in brackets show the number of individuals per location used for the genomics analyses.

## b) Genomics

### *DNA extraction & sequencing*

To investigate population genomic patterns, we used genome-wide Restriction Site Associated DNA (RAD) sequencing. We sequenced seven fish from Mabibi Island, four from Ruti Island, 34 fish from Makobe Island, one individual from Bihiru, three individuals from the rocky mainland shores of Senga, one individual from Kissenda Island from the western shore of the Mwanza Gulf, as well as ten individuals from Sweya and two individuals from Nyegezi, from Python and from Luanso each, and one individual from Lake Burigi and Lake Rwakajunju (figure 1). Individuals from Makobe Island and Sweya were caught by targeted sampling for this study in the field season 2014, while the individuals from other sites were collected while sampling for other species in other years, resulting in smaller numbers. Additionally, we used the published sequence from one individual from Lake Saka in the Lake Edward catchment (figure 1; Meier et al 2017a). We extracted DNA from fin clips using the phenol chloroform DNA extraction protocol provided by Sambrook and Russell (2001), and prepared RAD-libraries using the protocol by Baird et al. (2008) with minor changes indicated in Moser et al. (2018). Sequencing was conducted on an Illumina HiSeq2500 platform at the Centre of Integrative Genomics, University of Lausanne.



## *Population genomic & phylogenomic analyses*

We filtered the raw reads for high quality reads with an intact *Sbf*I restriction site, and de-multiplexed and barcode-trimmed those, using the FASTX toolkit v.0.0.13 (Gordon & Hannon 2010) and custom-made Python scripts. These reads were mapped against the reference genome of *Oreochromis niloticus* (Brawand et al 2014) with Bowtie2 version 1.0.1 (Langmead & Salzberg 2012) and default end-to-end alignment parameters. We used empirical error rate estimations from bacteriophage PhiX reads for base quality score recalibration (Marques et al 2016). Next, we called genotypes and variants with the GATK tool UnifiedGenotyper (McKenna et al 2010), and filtered the resulting vcf file for genotype depth (>3), genotype quality (>30) and missing data (<20% per site) and removed indels with vcftools (Danecek et al 2011). Further we removed individuals with more than 75% missing data. RAxML v8.0.0 (Stamatakis 2014) was used to build a maximum likelihood tree using all concatenated sequences (~1.69 Mio reads) from individuals from all sampling sites. For each of 100 bootstrap replicates, we resampled sites from the concatenated dataset. From each resampled dataset, the maximum likelihood tree was inferred using a GTRGAMMA model. Bootstrap support was then calculated based on these 100 topologies and was plotted on the edges of the best-scoring maximum-likelihood tree calculated from the complete data set.

For principal component analysis (PCA), we removed all monomorphic sites, used only SNPs with a genotype depth >20, genotype quality >30 and less than 20% missing data, and applied a 5% minor allele cutoff in vcftools. Individuals with more than 45% missing data were excluded. PCA was performed in R, using the packages gdsfmt and SNPRelate (Zheng et al 2012).

For STRUCTURE analysis (Pritchard et al 2000) we removed all monomorphic sites and indels, used only SNPs with a genotype depth > 10, genotype quality >30 and less than

20% missing data, and applied a 5% minor allele frequency cutoff in vcfTools. Additionally, we excluded individuals with more than 60% missing data and individuals with highly unbalanced allele frequencies at heterozygous sites, to reduce the effect of potential PCR-errors. We ran STRUCTURE for K=1-6 with six runs for every K. The best supported number of clusters was evaluated with StructureHarvester (Earl & Vonholdt 2012).

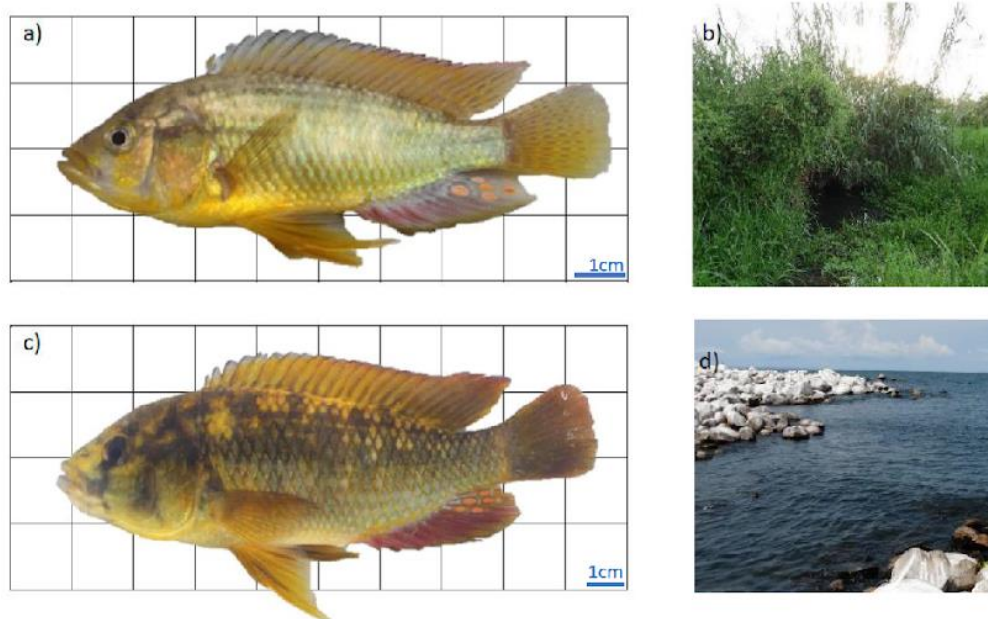
To test for isolation by distance (IBD) and isolation by ecology (IBE), we calculated identity by state (IBS) between all individuals. IBS was calculated using the same dataset that we used to calculate the PCA, but we retained individuals with up to 80% missing data. IBS was then calculated in R, using the Package SNPrelate (Zheng et al 2012). To get a measure of relatedness between populations, we calculated the average of each individual pairing between two populations (table S1). To test for IBD, we constructed a distance matrix with the geographical distances between the different populations using google maps (table S2). To test for isolation by ecology, we used five ecological variables: vegetation (categorical with three states: none, some, plenty), substrate (binary, rock or mud), turbidity (continuous, Secchi-depth), depth range (continuous, maximal sampled water depth at the sampling location in m) and distance to the mainland shore (continuous, in km; table S3). To estimate the ecological distance (habitat contrast) between any two sampling sites, we performed a PCA in R, including all these five ecological and environmental variables. The distance between two populations in the ecospace spanned by PC1 (explains 70.00% of variance) and PC2 (explains 16.96% of variance) was used as an estimate of the ecological habitat contrast between the two populations (table S4). To calculate effects consistent with IBD and IBE, we conducted Mantel tests in the R-package ecodist.

To test for genetic divergence between our two focal populations from Makobe Island and Sweya, we calculated F-statistics using a locus-by-locus AMOVA in Arlequin (Excoffier

et al 2005). For this, we used a stringent missing data per site cutoff ( $<0.8$ ), a minor allele frequency cutoff of 5% and removed monomorphic and multiallelic SNPs.

### c) Morphology

To quantify phenotypic variation between the *A. alluaudi* populations from Makobe Island and Sweya, representing the two major ecotypes (see results), we measured 14 morphometric distances that have been shown to be powerful for capturing taxonomic and ecomorphological differences between closely related haplochromine cichlid species (Banyankimbona et al 2013, Barel et al 1977, Parsons et al 2003). We used digital calipers to measure standard length (SL), head length (HL), head width (HW), body depth (BD), lower jaw length (LJL) and width (LJW), interorbital width (IOW), preorbital width (POW), snout width (SnW), snout length (SnL), eye length (EyL), eye depth (EyD), cheek depth (ChD) and preorbital depth (POD; figure S1) in 83 fish from our two focal populations (N=16 for Sweya, N=67 for Makobe Island; figure 2). Additionally, we measured the length of all fins. Fin length provides additional information, as it can be correlated with swimming performance (Plaut 2000, Qu et al 2013), physical health (Bakker & Mundwiler 1999) and mating success (Bischoff et al 1985). Finally, we counted the number of anal fin spines, a highly conserved meristic trait used in cichlid fish systematics, which was also used to discriminate between the two described *Astatoreochromis* species (Banyankimbona et al 2013).



**Figure 2:** A wild-caught male *A. alluaudi* collected from Sweya (a), a vegetation-rich inlet stream (b) and a wild-caught male *A. alluaudi* caught at Makobe Island (c), an offshore rocky island (d).

To investigate phenotypic plasticity as a possible cause of morphological differences, we took standardized pictures of 30 males raised to maturity in a common-garden environment (16 Sweya, 14 Makobe) and set 7 landmarks in the program tpsDig232 (Rohlf 2015; figure S2). These images of live fish were used to measure four traits corresponding to four traits measured also in the wild-caught fish: Standard length, dorsal fin length, caudal fin length and anal fin length. The other traits we had measured in the wild-caught fish were not measured in our common garden laboratory fish because they would need to be measured from preserved fish and we needed to keep our fish alive for mate choice experiments (see below).

For all analyses, all morphological traits (13 body measurements and five fin measurements), except the anal fin spine counts, were log-transformed. For size-correction we ran an ANCOVA with population and the log-transformed standard length (SL) as covariates for each log transformed morphometric trait. Then we used the standardized

residuals of each trait after extracting the effects of the coefficient of the linear relationship between the trait and SL to conduct t-tests for significant differences. First, we tested for differences between the two sexes of the same population, to check for sexual dimorphism, then between the two populations (Makobe and Sweya) to test for differences between the populations and finally between the two populations for each sex separately. For the anal fin spine number, we conducted a Chi-Square test of independence to test if the number of spines (4, 5 or 6) was independent of the population. Finally, to estimate overall morphological differentiation, we performed principal component analyses (PCA) on the 18 linear morphological traits. All analyses, if not stated differently, were performed in R (R-Core-Team 2015).

#### d) Trophic ecology

To evaluate potential differences in trophic ecology between the two populations, we compared the stomach contents of wild-caught *A. alluaudi* individuals of Makobe Island to those from Sweya. For this purpose, we dissected the stomachs of 36 Makobe and 17 Sweya individuals and quantified volumetric proportions of the different prey categories. Items belonging to four categories (snails, mussels, shrimps and fish) could be identified in the stomachs of our fish. Unidentifiable material was summarized in a fifth category (Hyslop 1980). Zooplankton, detritus, insects and insect larvae, prey typically of many other haplochromines from the same habitats, were not found in any of the stomachs. From the stomach contents, we calculated Schoener's percent resource overlap index (Schoener 1970) between *A. alluaudi* from Sweya and Makobe. Stomachs that were empty (estimated fullness < 25%, N = 7 for Sweya and N = 2 for Makobe) were removed from subsequent analysis.

## e) Mate choice

For the mate choice experiment, we used naive lab-bred common garden fish to exclude environmental factors as well as any previous experience with courtship and mate choice. We conducted the experiment in 143cm/77cm/32.5cm (length/depth/height) tanks, laid out with gravel. All tanks had a slow, but constant waterflow. Within the experimental tank, we used plastic mesh to create four semi-circular areas that were separated from the rest of the tank (figure S3). The mesh size (1.5 cm in the diagonal) was chosen to allow the females to enter and leave these fenced areas but keeping the slightly larger males from leaving their territories. This set-up prevented direct male-male interference while allowing females to directly interact with the males and mate, reducing the influence of male-male interactions on female mate choice to a minimum.

The evening before a trial started, we placed four males, two of each population (of two different clutches each), into the four compartments within the experimental tanks. This allowed them to interact visually but not physically to animate them to express nuptial coloration. Males of the same population were always placed diagonally to each other (figure S3). The position of the first placed male was randomized for each trial. The trial started with the introduction of females into the experimental tank in the subsequent morning. For each trial, we used females of a single family, assuring that none of the stimulus males was full-sibling to the focal females. We used one to seven females per trial, depending on how many females we had raised per clutch. The trials lasted until one of the females had successfully mated.

After each replicate, we took a standardized picture of every male in the trial to measure their standard length and fin lengths (figure S2). After anesthesia with a low dose of MS-222, we took a fin clip and measured their wet weight, before placing them into separate smaller individual tanks to recover. *Astatoreochromis alluaudi* are female mouth-brooders.

The females were retained in the experimental tank for an additional week after the males had been removed to minimize the stress for the mouth-brooding female which otherwise may lead to termination of a brood. We carefully collected all the eggs from the brooding female and incubated them for an additional week, to assure enough tissue for parentage analyses. Embryos were euthanized with an overdose of MS-222 and stored individually in 100% analytical ethanol until DNA extraction. The females that had not mated during the experiment were put back into their original family-specific female-only tank. The female that had mated was anesthetized in a 50mg/L MS-222 solution to take a fin clip for genetic parentage analysis and was then put into a separate tank for recovery.

For parentage analysis, we used tissue from the fin clips of the parents, and tissues from the caudal fins of the embryos. We extracted the DNA following the Chelex extraction protocol. We prepared a primer mix, containing 1.5 $\mu$ l of each primer (table 1; Sanetra et al 2009, Taylor et al 2002) and 55 $\mu$ l TE-Buffer. For each individual we mixed 0.75 $\mu$ l of the extracted DNA solution, with 4.25 $\mu$ l Master-Mix (57% Multiplex Kit, 40% deionized water, 3% Primer Mix) to conduct a PCR (table S5). Genotyping at the five microsatellite loci was performed on an ABI 3130xl Genetic Analyzer (DNA CoreLab) with Size Standard LIZ600. Among the experimental males, we found a total of 58 alleles, of which 11 were present in both populations, 16 were unique to individuals from Sweya and 31 were unique to individuals from Makobe Island (table S6).

To test for deviation from random mating, we used a General Linear Model with the mating status per replicate trial (assortative/disassortative), the father's weight ranking (R1=heaviest, R4=lightest) and the fathers compartment position in the experimental tank (C1=front left, C2=front right, C3=back left, C4=back right) as independent variables (table S11). We also conducted a t-test using the number of fertilized eggs from each replicate to

test if the number of fertilized eggs differed between assortative and disassortative mating events.

**Table 1:** Primers and their sequences used for the parentage analysis, with the repeat motifs and the published size ranges of the microsatellites (Sanetra et al 2009, Taylor et al 2002).

Primer	Sequence Forward	Sequence Reverse	Repeat	Size range
Ppun7	5'-TGACCATCTGCGACAAATAAC-3'	5'-AGGCCTAAGTCCCCCTAACC-3'	<b>GATA</b>	<b>189-299</b>
Ppun21	5'-GGTTGACAGCTGCAAAAATG-3'	5'-AGGCAGTGACCTCTGCTCTC-3'	<b>GATA</b>	<b>311-389</b>
OSU19T	5'-TGAAGGACAAAGCAGGACTG-3'	5'-TGCCCGAACCTTTTTATTTA-3'	<b>CA</b>	<b>126-142</b>
OSU20D	5'-GAAGTGGGATTTGCAGCTTG-3'	5'-CATGCTTACAAAGAACAGGGTTAC-3'	<b>GT</b>	<b>155-211</b>
OSU13d	5'-TAAGCTGATAGGAACCCAAC-3'	5'-ACTCCTATTTTGTATTTTGTGA-3'	<b>GT</b>	<b>105-151</b>

### 3. Results

#### a) Genetic differentiation between allo-parapatric populations of *A. alluaudi*

Phylogenetic analyses based on approximately 1.69 million concatenated sites in individuals from thirteen sampling sites revealed a deep split into two distinct clades (figure 3a, bootstrap support = 100%). The first well-supported clade contains all individuals from sampling sites at the eastern shore of the Mwanza Gulf (SW, NY, PY, LU) and the individuals from the geographically distant Lakes Burigi, Rwakajunju and Saka. Within this clade, the individuals from the distant lakes form a highly supported monophyletic group (Bootstrap support = 94%). The second clade contains all fish from all rocky mainland reefs and islands in the open Lake Victoria studied here (Bihiru, Senga, Makobe, Mabibi and Ruti) and the individual from a rocky island near the western shore of the Mwanza Gulf (Kissenda). Within this clade, Mabibi Island and Ruti Island, the two most offshore rocky

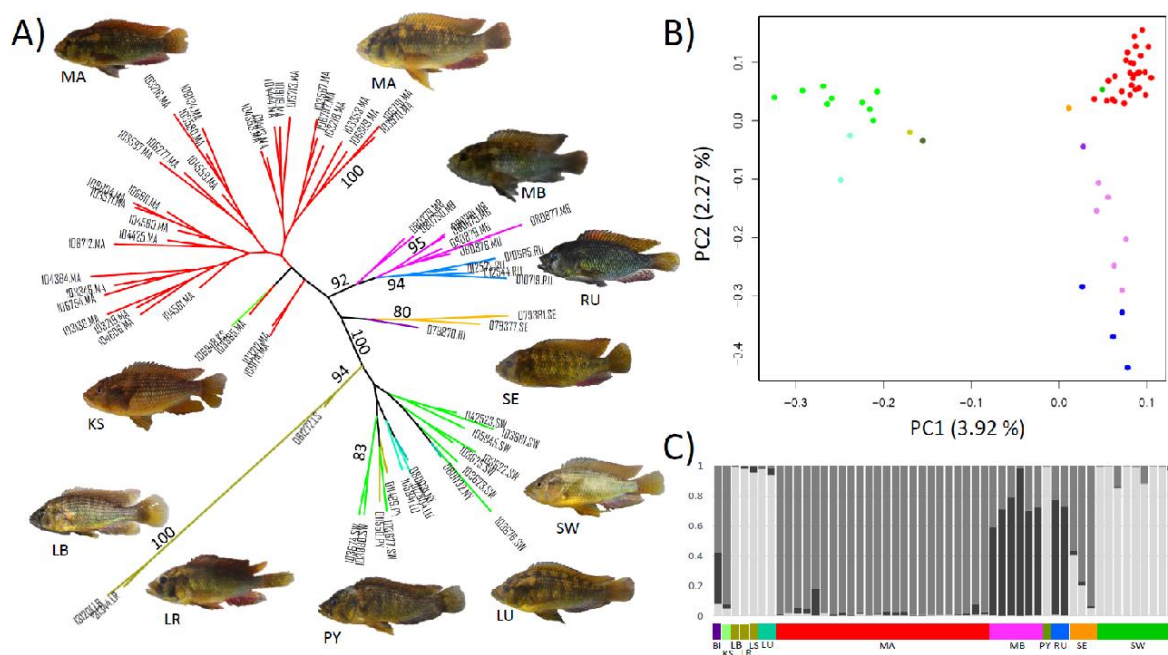


islands in the Speke Gulf, together make a well-supported group (BS = 92%). Within it, all individuals from Ruti form a well-supported monophyletic clade (BS = 94%) that is nested within a paraphyletic Mabibi clade, indicating geographic structure between the offshore rocky islands and between them and the more nearshore and inshore sites. The positions of the individuals from our two open lake mainland reef and nearshore island sites (Senga and Bihiru respectively), relative to the others are not well resolved even though Senga might form a monophyletic clade on its own (figure 3a). On the other hand, Makobe Island is not a strongly supported clade and our fish from Kissenda Island resides within the Makobe clade.

A principal component analysis (PCA) conducted with 2,580 SNPs derived from the RAD sequences of 54 individuals from 10 locations clearly separates the individuals from the open lake mainland reefs and islands (Mabibi, Ruti, Makobe, Senga, Bihiru) and Kissenda from all individuals from the eastern shore of the Mwanza Gulf (Sweya, Python and Luanso) and the individual from Lake Burigi on PC1, which explains 3.92% of the variance. PC2 (2.27%) separates the fish from Ruti and Mabibi Island from those of Makobe, Senga, Bihiru and Kissenda, as well as all others (figure 3b).

STRUCTURE analyses based on 1,277 SNPs revealed a best fit for three genetic clusters (K=3; figure 3c). The first split, revealed by K=2, separates all individuals from the eastern Mwanza Gulf sites (SW, PY and LU) together with Lake Saka (LS), Lake Burigi (LB) and Lake Rwakajunju (LR) from those from the mainland reefs and islands in the open lake (MA, MB, RU, BI, SE) and the northwestern Mwanza Gulf (KS; figure S4). Best supported K=3 further separates the fish from both offshore reefs MB and RU from the other representatives of the second cluster, with about 30-50% assignment to cluster 2. The two open lake mainland and nearshore reef populations (SE and BI) mainly were assigned to the Makobe cluster; but there are indications of introgression from the offshore island cluster into BI and from the eastern Mwanza Gulf cluster into SE. This confirms the pattern revealed by

the genomic PCA with the clearest separation between the individuals from the eastern Mwanza Gulf, LB, LR and LS, and the other locations (along PC1 and in K=2), and an additional split, between Makobe and Kissenda and the two offshore rocky reef populations from the Speke Gulf (RU and MB).



**Figure 3:** a) Maximum Likelihood phylogenetic tree of *A. alluaudi* populations based on over 2 million concatenated sites obtained by RAD sequencing reveals two well-supported clades: Eastern Mwanza Gulf (Sweya = SW, Nyegezi = NY, Python = PY, Luanso = LU; green) plus the lakes west of Lake Victoria (Lake Saka = LS, Lake Rwakajunju = LR, Lake Burigi = LB) on one hand, and the offshore rocky islands Mabibi (MB, purple) and Ruti (RU, blue) on the other hand. Makobe Island (MA, red) and Kissenda Island (KS, green) form a third clade, which is less strongly supported (13% BS). The open lake mainland reef population Senga (SE, violet) and nearshore island population Bihiru (BI, orange) fall outside those three clades. Bootstrap support higher than 70% are reported. B) Principal Component Analyses based on 2,580 SNPs shows a similar pattern. Whereas PC1 separates the fish from the Eastern Mwanza Gulf and the western lakes from the rocky islands and reefs of the open lake, PC2 separates Ruti and Mabibi from Makobe, Kissenda, Senga and Bihiru. C) STRUCTURE assignments of individuals to genetic populations. Analysis with K = 3 reveals a

separation into (i) a population from offshore islands in the Speke Gulf (MB, RU, dark grey); (ii) a population from Makobe (MA) Island, Kissenda (KS) Island, Senga (SE) and Bihiru Island (BI; grey); (iii) a population from all eastern Mwanza Gulf sites (SW, NY, PY, LU) and the distant lakes (LB, LR, LS, light grey). Note signals of introgression between two or all three clades, especially in individuals from the mainland and nearshore reefs outside the Mwanza Gulf (SE, BI).

## b) Genetic differentiation is mainly explained by habitat contrast

The Mantel tests revealed a highly significant correlation between genetic distance and ecological distance (permutations=10000,  $r_m=0.416$ ,  $p=0.003$ ), whereas no correlation between genetic and geographic distance (permutations=10000,  $r_m=-0.145$ ,  $p=0.374$ ). Using partial Mantel tests to test for Isolation by Ecology (IBE) after correcting for geographic distance resulted in a highly significant pattern of IBE (permutations=10000,  $r_m=0.507$ ,  $p<0.001$ ), whereas isolation by distance (IBD) after correcting for ecological distance is significant too but with a much higher p-value (permutations=10000,  $r_m=0.348$ ,  $p=0.033$ ). The genetic structure among *Astatoreochromis* populations therefore is mainly explained by ecological differences between habitats, with only a weak isolation by distance.

The two distinct genetic clusters that we found along PC1 and in STRUCTURE with  $K=2$ , correspond to the two habitat types with very distinct ecological characteristics. One of the two clusters contains all individuals from the three rocky islands in the open lake (Makobe, Ruti and Mabibi), and also the rocky nearshore island (Bihiru) and mainland shore outside the Mwanza Gulf (Senga) but occurs also at Kissenda Island in the northwestern Mwanza gulf, an inshore site with some vegetation. The habitat of this group is characterized by extensive rocky (granite) boulders, general lack of aquatic vegetation and clear water with a Secchi-depth of 2 meters or more (Kissenda is the exception with some vegetation and turbid water). The second genetic cluster contains all individuals from vegetation-rich sites along the eastern shore of the Mwanza Gulf, a shoreline swamp as well as Lake Burigi, Lake

Rwakajunju and Lake Saka. These sampling sites have a medium to large amount of aquatic (usually emergent) vegetation, low water clarity (Secchi-depth of less than 2 meters; table S3) and either soft bottoms or considerable sediment cover on the rocks. The two genetic clusters of *Astatoreochromis alluaudi* can therefore be considered two different ecotypes. Henceforth, we refer to them as the reef ecotype and the vegetation ecotype.

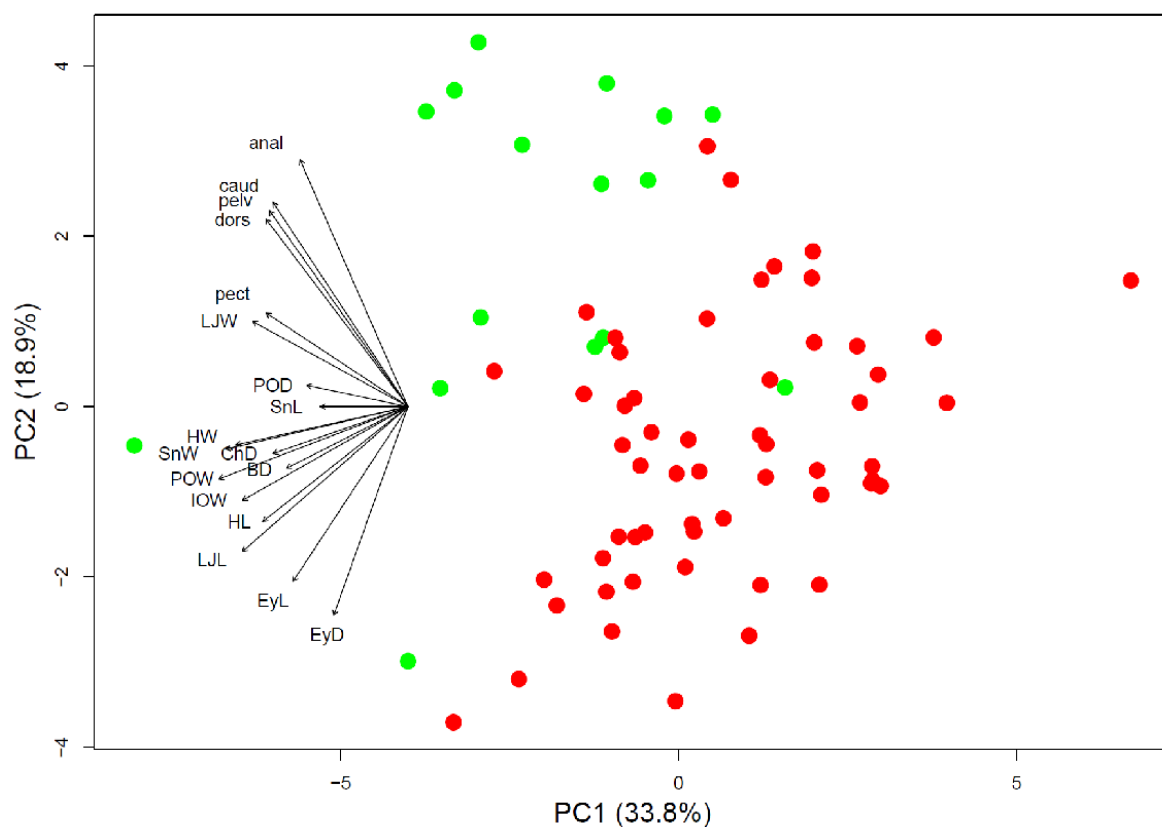
To further test for phenotypic and ecological differentiation and potentially behavioural reproductive isolation between these ecotypes, and hence to test if speciation has occurred, we subsequently focus on the two populations of different ecotype that are geographically the closest to each other: The population at Makobe Island representing the reef ecotype, and the population at Sweya, representing the vegetation ecotype. These two populations show significant genomic differentiation ( $F_{ST}=0.0637$ ,  $p<0.001$ ).

### c) Phenotypic divergence is pronounced between the ecotypes of *A. alluaudi*

We found significant differences between the Makobe and Sweya populations in fin length (pelvic, pectoral, caudal, dorsal and anal fin all longer in the vegetation ecotype; figure 4; table S7), as well as significant differences in lower jaw width (LJW) and snout width (SnW, table S7), with Sweya individuals having longer fins, wider snouts and wider lower jaws than individuals from Makobe. These differences, especially in the fins, are more pronounced in males and hardly present in females (figures S7 & S8), and there is a pronounced sexual dimorphism (figure S9, table S7). This sexual dimorphism is stronger in *A. alluaudi* at Makobe Island than in *A. alluaudi* at Sweya (table S10). The PCA based on all morphological traits shows strong divergence between the two populations and ecotypes (figure 4). PC1 is mainly explained by preorbital width (POW) and snout width (SnW), while PC2 is mainly explained by fin length and eye size (EyD and EyL; table S8).

We also found a significant difference in number of anal fin spines (Chi-Squared=61.83,  $df=2$ ,  $p<0.001$ ). All individuals from Sweya (N=16) have either five (N=10) or six (N=6) anal fin spines. Five spines represents the typical anal fin spine count in this species (Greenwood 1959). Most individuals from Makobe, in contrast, have only four spines (N=50), few had five (N=7) and none had six spines (table S9, figure S5).

In our common-garden raised fish, the significant differences between the two wild populations were recovered in all traits that we could measure (anal fin length, caudal fin length and dorsal fin length; figure S6), indicating a heritable basis for these trait differences.



**Figure 4:** Principal component analysis of linear morphometric traits of *A. alluaudi* from a reef ecotype population (Makobe, red) and a vegetation ecotype population (Sweya, green). It reveals phenotypic differences between the populations with Sweya individuals having a wider head (negative PC1) longer fins and smaller eyes (positive PC2) than individuals from Makobe. The black

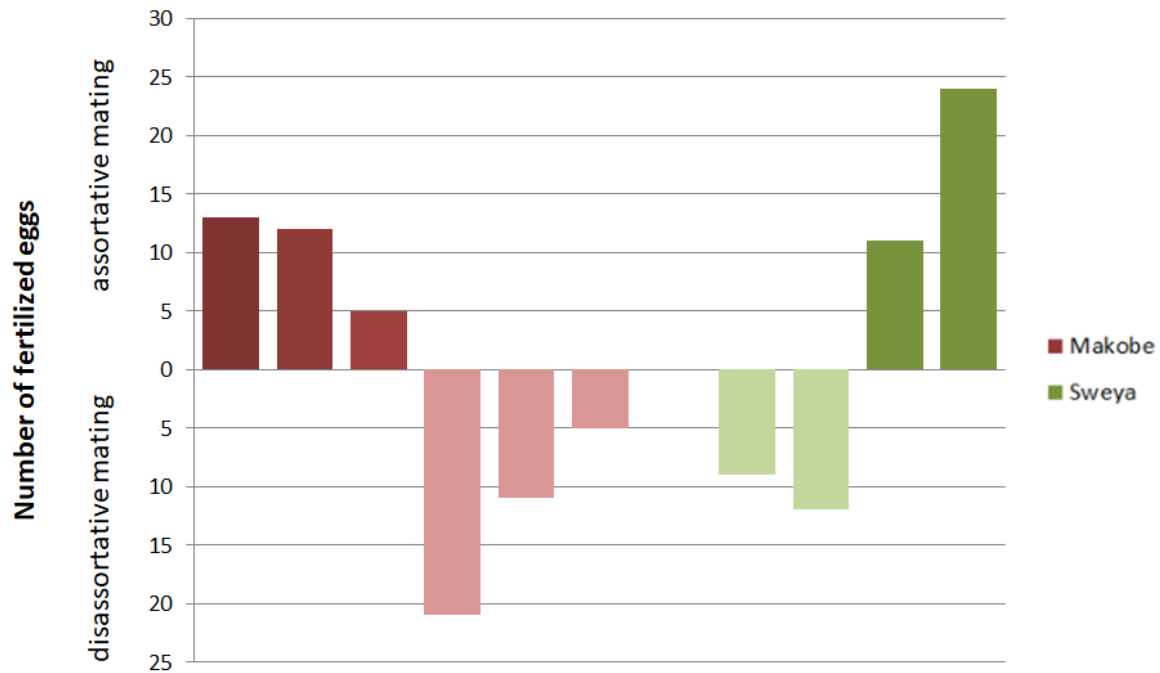
arrows on the left indicate the loadings of the different morphological traits with the different fin lengths in small letters and the other morphological traits in capital letters.

#### d) Trophic ecology is not markedly differentiated

The main prey item in both populations was snails (>50%, figure S10), followed by crushed mussels (~18%). In Sweya, the remaining content could not be identified, whereas for Makobe Island, a small amount of prey items consisted of small fish (7.1%), shrimps (6.8%), as well as some unidentifiable items. Generally, the prey composition of both population samples did not differ significantly (Schoener percent resource overlap index=91.3%).

#### e) No behavioural assortative mating

The ten successful mate choice trials (four replicates with Sweya females and six with Makobe females) resulted in 123 fertilized eggs (table S11), which were used for parentage analyses. Although each female had the simultaneous choice between two Sweya and two Makobe males, every clutch was sired by a single male only. Two of the Sweya females mated with a Makobe male, and two mated with a Sweya male. Three of the Makobe females mated with a Sweya male, and three mated with a Makobe male (figure 5). The generalized linear model shows no significant effect of either male ecotype, weight of the male or experimental compartment of the male, indicating random mating and no indications of any ecotype or population-based assortative female mate choice. Females on average were carrying a similar number of eggs independent of whether they mated with a male from their own or from the other ecotype and population ( $p = 0.739$ , table S11, figure 5).



**Figure 5:** Number of fertilized eggs per trial. Each clutch was sired by a single male. The length of bars indicates the number of fertilized eggs: above zero are eggs resulting from matching ecotypes, below zero are eggs from matings between the two ecotypes.

#### 4. Discussion

We studied the genomic, morphological and ecological differentiation between populations of *Astatoreochromis alluaudi* occupying contrasting habitats. We wanted to investigate the factors that explain absence of radiation in this lineage despite its long presence in the East African Great Lakes. We found strong genetic differentiation and reciprocal monophyly among the rock and vegetation ecotypes. Surprisingly this was robust even across large geographical distances, namely including smaller lakes in the Lake Victoria and Edward catchments. Overall, ecological/habitat differences had a larger effect on genetic divergence than the geographic distance between the populations. We also detected significant phenotypic differences between two populations of the ecotypes, for instance in fin length and jaw morphology. Most of these differences between the ecotypes were retained

in laboratory-bred F1 that were raised under common garden conditions. We did not detect significant differences in diet between ecotype populations. Finally, individuals of both ecotypes mated randomly in behavioral mating experiments under laboratory conditions.

a) Genomic divergence between habitat-specific ecotypes is stronger than isolation by geographic distance

*Astatoreochromis alluaudi* can be found all over Lake Victoria, the smaller lakes and swamps in its catchment and in the Lake Edward catchment, and it populates ecologically very distinct habitats (Slootweg 1989), possibly indicating generalist habits with high dispersal rates and few dispersal limitations. High dispersal rates can cause gene flow between populations and therefore might curtail local adaptation as well as neutral genetic divergence between populations (e.g. Arendt 2015, Claramunt et al 2012, Gavrilets & Vose 2005). However, our results do not support this hypothesis. Several different populations in Lake Victoria are genetically clearly distinct from one another, and populations from the two habitat types even form reciprocally monophyletic clades. Surprisingly, the vegetation ecotype clade of southern Lake Victoria is more closely related to individuals from vegetated habitats in geographically distant lakes (Lakes Burigi, Rwakajunju, Saka, the latter of which is in the Lake Edward catchment) than to the rocky reef ecotype clade from places nearby in southern Lake Victoria. Within Lake Victoria, we detect further subdivision within the rocky reef clade, where the populations of offshore islands form clades that are distinct from those of more nearshore islands and mainland reefs. Taken together, this indicates that the strongest genetic divergence within *A. alluaudi* occurs between populations occupying distinct habitats resulting in a strong pattern of isolation by ecology (IBE), whereas further genetic structure within the reef ecotype is probably attributed to geographic isolation, explaining the comparatively weak pattern of isolation by distance (IBD). This pattern and the fact that the



genomic divergence between the population at Makobe Island and the one in Sweya are in a similar range as between closely related species in similarly distant populations ( $F_{ST} = 0.097$  between *Pundamilia pundamilia* from Makobe Island and *P. pundamilia*-like from Python,  $F_{ST} = 0.062$  between *P. nyererei* from Makobe and *P. nyererei*-like from Python (Meier et al 2017b)) indicate some barriers to gene flow between the different habitats.

#### b) Heritable phenotypic divergence between ecotypes

Species with a generalist strategy tend to undergo muted if any evolutionary phenotypic divergence between different habitats (Martin & Wainwright 2013, Rueffler et al 2006). As a potential habitat generalist (Binning & Chapman 2008, Witte 1981), *A. alluaudi* would not have to divergently adapt to the different habitats and would not be expected to diverge morphologically. Morphological divergence may be an important factor in the speciation process (Kocher 2004) but especially so in the process of establishing sympatry between closely related species (Seehausen 2015).

We found significant morphological differences between the two ecotypes of *A. alluaudi* in multiple traits that may play a role either in ecological adaptation (Bouton et al 1999) or evolve under sexual selection (Baerends & Baerends 1950). Sexual selection can differ in strength between habitats for instance due to differential predation regimes and differences in other ecological variables (Cole & Endler 2015, Endler 1988). This might indicate that the phenotypic differences between the ecotypes might have evolved in response to divergent ecological and sexual selection between the habitats. The fact that these differences were retained in the fish bred and raised in a common-garden environment further suggests that they have a heritable component.

### c) Lack of assortative mating prevents speciation

The ability for sister species to establish sympatry and persist and perhaps diverge further through ecological character displacement is key for the build-up of species diversity in an adaptive radiation (Seehausen 2015). In most animal systems, prezygotic mating barriers evolve as the first cause of reproductive isolation that may allow for persistence upon secondary contact, and eventually co-existence in sympatry (Coyne & Orr 2004, Seehausen 2006). One such mechanism is behavioral assortative mating, which has been suggested as being very important in rapid speciation of cichlid fish (Dominey 1987, Knight & Turner 2004, Maan et al 2004, Seehausen 2000, Selz et al 2014b).

In our mate choice experiment with the two ecotypes, we did not detect any indication of ecotype assortative mating, despite heritable phenotypic and genetic differences between the ecotypes. Similar experiments with member species of the Lake Victoria radiation revealed highly assortative mating (>90%) between individuals of closely related sister species that coexist in sympatry (Selz et al. 2014(a), Selz et al. 2014(b), Selz et al. 2016). While our small sample size does not allow us to rule out very weak mating preferences, a signal of assortative mating similar to that in *Pundamilia* species should have shown. This indicates that the phenotypic differences between the ecotypes do not impact mating decisions, perhaps because the divergent traits are not coupled to mate-choice-cues. Such a lack of female preference for own ecotype despite divergent phenotypic adaptations between the ecotypes may indeed explain the lack of speciation or its completion towards persistence in sympatry in *A. alluaudi*, a constraint that has been described in several other lineages (Maan & Seehausen 2011 and references therein).

#### d) Constraints to speciation in cichlids

Our study provides evidence for the existence of different ecotypes within the species *Astatoreochromis alluaudi*. These ecotypes are genetically and phenotypically distinct and occur in ecologically distinct habitats with no strong physical dispersal barriers between them. The most striking phenotypic differences were found in fin length and head width. All fins are significantly longer in individuals from Sweya than in individuals from Makobe. This pattern is retained in fish bred and raised in a common garden environment, indicating that it is heritable. It is heavily driven by the males, while these differences are less pronounced in females, reflecting the sexual dimorphism in fin length that is more pronounced in the vegetation ecotype. Male cichlids often use the display of their fins in intrasexual competition and for courtship behavior towards females (Baerends & Baerends 1950). Especially important for courtship is the anal fin, which often displays conspicuous background coloration and contains the egg spots, which are considered an important trait for fertilization and have been suggested to play into cichlid speciation (Goldschmidt 1991, Theis et al 2012). Divergent sexual selection on other male nuptial traits is an important factor in haplochromine cichlid speciation (Knight & Turner 2004, Seehausen 2015, Seehausen et al 2008). Significant differences in anal fin length between the males may indicate divergent sexual selection between the ecotypes (Collins et al 2009, Debelle et al 2014), although we were unable to find evidence for sexual selection in our female mate choice trials. The difference in spine numbers is note-worthy, even though its functional importance remains unclear. Anal fin spine number is a highly conserved trait and is used to differentiate *A.alluaudi* (with four to seven spines) from *A. straeleni* (three to four spines) and most other haplochromine cichlids (three spines; Banyankimbona et al 2013).

Despite pronounced genomic and phenotypic differences, *A. alluaudi* ecotypes do not show any indications of assortative mating in experimental mate choice trials. In line with

this, we found evidence for considerable genomic introgression between the wild populations of the different ecotypes, mainly at sites between the range of the vegetation type and the offshore islands occupied by the reef type. Therefore, the differences between the ecotypes that accumulated in allo-parapatry do not seem to be linked to traits causing strong reproductive isolation. This is similar to the pattern detected in another cichlid lineage that occurs in Lake Tanganyika and its inlet streams which has not speciated either (Theis et al 2014), thus pointing to the importance of divergent sexual selection in the evolution and maintenance of the haplochromine cichlid radiations (Allender et al 2003, Seehausen 2015, Seehausen et al 2008). In most cases where genetically close sister species co-occur in sympatry, there is evidence for divergent sexual selection, often triggered by male nuptial coloration sometimes in association with environmental conditions (Maan & Seehausen 2011, Seehausen et al 2008, Selz et al 2014(b)).

Our study and previous studies (Dingle et al. 2008, Lucek et al. 2014, Li et al. 2019, Conith et al. 2020) thus provide evidence that the buildup of genomic and potentially adaptive and heritable phenotypic differences can occur rather frequently in populations living in distinct ecological habitats. However, corresponding mechanisms of reproductive isolation that would allow the transition from local adaptation and ecotype formation to speciation and persistence in sympatry do not necessarily follow. Intrinsic postzygotic reproductive isolation may evolve between allopatric populations when the populations diverge genetically over time (e.g. Ludlow & Magurran 2006). We did not find any evidence for postzygotic isolation in *A. alluaudi*, as mating within and between ecotypes resulted in similar numbers of offspring. Prezygotic isolation however, may arise when immigrants from an ecologically different habitat have a reduced viability (e.g. Matute et al 2009) or reduced mating success (Snowberg & Benkman 2009, Tobler et al 2009). Whereas the former was not addressed in this study, we could not find evidence for the latter.

## 5. Conclusion

We discovered that *Astatoreochromis alluaudi*, an old occupant of the Lake Victoria basin that has never speciated, is composed of two genetically and phenotypically distinct ecotypes with wide geographical distributions. Additionally, there is genetic differentiation between populations from different rocky reefs. The genetic divergence between allopatric or allo-parapatric ecotypes is comparable in extent to that found in sympatric sister species of the radiating lineage. The differentiation is tightly linked to environmental condition rather than geographical distance among populations. The morphological differences are seen in traits that are often linked to either ecological adaptation or sexual selection. As these phenotypic differences are retained under laboratory conditions, it seems likely that they are heritable.

Despite significant phenotypic and genetic divergence, females did not show any preference for males of their own ecotype in our mate choice trials. This indicates that the identification of genomic and phenotypic differences between para- or allopatric populations, even though of importance for management consideration, does not per se allow inferences about speciation without additional evidence for reproductive isolation in sympatry.

We conclude, that the lack of coupling between phenotypic and genetic differences on the one hand and behavioral reproductive isolation on the other hand may be a major constraint to speciation and adaptive radiation.

## 6. References

- Agrawal AA. 2001. Ecology - Phenotypic plasticity in the interactions and evolution of species. *Science* 294: 321-26
- Allender CJ, Seehausen O, Knight ME, Turner GF, Maclean N. 2003. Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. *P Natl Acad Sci USA* 100: 14074-79
- Arbogast BS, Drovetski SV, Curry RL, Boag PT, Seutin G, Grant PR, Grant BR, Anderson DJ. 2006. The origin and diversification of Galapagos mockingbirds. *Evolution* 60: 370-82
- Arendt JD. 2015. Effects of dispersal plasticity on population divergence and speciation. *Heredity* 115: 306-11
- Baerends GP, Baerends R. 1950. An introduction to the study of the ethology of cichlid fishes. *Behaviour supplement* 1: 1-242
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *Plos One* 3: e3376.doi:10.1371/journal.pone.0003376
- Bakker TCM, Mundwiler B. 1999. Pectoral fin size in a fish species with paternal care: a condition-dependent sexual trait revealing infection status. *Freshwater Biol* 41: 543-51
- Banyankimbona G, Vreven E, Snoeks J. 2013. A revision of the genus *Astatoreochromis* (Teleostei, Cichlidae), East-Africa. *Eur J Taxon* 39: 1-21
- Barel CDN, Van Oijen MJP, Witte F, Wittemaas ELM. 1977. Introduction to Taxonomy and Morphology of Haplochromine Cichlidae from Lake Victoria - Manual to Greenwood's Revision Papers. *Neth J Zool* 27: 333-89

- Bischoff RJ, Gould JL, Rubenstein DI. 1985. Tail Size and Female Choice in the Guppy (*Poecilia Reticulata*). *Behav Ecol Sociobiol* 17: 253-55
- Bolnick DI, Fitzpatrick BM. 2007. Sympatric speciation: Models and empirical evidence. *Annu Rev Ecol Evol S* 38: 459-87
- Bouton N, Seehausen O, van Alphen JJM. 1997. Resource partitioning among rock-dwelling haplochromines (Pisces: Cichlidae) from Lake Victoria. *Ecol Freshw Fish* 6: 225-40
- Bouton N, Witte F, van Alphen JJM, Schenk A, Seehausen O. 1999. Local adaptations in populations of rock-dwelling haplochromines (Pisces: Cichlidae) from southern Lake Victoria. *P Roy Soc B-Biol Sci* 266: 355-60
- Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, Fan S, Simakov O, Ng AY, Lim ZW, Bezault E, Turner-Maier J, Johnson J, Alcazar R, Noh HJ, Russel P, Aken B, Alföldi J, Amemiya C, Azzouzi N, Baroiller J, ... Di Palma F. 2014. The genomic substrate for adaptive radiation in African cichlid fish. *Nature* 513: 375-81
- Claramunt S, Derryberry EP, Renssen JV, Brumfield RT. 2012. High dispersal ability inhibits speciation in a continental radiation of passerine birds. *P Roy Soc B-Biol Sci* 279: 1567-74
- Cole GL, Endler JA. 2015. Variable Environmental Effects on a Multicomponent Sexually Selected Trait. *Am Nat* 185: 452-68
- Collins SA, de Kort SR, Perez-Tris J, Telleria JL. 2009. Migration strategy and divergent sexual selection on bird song. *P Roy Soc B-Biol Sci* 276: 585-90
- Conith AJ, Kidd MR, Kocher TD, Albertson RC. 2020. Ecomorphological divergence and habitat lability in the context of robust patterns of modularity in the cichlid feeding apparatus. *BMC Evol Biol* 20: 95
- Coyne JHO, Orr HA. 2004. *Speciation*. Sinauer, Sunderland.

- Crother BI, Guyer C. 1996. Caribbean historical biogeography: Was the dispersal-vicariance debate eliminated by an extraterrestrial bolide? *Herpetologica* 52: 440-65
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MQ, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R. 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156-58
- Debelle A, Ritchie MG, Snook RR. 2014. Evolution of Divergent Female Mating Preference in Response to Experimental Sexual Selection. *Evolution* 68: 2524-33
- Dingle C, Halfwerk H, Slabbekoorn H. 2008. Habitat-dependent song divergence at subspecies level in the grey-breasted wood-wren. *J Evol Biol* 21: 1079-89
- Dominey WJ. 1987. Effects of sexual selection and life history on speciation: species flocks in African cichlids and Hawaiian *Drosophila*. In: Echelle AA, Kornfield I. *Evolution of Fish Species Flocks*. University of Maine at Orono Press: 231-50
- Earl DA, Vonholdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4: 359-61
- Endler JA. 1988. Sexual Selection and Predation Risk in Guppies. *Nature* 332: 593-94
- Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform* 1: 47-50
- Feller AF, Haesler MP, Peichel CL, Seehausen O. 2020. Genetic architecture of a key reproductive isolation trait differs between sympatric and non-sympatric sister species of Lake Victoria cichlids. *P Roy Soc B-Biol Sci* 287: 20200270
- Gavrilets S, Vose A. 2005. Dynamic patterns of adaptive radiation. *P Natl Acad Sci USA* 102: 18040-45



- Genner MJ, Seehausen O, Cleary DFR, Knight ME, Michel E, Turner GF. 2004. How does the taxonomic status of allopatric populations influence species richness within African cichlid fish assemblages? *J Biogeogr* 31: 93-102
- Genner MJ, Turner GF. 2005. The mbuna cichlids of Lake Malawi: a model for rapid speciation and adaptive radiation. *Fish Fish* 6: 1-34
- Givnish TJ, Millam KC, Mast AR, Paterson TB, Theim TJ, et al. 2009. Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *P Roy Soc B-Biol Sci* 276: 407-16
- Goldschmidt T. 1991. Egg Mimics in Haplochromine Cichlids (Pisces, Perciformes) from Lake Victoria. *Ethology* 88: 177-90
- Gordon H, Hannon G. 2010. Fastx-toolkit: FASTQ/A short-reads pre-processing tools. [http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)
- Grant PR, Grant BR. 2008. *How and why species multiply: the radiation of Darwin's finches*. Princeton University Press. JSTOR.
- Greenwood PH. 1959. The monotypic genera of cichlid fishes in Lake Victoria. *Bulletin of British Museum (Natural History)* part II
- Harrison RG. 2012. The Language of Speciation. *Evolution* 66: 3643-57
- Hugall AF, Stuart-Fox D. 2012. Accelerated speciation in colour-polymorphic birds. *Nature* 485: 631-34
- Huyseune A. 1995. Phenotypic Plasticity in the Lower Pharyngeal Jaw Dentition of *Astatoreochromis alluaudi* (Teleostei, Cichlidae). *Arch Oral Biol* 40: 1005-14
- Hyslop EJ. 1980. Stomach Contents Analysis - a Review of Methods and Their Application. *J Fish Biol* 17: 411-29
- Jablonski D. 2008. Species Selection: Theory and Data. *Annu Rev Ecol Evol S* 39: 501-24

- Johnson TC, Scholz CA, Talbot MR, Kelts K, Ricketts RD, et al. 1996. Late pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science* 273: 1091-93
- Knight ME, Turner GF. 2004. Laboratory mating trials indicate incipient speciation by sexual selection among populations of the cichlid fish *Pseudotropheus zebra* from Lake Malawi. *P Roy Soc B-Biol Sci* 271: 675-80
- Kocher TD. 2004. Adaptive evolution and explosive speciation: The cichlid fish model. *Nat Rev Genet* 5: 288-98
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357-59
- Lerner HRL, Meyer M, James HF, Hofreiter M, Fleischer RC. 2011. Multilocus Resolution of Phylogeny and Timescale in the Extant Adaptive Radiation of Hawaiian Honeycreepers. *Curr Biol* 21: 1838-44
- Li MR, Wang HY, Ding N, Lu T, Huang YC, Xiao HX, Liu B, Li LF. 2019. Rapid Divergence Followed by Adaptation to Contrasting Ecological Niches of Two Closely Related Columbine Species *Aquilegia japonica* and *A. oxysepala*. *Genome Biology and Evolution* 11: 919-30
- Losos JB. 2009. *Lizards in an evolutionary tree: ecology and adaptive radiation of anoles*. Berkeley, CA, USA: University of California Press
- Lucek K, Sivasundar A, Kristjánsson BK, Skúlason S, Seehausen O. 2014. Quick divergence but slow convergence during ecotype formation in lake and stream stickleback pairs of variable age. *J Evol Biol* 27: 1878-92
- Ludlow AM, Magurran AE. 2006. Gametic isolation in guppies (*Poecilia reticulata*). *P Roy Soc B-Biol Sci* 273: 2477-82

- Maan ME, Seehausen O. 2011. Ecology, sexual selection and speciation. *Ecol Lett* 14: 591-602
- Maan ME, Seehausen O, Soderberg L, Johnson L, Ripmeester EAP, Mrosso HDJ, Taylor MI, van Dooren TJM, van Alphen JJM. 2004. Intraspecific sexual selection on a speciation trait, male coloration, in the Lake Victoria cichlid *Pundamilia nyererei*. *P Roy Soc B-Biol Sci* 271: 2445-52
- Marques DA, Lucek K, Meier JI, Mwaiko S, Wagner CE, Excoffier L, Seehausen O. 2016. Genomics of Rapid Incipient Speciation in Sympatric Threespine Stickleback. *PLoS Genet* 12: e1005887
- Martin CH, Wainwright PC. 2013. Multiple Fitness Peaks on the Adaptive Landscape Drive Adaptive Radiation in the Wild. *Science* 339: 208-11
- Matute DR, Novak CJ, Coyne JA. 2009. Temperature-based extrinsic reproductive isolation in two species of drosophila. *Evolution* 63: 595-612
- Mayr E. 1963. *Animal Species and Evolution*. Cambridge, MA.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20: 1297-303
- Meier JI, Marques DA, Mwaiko S, Wagner CE, Excoffier L, Seehausen O. 2017a. Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat Commun* 8: 14363
- Meier JI, Sousa VC, Marques DA, Selz OM, Wagner CE, Excoffier L, Seehausen O. 2017b. Demographic modelling with whole-genome data reveals parallel origin of similar *Pundamilia* cichlid species after hybridization. *Mol Ecol* 26: 123-41
- Moser FN, van Rijssel JC, Mwaiko S, Meier JI, Ngatunga B, Seehausen O. 2018. The onset of ecological diversification 50 years after colonization of a crater lake by haplochromine cichlid fishes. *P Roy Soc B-Biol Sci* 285: 20180171

- Moser FN, Van Rijssel JC, Seehausen O. unpublished. A hybrid past and ecological opportunity together predict rapid speciation in sympatric cichlids.
- Nonaka E, Svanback R, Thibert-Plante X, Englund G, Brannstrom A. 2015. Mechanisms by Which Phenotypic Plasticity Affects Adaptive Divergence and Ecological Speciation. *Am Nat* 186: E126-43
- Parsons KJ, Robinson BW, Hrbek T. 2003. Getting into shape: An empirical comparison of traditional truss-based morphometric methods with a newer geometric method applied to New World cichlids. *Environ Biol Fish* 67: 417-31
- Pfennig DW, Rice AM, Martin RA. 2006. Ecological opportunity and phenotypic plasticity interact to promote character displacement and species coexistence. *Ecology* 87: 769-79
- Pfennig DW, Wund MA, Snell-Rood EC, Cruickshank T, Schlichting CD, Moczek AP. 2010. Phenotypic plasticity's impacts on diversification and speciation. *Trends Ecol Evol* 25: 459-67
- Plaut I. 2000. Effects of fin size on swimming performance, swimming behaviour and routine activity of zebrafish *Danio rerio*. *J Exp Biol* 203: 813-20
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-59
- Qu Y, Duan M, Yan J, Feng G, Liu J, et al. 2013. Effects of lateral morphology on swimming performance in two sturgeon species. *J Appl Ichthyol* 29: 310-15
- R-Core-Team. 2015. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing, Vienna, Austria.*
- Rabosky DL. 2016. Reproductive isolation and the causes of speciation rate variation in nature. *Biol J Linn Soc* 118: 13-25

- Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, Chang J, Alfaro ME. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat Commun* 4: 1958
- Rohlf FJ. 2015. The tps series of software. *Hystrix* 26: 9-12
- Rolland J, Condamine FL, Jiguet F, Morlon H. Faster Speciation and Reduced Extinction in the Tropics Contribute to the Mammalian Latitudinal Diversity Gradient. *PloS Biol* 12: e1001775
- Ronco F, Matschiner M, Böhne A, Boila A, Büscher HH, El Taher A, Indermaur A, Malinsky M, Ricci V, Kahmen A, Jentoft S. 2020. Drivers and dynamic of a massive adaptive radiation in cichlid fishes. *Nature* 589: 76-81
- Rueffler C, Van Dooren TJM, Leimar O, Abrams PA. 2006. Disruptive selection and then what? *Trends Ecol Evol* 21: 238-45
- Sambrook J, Russell D. 2001. *Molecular Cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press.
- Sanetra M, Henning F, Fukamachi S, Meyer A. 2009. A Microsatellite-Based Genetic Linkage Map of the Cichlid Fish, *Astatotilapia burtoni* (Teleostei): A Comparison of Genomic Architectures Among Rapidly Speciating Cichlids. *Genetics* 182: 387-97
- Schluter D. 2000. *The Ecology of Adaptive Radiation*. Oxford Series in Ecology and Evolution.
- Schluter D, Pennell MW. 2017. Speciation gradients and the distribution of biodiversity. *Nature* 546: 48-55
- Schneider RF, Meyer A. 2017. How plasticity, genetic assimilation and cryptic genetic variation may contribute to adaptive radiations. *Mol Ecol* 26: 330-50
- Schoener TW. 1970. Nonsynchronous Spatial Overlap of Lizards in Patchy Habitats. *Ecology* 51: 408-18

- Seehausen O. 2000. Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: Effects of sexual selection. *Adv Ecol Res* 31: 237-74
- Seehausen O. 2002. Patterns in fish radiation are compatible with Pleistocene desiccation of Lake Victoria and 14 600 year history for its cichlid species flock. *P Roy Soc B-Biol Sci* 269: 491-97
- Seehausen O. 2006. African cichlid fish: a model system in adaptive radiation research. *P Roy Soc B-Biol Sci* 273: 1987-98
- Seehausen O. 2009. Progressive levels of trait divergence along a "speciation transect" in the Lake Victoria cichlid fish *Pundamilia*. *Speciation and patterns of diversity (Cambridge, UK: Cambridge University Press)*: 155-76
- Seehausen O. 2015. Process and pattern in cichlid radiations - inferences for understanding unusually high rates of evolutionary diversification. *New Phytol* 207: 304-12
- Seehausen O, Bouton N. 1997. Microdistribution and fluctuations in niche overlap in a rocky shore cichlid community in Lake Victoria. *Ecol Freshw Fish* 6: 161-73
- Seehausen O, Terai Y, Magalhaes IS, Carleton KL, Mrosso HDJ, Miyagi R, van der Sluijs I, Schneider MV, Maan ME, Tachida H, Imai H, Okada N. 2008. Speciation through sensory drive in cichlid fish. *Nature* 455: 620-23
- Selz OM, Thommen R, Maan ME, Seehausen O. 2014 (a). Behavioural isolation may facilitate homoploid hybrid speciation in cichlid fish. *J Evol Biol* 24: 275-89
- Selz OM, Pierotti MER, Maan ME, Schmid C, Seehausen O. 2014 (b). Female preference for male color is necessary and sufficient for assortative mating in 2 cichlid sister species. *Behav Ecol* 25: 612-26
- Selz OM, Thommen R, Pierotti MER, Anaya-Rojas JM, Seehausen O. 2016. Differences in male coloration are predicted by divergent sexual selection between populations of a cichlid fish. *P Roy Soc B-Biol Sci* 283: 20160172

- Slootweg R. 1989. Proposed introduction of *Astatoreochromis alluaudi*, an East African mollusk crushing cichlid, as a mean of snail control. *Musee Royal de l'Afrique Central, Annales Sciences Zoologique* 257: 61-64
- Smits JD, Witte F, VanVeen FG. 1996. Functional changes in the anatomy of the pharyngeal jaw apparatus of *Astatoreochromis alluaudi* (Pisces, Cichlidae), and their effects on adjacent structures. *Biol J Linn Soc* 59: 389-409
- Snowberg LK, Benkman CW. 2009. Mate choice based on a key ecological performance trait. *J Evol Biol* 22: 762-69
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312-13
- Taylor MI, Meardon F, Turner G, Seehausen O, Mrosso HDJ, Rico C. 2002. Characterization of tetranucleotide microsatellite loci in a Lake Victorian, haplochromine cichlid fish: a *Pundamilia pundamilia* x *Pundamilia nyererei* hybrid. *Mol Ecol Notes* 2: 443-45
- Theis A, Ronco F, Indermaur A, Salzburger W, Egger B. 2014. Adaptive divergence between lake and stream populations of an East African cichlid fish. *Mol Ecol* 23: 5304-22
- Theis A, Salzburger W, Egger B. 2012. The Function of Anal Fin Egg-Spots in the Cichlid Fish *Astatotilapia burtoni*. *Plos One* 7: e29878
- Tobler M, Riesch R, Tobler CM, Schulz-Mirbach T, Plath M. 2009. Natural and sexual selection against immigrants maintains differentiation among micro-allopatric populations. *J Evolution Biol* 22: 2298-304
- Turner GF. 2007. Adaptive radiation of cichlid fish. *Curr Biol* 17: 827-31
- Turner GF, Seehausen O, Knight ME, Allender CJ, Robinson RL. 2001. How many species of cichlid fishes are there in African lakes? *Mol Ecol* 10: 793-806
- Wagner CE, Harmon LJ, Seehausen O. 2012. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* 487: 366-69

- Wagner CE, Harmon LJ, Seehausen O. 2014. Cichlid species-area relationships are shaped by adaptive radiations that scale with area. *Ecol Lett* 17: 583-92
- Weir JT, Schluter D. 2007. The Latitudinal Gradient in Recent Speciation and Extinction Rates of Birds and Mammals. *Science* 315: 1574-76
- Witte F, Van OijenMJP.1990. Taxonomy, ecology and fishery of Lake Victoria haplochromine trophic groups. *Zoologische Verhandelingen* 262: 1-47
- Young RL. 2013. Linking conceptual mechanisms and transcriptomic evidence of plasticity-driven diversification. *Mol Ecol* 22: 4363-65
- Zheng XW, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28: 3326-28

Accepted Manuscript



## TABLES AND FIGURE CAPTIONS

**Table 1:** Primers and their sequences used for the parentage analysis, with the repeat motifs and the published size ranges of the microsatellites (Sanetra et al 2009, Taylor et al 2002).

Primer	Sequence Forward	Sequence Reverse	Repeat	Size range
Ppun7	5'-TGACCATCTGCGACAAATAAC-3'	5'-AGGCCTAAGTCCCCCTAACC-3'	<b>GATA</b>	<b>189-299</b>
Ppun21	5'-GGTTGACAGCTGCAAAAATG-3'	5'-AGGCAGTGACCTCTGCTCTC-3'	<b>GATA</b>	<b>311-389</b>
OSU19T	5'-TGAAGGACAAAGCAGGACTG-3'	5'-TGCCCGAACCTTTTTATTTA-3'	<b>CA</b>	<b>126-142</b>
OSU20D	5'-GAAGTGGGATTTGCAGCTTG-3'	5'-CATGCTTACAAAGAACAGGGTTAC-3'	<b>GT</b>	<b>155-211</b>
OSU13d	5'-TAAGCTGATAGGAACCCAAC-3'	5'-ACTCCTATTTTGTATTTTGTGA-3'	<b>GT</b>	<b>105-151</b>

Accepted Manuscript

**Figure 1:** Map of sampling sites for *A. alluaudi* used in this study. We sampled from three more offshore rocky islands in Lake Victoria, Makobe Island (MA), Ruti Island (RU) and Mabibi Island (MB), two mainland sites or nearshore islands in the open lake, Bihiru (BI) and Senga (SE) and from five populations within the Mwanza Gulf: Kissenda (KS) in the NW of the Gulf, Sweya (SW), Nyegezi (NY), Python (PY) and Luanso (LU) in the East. Additionally, we included two individuals from lakes west (and upstream) of Lake Victoria (Lake Burigi and Lake Rwakajunju) and the published sequence of an individual from Lake Saka (LS) in the Lake Edward drainage. The numbers in brackets show the number of individuals per location used for the genomics analyses.

**Figure 2:** A wild-caught male *A. alluaudi* collected from Sweya (a), a vegetation-rich inlet stream (b) and a wild-caught male *A. alluaudi* caught at Makobe Island (c), an offshore rocky island (d).

**Figure 3:** a) Maximum Likelihood phylogenetic tree of *A. alluaudi* populations based on over 2 million concatenated sites obtained by RAD sequencing reveals two well-supported clades: Eastern Mwanza Gulf (Sweya = SW, Nyegezi = NY, Python = PY, Luanso = LU; green) plus the lakes west of Lake Victoria (Lake Saka = LS, Lake Rwakajunju = LR, Lake Burigi = LB) on one hand, and the offshore rocky islands Mabibi (MB, purple) and Ruti (RU, blue) on the other hand. Makobe Island (MA, red) and Kissenda Island (KS, green) form a third clade, which is less strongly supported (13% BS). The open lake mainland reef population Senga (SE, violet) and nearshore island population Bihiru (BI, orange) fall outside those three clades. Bootstrap support higher than 70% are reported. B) Principal Component Analyses based on 2,580 SNPs shows a similar pattern. Whereas PC1 separates the fish from the Eastern Mwanza Gulf and the western lakes from the rocky islands and reefs of the open lake, PC2 separates Ruti and Mabibi from Makobe, Kissenda, Senga and Bihiru. C) STRUCTURE assignments of individuals to genetic populations. Analysis with  $K = 3$  reveals a separation into (i) a population from offshore islands in the Speke Gulf (MB, RU, dark grey); (ii) a population from Makobe (MA) Island, Kissenda (KS) Island, Senga (SE) and Bihiru Island (BI; grey); (iii) a population from all eastern Mwanza Gulf sites (SW, NY, PY, LU) and the distant lakes (LB, LR, LS, light grey). Note signals of introgression between two or all three clades, especially in individuals from the mainland and nearshore reefs outside the Mwanza Gulf (SE, BI).

**Figure 4:** Principal component analysis of linear morphometric traits of *A. alluaudi* from a reef ecotype population (Makobe, red) and a vegetation ecotype population (Sweya, green). It reveals phenotypic differences between the populations with Sweya individuals having a wider head (negative PC1) longer fins and smaller eyes (positive PC2) than individuals from Makobe. The black arrows on the left indicate the loadings of the different morphological traits with the different fin lengths in small letters and the other morphological traits in capital letters.

**Figure 5:** Number of fertilized eggs per trial. Each clutch was sired by a single male. The length of bars indicates the number of fertilized eggs: above zero are eggs resulting from matching ecotypes, below zero are eggs from matings between the two ecotypes.

Accepted Manuscript