Recent sympatric speciation involving habitat-associated nuptial colour polymorphism in a crater lake cichlid

MELISSA LEMOINE[1,2], MARTA BARLUENGA[3], KAY LUCEK[4], SALOME MWAIKO[1,2], MARCEL HAESLER[1,2], LAUREN J. CHAPMAN[5], COLIN A. CHAPMAN[6], and OLE SEEHAUSEN[1,2]

[1]*Institute of Ecology and Evolution, University of Bern, CH-3012 Bern, Switzerland*

[2]*EAWAG Center of Ecology, Evolution and Biochemistry, CH-6047 Kastanienbaum, Switzerland*

[3]*Museo Nacional de Ciencias Naturales, CSIC, Jose´ Gutierrez Abascal 2, 28006, Madrid, Spain*

[4]*Department of Environmental Sciences, University of Basel, CH-4056 Basel, Switzerland*

[5]*Department of Biology, McGill University, Montréal, QC, Canada*

[6]*Department of Anthropology and McGill School of Environment, McGill University, Montréal, QC, Canada*

Keywords: crater lakes, Lake Victoria region, sexual selection, microsatellites, Cichlidae

Correspondence: OLE SEEHAUSEN[1]; E-mail: ole.seehausen@eawag.ch

Running title: Colour polymorphism in a crater lake cichlid

ABSTRACT

Whereas the idea that modes of speciation other than allopatric speciation are possible in nature is now widespread, compelling examples of ecological speciation in sympatry remain rare. We studied an undescribed radiation of haplochromine cichlids in a young crater lake in western Uganda, and in the small river that is nearby but has currently no known surface connection to the lake. We describe two different modes of speciation that occurred in this cichlid lineage within the past 1'500-10'000 years. Not constrained by gene flow, allopatric divergence between river and lake cichlids affects many different morphological traits as well as nuptial coloration – muted in the river, but intensified and polymorphic in lake cichlids – and neutral genetic differentiation. More surprisingly, we demonstrate a case for sympatric speciation within the small lake that is associated with dramatic differences in male breeding colouration (yellow with bright red-chest versus bright blue) and subtle differences in microhabitat, feeding regime and morphology. Reproductive isolation by assortative mating is suggested by significant differentiation between yellow and blue males in neutral markers of gene flow despite complete sympatry. We hypothesize speciation is mediated by divergent selection on sexual signalling between microhabitats.

INTRODUCTION

For many decades, allopatric speciation was the only widely accepted mode of speciation (Futuyma & Mayer, 1980). This view has been increasingly challenged by theoretical and empirical studies, and evidence that other geographical modes of speciation are possible in nature is strong (Nosil, 2008; Santini et al., 2012). However, there is still much debate regarding the conditions promoting these alternative modes of speciation (Bolnick & Fitzpatrick, 2007; Santini et al., 2012; Feder et al., 2013; Seehausen et al., 2014). Defining the spatial scale of speciation does not in itself constitute a characterization of the mechanisms (Crow et al., 2010). Thus, identification of the driving force for divergence (e.g., divergent ecological or sexual selection between habitats, disruptive natural selection or disruptive sexual selection) is critical (Butlin et al., 2008). In this context, the diversity of cichlid fish in lakes of various size and isolation and with diverse ecological conditions, provides a suitable study system for intraspecific divergence and sympatric speciation. Cichlid fishes combine great richness of sympatric species with substantial phenotypic divergence between species and evidence for rapid speciation (Kocher, 2004; Barluenga et al., 2006; Seehausen, 2006; Elmer et al. 2010; Malinsky et al., 2015; Kautt et al. 2016b; Meier et al., 2017b; Moser et al. 2018). The East African Great Lakes (Victoria, Malawi, Tanganyika), with their extraordinarily rich cichlid faunas have been extensively studied and provide great opportunities for studying speciation and adaptive radiation (Muschick et al., 2012; Wagner et al., 2014). However, the large size of these lakes, and their historical and ecological complexity make it difficult to ask questions specifically about the role of space in speciation (Kisel & Barraclough, 2010; Nosil, 2012).

Small crater lakes that constitute young and relatively isolated habitats but host endemic monophyletic species pairs or clades, provide powerful model systems for studying the role of space in cichlid speciation. Source populations can be identified, introgression from non-sister taxa can be tested, and the spatial scale of population differentiation can be explicitly measured (Malinsky et al., 2015; Kautt et al., 2016 a&b). Endemics of crater lakes are considered among the strongest empirical examples of sympatric speciation (Schliewen et al., 1994; Coyne & Orr, 2004; Barluenga et al. 2006). To date, sympatric cichlid speciation in crater lakes has been invoked in the genus *Sarotherodon* (two crater lakes in Cameroon), in the genus *Coptodon* (two crater lakes in Cameroon), the genus *Amphilophus* (several crater lakes in Nicaragua), and the genus *Astatotilapia* (one crater lake in Tanzania). Although recent work revealed that many of these lakes have been colonized more once (Martin et al. 2015; Kautt et al. 2016 a&b), the data are still consistent with intra-lake speciation, albeit perhaps with genetic input from outside. Whether the gene flow from outside the lake has facilitated this intralacustrine speciation remains to be investigated. With one exception (Malinsky et al. 2015), sympatric speciation in crater lakes has been demonstrated in lineages that are not part of the repeated large-scale radiations in East Africa. Whereas rapid speciation in the haplochromine

cichlids of the Great Lakes is often attributed to the action of divergent sexual selection and its interaction with ecology (Allender et al. 2003; Seehausen et al. 2008; Wagner et al. 2012), sympatric speciation in the crater lakes has been attributed mainly to disruptive ecological selection (Schliewen et al. 2001; Barluenga et al. 2006) and indeed many of the crater lake cichlids do not show bright sexually selected coloration, have monogamous mating systems and no strong sexual dichromatims. Several authors have studied crater lake populations of haplochromine cichlids, but found no evidence of speciation (Sato et al. 2003, Samonte et al. 2007, Machado-Schiaffino et al. 2015). To what extent mode and mechanism of speciation within the non-haplochromine crater lake cichlids can inform us about mechanisms operating in the Great Lakes radiations therefore remained an open question. To our knowledge, Malinsky et al. (2015) is the only study to show sympatric speciation in a crater lake population of haplochromine cichlids, and this study found patterns of divergence similar to those among sister species in the Great Lake radiations (Seehausen et al. 2008), i.e. divergence mainly in male nuptial coloration associated with habitat. In this paper we summarize our evidence for a second case of crater lake speciation in a haplochromine.

Lake Saka is a young and small crater lake in western Uganda (Fig. 1) that is situated within the catchment of lakes Edward and George on the slopes of the Rwenzori Mountains North of Lake George. The lake formed in a shallow depression around a small explosion crater. The latter is 12 m deep but barely 20 m in diameter, and the water is anoxic in the explosion crater below 2-3 m depth (Mills, 2009). The lake around it is only about 4 m deep with a surface of 0.64 km^2 (Mills, 2009). Lake Saka is part of the Fort Portal volcanic field that is thought to be upper-Pleistocene to Holocene in origin (Nixon & Hornung, 1973) but perhaps just 6000-4000 yrs old (Vinogradov et al., 1978). The paleolimnology of the lake is not well resolved, but work on other crater lakes in western Uganda suggest major droughts as recently as 1500 and 1750 years ago (Russell et al., 2007). The lake is home to a population of haplochromine cichlids with conspicuously bright and polymorphic male nuptial coloration that closely resembles polymorphisms classically associated with speciation in Lake Victoria cichlids (Seehausen & van Alphen, 1999; Seehausen & Schluter, 2004; Seehausen et al., 2008). These include two common morphs, one with bright metallic blue males and one with yellow males that have a bright red-chest, and a third but rare morph that is yellow with an orange dorsum. Lake Saka is so small (approx. 1.3 x 0.4 km) and shallow (average 3.4 m, 12 m max) that ongoing speciation would have to be sympatric. We investigated ecology, morphology, population genetics, and phylogeography of the two common colour morphs of Lake Saka cichlids and of cichlids from two sites in the nearby Mpanga river to evaluate if the diversity of the cichlids in Lake Saka might result from sympatric speciation. First, using two mitochondrial DNA markers, we examined whether the colour morphs of Lake Saka constitute a monophyletic mitochondrial lineage within the haplochromine cichlid fishes of the Lake Victoria region superflock (Verheyen et al., 2003; Meier et

al., 2017a); and we examined their relationship to the Mpanga River cichlids. Second, we assessed genetic and morphological divergence between cichlids of Lake Saka and cichlids inhabiting the Mpanga River. Third, we evaluated genetic, ecological, and morphological divergence among the sympatric colour morphs within Lake Saka, using mtDNA, microsatellite DNA, sequences of the LWS opsin gene, morphometric data, and habitat use data. Finally we tested whether spawning time allochrony could explain reproductive isolation among the sympatric colour morphs.

MATERIALS & METHODS

Morphological analyses

Specimens for morphological analysis were collected in 2003 by OS from the crater lake Saka (N=145) and the Mpanga River (N=11; see Fig. 1) and in 2000 by LJC, CAC and OS (N: 83 blue and 103 yellow males). Fishing in the lake was done using gill nets and minnow traps (baited with bread) and the Mpanga River was fished by small seine net and minnow traps. Fish were euthanized immediately after capture with MS-222 and fixed in 10% buffered formalin until manipulation. Lake fish were sexed and males were classified according to their nuptial colour as either yellow with bright red chest or bright blue (Fig. 1). Males of the blue morph varied from bright metallic blue to blue with bright red chest while males of the yellow morph varied from yellow, yellow with bright red chest to yellow orange. Established morphometric distances that capture subtle eco-morphological variation among haplochromine cichlids were collected on preserved specimens using a digital calliper. These included: standard length, body depth, head length, head width, lower jaw length, lower jaw width, snout length, snout width, eye length, inter orbital width, and cheek depth (see Barel et al., 1977). For the analysis, we pooled the 18 blue males and the 41 yellow males captured in 2003 with fish captured in 2000. In total we has 83 blue and 103 yellow males. For 20 blue males and 19 yellow males captured in 2000, snout length, snout width, head width and cheek depth were not measured but the number and size of egg dummies on the anal fin of the males were recorded. For 45 blue males and 43 yellow males, also caught in 2000, eye depth and pre-orbital depth were recorded in addition to the distances described above (Table 1).

A multivariate analysis of covariance (MANCOVA) was used to assess the overall morphological differentiation between river and lake fish as well as between male colour morphs of the lake population. Sex, standard length and fish origin (lake vs. river) were included sequentially in the model. To compare colour morphs, the model included respectively the sampling event, standard length and colour morph. Residuals of each response variable were visually checked for normality and heteroscedasticity. All morphological distances were further analysed separately. Each distance was regressed against standard length or against standard length in interaction with sampling event

for datasets with more than one sampling event to correct for size heterogeneity among individuals. Standardized residuals from these regressions as well as standard length were used as response variables for individual morphological analyses. Due to heteroscedasticity across lake and river fish in standard length (Fligner-Killeen test: χ^2_1= 27.32, P<0.001), a Kruskal-Wallis rank sum test (hereafter KW test) was used to compare standard length between lake and river fish. The size correction based on univariate regressions resolved the problem of heterogeneity of variance across groups for all other morphometric distances (Fligner-Killeen test: all χ^2_1< 3.46, all P>0.06). Variance in standard length between colour morphs within Lake Saka did not deviate from homogeneity (Fligner-Killeen test: χ^2_1= 2.84, P = 0.09). Therefore all morphological variables were analysed using a one-way ANOVA, except for standard length when comparing lake and river fish and for number of egg dummies which were analysed using a KW test. All p-values were corrected for multiple testing with a sequential Bonferroni procedure (Rice, 1989).

Ecological analyses

(1) Stomach content analysis

A subset of fish collected in 2003 was analysed for stomach contents (N_{Stream}=10, N_{Lake}=46; male colour morphs: N_{Blue}=11, N_{Yellow}=14; plus an additional 21 female lake fish). A few yellow males with orange-dorsum were pooled with all others yellow morphs for the ecological analyses. All these fish were collected on a single day between 12 am and 3 pm. Stomach contents were placed in a petri dish and examined under a dissection microscope at Makerere Biological Field Station on the same day. Stomach fullness was assessed using a 5-point scale ranging from 1 (empty) to 5 (full). When stomachs were empty, intestines were dissected to identify their contents using the same procedure as for stomachs. Eight categories of food items were identified: filamentous algae, planktonic green algae, planktonic blue-green algae, diatoms, zooplankton, macrophytes, insects, and miscellaneous (e.g. fish larvae, sand or fungus). Stomach fullness was compared across groups using a generalized linear model with Poisson distribution and fish origin (lake or stream) or colour morph as an explanatory variable. The volumetric percentage of each item contained in the stomach (or intestine) was determined using the points method (Hyslop, 1980) and analysed using a generalized linear mixed model (GLMM) with binomial distribution. Fish origin (lake or stream) or colour morph, food items and their interaction were included as fixed effects and individual identity as a random effect. The significance of diet difference between river and lake fish was assessed by testing the interaction between food items and fish origin (lake or river) with a likelihood ratio test comparing models (with and without interaction) fitted using maximum likelihood. The food items that differed between fish groups were determined using z-tests associated with fixed effect parameters of the model fitted with restricted maximum likelihood. Fish with stomach parasites and/or heavily digested food items

that precluded identification were excluded from diet analysis producing comparisons between 10 river and 36 lake fish respectively. A permutational MANOVA provided similar results suggesting that the use of GLMM did not lead to inflated type II errors despite our low sample size. However due to difference in the diurnal feeding cycle (see results), unevenly affecting the accuracy with which we could identify dietary items, comparison of diet between color morphs was not performed.

(2) Habitat segregation

To test for possible habitat segregation between male nuptial colour morphs, three different inshore habitats were identified within Lake Saka based on the main vegetation (*Cladium, Potamogeton, emergent Phragmites*) and the abundance of colour morphs within these three inshore habitats and the offshore open water habitat was quantified in May 2000. Over three days, 30m long benthic gill nets (four panels: 25.4 mm, 50.8 mm, 76.2 mm, 101.6 mm stretched mesh, and 1.5 m in depth) were set for approximately 1 hour at 14 sites, randomly distributed around the lake. *Cladium* and *Potamogeton* habitats were approximately 0.5 to 1.5 m deep, whereas the depth of emergent *Phragmites* habitat ranged from 1.6 to 2.75 m and that of the open water habitat from 2.75 to 3.35 m depth. A minimum of eight males was collected from each site except for two sites within the *Cladium*-habitat where no mature males were caught. These two sites were excluded from the analyses and, thus we had data for three sites of each of four habitat types. The vertical position of fish in the net (bottom, middle and top) was recorded. The presence and vertical position of a total of 328 males in nuptial colouration were recorded, pooled by habitat type, and analysed using a generalized linear model with Poisson distribution. Differences in habitat use between colour morphs were assessed by testing the three-way interaction among habitat type, vertical position and colour morph while habitats and positions that differed between colour morphs were identified using z-tests associated with model parameters. Females could not be included in this analysis because there is no way to assign them to colour morph (or species, see results).

(3) Spawning seasonality

Between November 1998 and September 1999, fish were sampled approximately monthly, euthanized by emersion in buffered MS 222 and preserved in formalin. In the laboratory, they were transferred to 70% ethanol, dissected, and their gonads examined to determine the stage of maturity under a dissecting microscope. Stages of maturation were classified as follows: I, immature; II, developing; III, maturation; IV, ripe; V, spawning (running); VI, spent (Seehausen et al., 1998). Mean gonad stage and proportion of reproductively active fish (stages IV and V) were analysed using a general linear model and a generalized linear model with binomial distribution where month and colour morph were included as categorical explanatory variables, respectively. A difference in

spawning seasonality between colour morphs was assessed by testing the interaction between month and colour morph, while differences between colour morphs by month were determined using t/z-tests associated with model parameters. Samples with less than five males within one morph were excluded from analyses. Overall, the gonad stage of 107 blue males and 78 yellow red chest males was assessed for a total of seven months. Females could not be included in this analysis because there is no way to assign them to colour morph. All statistical analyses were done in R 3.0.2.

Molecular analyses

(1) Samples and DNA extraction.

Adult *haplochromines* from Lake Saka and the Mpanga River used for molecular genetic analyses are a subsample of the collections made in 2000 and 2003 (see morphology and stomach content analysis). Additional samples were collected by LJC and CAC from River Mpanga in 2008 for population genetic analyses (Fig. 1). Samples from other lakes in the region for phylogenetic analyses were collected during several sampling expeditions (see Table S1 for a list of all samples included for mitochondrial sequencing). Fin clips and muscle tissue from each fish were preserved in 100% ethanol for DNA analyses. Total DNA was extracted using the QIAGEN BioSprint (Qiagen, Zug, Switzerland) DNA animal tissue kit on a Qiagen-BioSpring96 robot. DNA concentrations were adjusted to 50 ng/µl.

(2) Mitochondrial DNA sequencing.

Two regions of the mitochondrial genome were sequenced: 973 bp of the control region (D-Loop) using the primers FISHL15926-F 5'-GAG CGC CGG TCT TGT AA-3' and FISH12s-R 5'-TGC GGA GAC TTG CAT GTG TAA G-3' (Kocher et al., 1989) and 1071 bp of the NADH Dehydrogenase Subunit 2 (ND2) using primers ND2Met-F 5'-CAT ACC CCA AAC ATG TTG GT-3' and ND2Trp-R 5'-GTS GST TTT CAC TCC CGC TTA-3' (Kocher et al., 1995). The PCR products were Sanger sequenced on a CEQ 8000 Automated Capillary Sequencer (Beckman Coulter, Switzerland). Sequences were aligned using Sequencher v. 4.9 (Gene Codes Corporation, Ann Arbor, MI USA) and alignments verified by eye. The alignment was collapsed into the representative haplotypes of 1678bp and a haplotype network was constructed in TCS 1.2 (Clement et al., 2000), excluding gaps. To infer the colonization process of the isolated crater Lake Saka, we included in the haplotype network ten representative haplotypes from Lake Victoria, eight from Lake Edward, nine from Lake Albert and four from Lake Kivu, besides the haplotypes from Lake Saka and the Mpanga River and their frequencies.

(3) Microsatellite analyses.

To estimate neutral genetic variation among and differentiation between river and lake fish, and between the male colour morphs in Lake Saka, individuals from Lake Saka and Mpanga River were genotyped at nine microsatellite loci. Additionally we genotyped 12 individuals from Lake Edward at the same loci (N_{Total}=12), representing the four Lake Edward species that most closely resemble phenotypically the colour morphs of Lake Saka (OS personal observation).

The loci, developed for other cichlids, were amplified using two multiplexing PCR reactions with the QIAGEN Multiplex PCR kit (i.e., Ppun5, Ppun7, Ppun17, Ppun21 and Ppun32 then Osu16, Osu19, Osu20 and Tmo5). Detailed marker description and PCR conditions can be found in Magalhaes *et al.* (2010). Fragment length was analysed with an internal size marker of 400-bp (Beckman Coulter) on a CEQ 8000 and scored with GENEMARKER v. 1.75 (SoftGenetics, USA). Overall, 150 individuals (119 Lake Saka, 19 Mpanga River and 12 Lake Edward) were successfully genotyped for the nine microsatellites.

Genotypes were checked for scoring errors using MICRO-CHECKER v. 2.3 (Van Oosterhout et al., 2004). Neutrality of microsatellite markers (excluding individuals from Lake Edward) was tested using BAYESCAN 2.1 using default settings (Foll et al. 2008) and a false discovery rate (FDR) of 0.01. F_{IS}, allelic richness (AR) and gene diversity (GD) were compiled for each group (i.e. crater lake, river and each crater lake colour morph separately) using FSTAT v.2.9.3 (Goudet, 2002). Departure from Hardy-Weinberg equilibrium was calculated on the overall dataset as well as on the two Saka colour morphs in FSTAT. Finally, linkage disequilibrium was tested for all possible pairs of loci in each group using ARLEQUIN 3.5.1.2 (Excoffier & Lischer, 2010) and *P*-values were corrected for multiple testing with a sequential Bonferroni procedure (Rice, 1989). Global F_{ST} between lake and river fish and between colour morphs within Lake Saka were assessed with a locus-by-locus AMOVA. To further infer the genetic relationships among groups, a neighbour joining tree was calculated from Cavalli-Sforza chord distances among groups based on microsatellite allele frequencies. Statistical support for each node of the inferred tree was obtained using a bootstrap procedure with 1000 replicates in PHYLIP 3.695 (Felsenstein 2017). Finally, a factorial correspondence analysis of individual diploid genotypes was performed with GENETIX v. 4.05 (Belkhir *et al.* 1996-2002) to visualize clustering of groups.

(4) LWS opsin gene sequencing.

To test for divergent selection on the visual system, a 346 bp fragment of the long wavelength sensitive (*LWS*) opsin gene containing the variable and informative exons 4-6 was amplified using primers F3 and R4 from Carleton & Kocher (2001). The PCR cycle included an initial 5 min denaturing step at 95°C, followed by 35 cycles of 95°C for 0.5 min, 58°C for 0.5 min and 72°C for 1 min and a final 10 min extension at 72°C. Sanger sequencing was conducted on an ABI 3130xl sequencer (Applied Biosystems, Switzerland). Electropherograms were aligned in BIOEDIT 7.2.5 (Hall, 1999). Five

polymorphic SNPs in exons 4 and 5 that are associated with divergent adaptation between sister species of Lake Victoria cichlids with red and blue nuptial colouration (Seehausen et al., 2008) were used to assign the alleles present in Lake Saka. Identification of heterozygote individuals was based on visual inspection of the shape and the size of peaks. Sixteen blue and 32 yellow red chest Lake Saka males as well as 17 river fish were sequenced. A haplotype network including all major *LWS* haplotypes known from Lake Victoria (n=21 species; Seehausen et al., 2008) and Lake Edward (n=9 species; Meier et al. 2017a) as well as the representative haplotype of *Astatoreochromis alluaudi* (5 individuals), a species that does not belong to the endemic Lake Victoria region species flock, but belongs to an older lineage occurring in the region (including Lake Saka) was constructed with TCS excluding gaps. Differentiation between Lake Saka and river cichlids as well as between the two Lake Saka colour morphs was assessed with an AMOVA in ARLEQUIN.

All raw data as well as sequence alignments are deposited on ZENODO: doi:XXXXXXX.

RESULTS

A monomorphic Lake Saka mitochondrial haplotype originated from an ancient Mpanga River lineage

The haplotype network approach based on two segments of the mitochondrial genome was used to infer genealogical relationships between haplotypes found in Lake Saka and the Mpanga River in relation to haplotypes from members of the Lake Victoria 'superflock' in all other larger lakes of the region. The haplotype network showed that all haplotypes from Lake Victoria formed a single monophyletic group, connected to the lineages of the western rift lakes, Lakes Edward, Albert, and Kivu via five haplotypes that we found in Mpanga River and (just one of them) in Lake Saka (Fig. 2A). Reducing our sequence data set to only 782 bp D-loop allowed us to place our sequences into the larger haplotype network of the Lake Victoria Region superflock (Verheyen et al. 2003, Supplementary Method 1). This revealed that the haplotype of Lake Saka that is also dominant in the Mpanga River is shared with all large lakes in the region and is rather central in the haplotype network of the entire 'superflock' radiation, connecting Lake Victoria to the older western rift lakes, but closer to the haplotypes of the western rift lakes. Interestingly, the single haplotype that we found in Lake Saka was shared with two thirds of the fish from Mpanga River, where additional haplotypes (two based on the 782 bp D-loop and four on the 1678bp segment) were also identified, mostly closely related to the Saka haplotype.

Based on five polymorphic SNPs in exons 4-6 of the *LWS* opsin gene, we identified two major alleles that occurred both in Lake Saka and Mpanga River and were described previously from

cichlids living in Lakes Victoria and Edward (Seehausen et al., 2008). The first corresponds to the so-called "H-type" class of alleles (overlapping with alleles I, II, IV and V in Seehausen et al., 2008) and is part of the larger "class I" of LWS haplotypes (Sugawara et al. 2002; Terai et al. 2002). The other one corresponds to the "A2-type" class of alleles (overlapping with alleles 12, ed2 or yp2) and is part of the "class II" of *LWS* haplotypes (Terai et al. 2002). Using a genomic DNA fragment of 346bp including 13 SNPs, sequenced in 32 species collected in Mpanga River, Lake Saka, Lake Victoria and Lake Edward to build an haplotype network, fish of each lake and those of Mpanga river were split into two main groups corresponding to the *LWS* class I and class II described above (Fig. 2B). Meier et al. (2017a) have shown that these haplotype classes derive from two distantly related haplochromine species, that this polymorphism in the radiation is due to an ancient hybridization event between those species prior to the formation of the Lake Victoria Region Superflock (LVRS), and that a polymorphic hybrid population seeded all lakes in the region including Lake Saka (Meier et al. 2017a). The *LWS* haplotype of *Astatoreochromis alluaudi*, a much older and only distantly related species that occurs in lakes Victoria, Edward and Saka, took a central position between the two LVRS groups in the network.

Comparison between crater lake and river fish

Genetic population structure – Using microsatellites, we found populations from Lake Saka and Mpanga River were well separated from those of Lake Edward in the multilocus genotype space (Fig. 3A). Despite the smaller sample size, our samples from Lake Edward were a lot more diverse than those from Mpanga River and Lake Saka (Fig. 3A). Indeed, 83 private alleles (out of 107 private alleles from the pooled Mpanga River and Lake Saka dataset) occurred only in Lake Edward in our nine markers (i.e., an average of 9.22 private alleles per marker). Out of the 24 private alleles within Mpanga River/Lake Saka when comparing to Lake Edward, seven each were unique to Lake Saka and Mpanga River (ten were shared between Lake Saka and Mpanga River), leading to on average 0.78 private alleles per marker in Lake Saka as well as in Mpanga River when compared to Lake Edward, and 1.56 private alleles per marker for both when comparing Lake Saka with Mpanga River. Overall, 40 alleles were shared between the 12 individuals from Lake Edward and the 138 individuals from Lake Saka and Mpanga River, corresponding to 32.5% of the alleles in Lake Edward and 62.5% of those in Lake Saka/ Mpanga River. The 19 fish from Mpanga River shared 72% of their alleles with Lake Saka fish, and the 119 fish from Lake Saka also shared 72% of their alleles with fish from Mpanga River. Gene diversity approximated 0.758 and 0.684 for Mpanga River and Lake Saka respectively (Wilcoxon Signed Rank Test: $P = 0.02$), with an allelic richness of 5.55 and 4.23 (based on 19 individuals; Wilcoxon Signed Rank Test: $P = 0.008$, Fig. 3D). Therefore fish from Mpanga River were hence genetically more diverse than Lake Saka fish (Fig. 3B). None of the 36 pairs of loci were in

significant linkage disequilibrium after sequential Bonferroni correction and no marker showed a significant pattern of selection in BAYESCAN. Lake Saka fish and Mpanga River fish were significantly genetically differentiated (F_{ST} = 0.033, P =0.001). The neighbour-joining tree supports this genetic differentiation between lake and stream populations (Fig.3E).

Morphology – Although lake and river fish did not significantly differ in standard length (N_{Lake} = 145, N_{River} = 11; KW test: χ^2_1 = 0.39, P = 0.53), sexual size dimorphism differed between lake and river ($N_{Female\ Lake}$ = 76, $N_{Female\ River}$ = 6, $N_{Male\ Lake}$ = 63, $N_{Male\ River}$ = 5; KW test: χ^2_3 = 12.85, P = 0.005). Sexual size dimorphism was significant in Lake Saka, where males were larger than females ($N_{Female\ Lake}$ = 76, $N_{Male\ Lake}$ = 63; Median: $SL_{Female\ Lake}$= 6.10 cm, $SL_{Male\ Lake}$ = 6.20 cm; KW test: χ^2_1 = 4.82, P = 0.03). In contrast, sexual size dimorphism was not significant in the river population, but males tended to be smaller than females ($N_{Female\ River}$ = 6, $N_{Male\ River}$ = 5; Median: $SL_{Female\ River}$= 6.83 cm, $SL_{Male\ River}$ = 4.49 cm; KW test: χ^2_1 = 2.70, P = 0.10). Overall, males in Lake Saka were larger than males in Mpanga River (KW test: χ^2_1 = 7.36, P = 0.007) whereas females did not differ between lake and river (KW test: χ^2_1 = 2.43, P = 0.12). After correction for sex and standard length, lake and river fish were differentiated in shape (MANCOVA: Sex: $F_{20, 286}$ = 3.60, SL: $F_{10, 142}$ = 192.00, fish origin: $F_{10, 142}$ = 10.80, all P<0.001). Out of the 10 eco-morphological distances measured on both lake and river fish, body depth, head width, cheek depth, inter orbital width and eye length were different between lake and river fish after sequential Bonferroni correction (Table 1).

Diet – While stomach fullness did not differ between lake and river fish (dispersion parameter = 0.99, χ^2_1= 0.27, P = 0.60), their diet was strongly differentiated (LR χ^2_6 = 1911.50, P<0.001, Schoener's niche overlap index = 0.36). Filamentous algae, planktonic green algae, and planktonic blue-green algae were found in larger proportions in stomachs and intestines of lake fish (all P<0.002), whereas zooplankton, macrophytes, and insects were found in larger proportions in stomachs and intestines of river fish (all P<0.001).

Comparison between colour morphs within Lake Saka

Habitat segregation – Male colour morphs differed in their distribution over habitats and water depth (Fig. 4; χ^2_6 = 18.58, P = 0.005). Blue males were found more often than yellow-red chest males on the bottom in open water (P = 0.01) whereas yellow-red chest males were found more often than blue males in *Cladium* at intermediate depth (P = 0.01) and in emergent *Phragmites* near the surface (P = 0.05).

Diet – Stomachs of all blue males were empty, whereas only 3 out of 14 yellow-red chest males had an empty stomach ($N_{Blue\ males}$ = 11, $N_{Yellow\ males}$ = 14; dispersion parameter = 0.42, χ^2_1 = 11.94, P < 0.001) suggesting a difference between the morphs in the timing of feeding (all fish for this analysis were collected between 12am and 3pm).

Morphology – Difference in SL between the colour morphs was not influenced by the sampling event ($F_{2,180}$ = 1.89, P = 0.15), although average SL (of both morphs) differed among sampling events ($F_{2,182}$ = 7.53, P <0.001). After correction for sampling event, yellow-red chest males were significantly smaller than blue males (Table 1), but did not differ in multivariate shape (MANCOVA: sampling event: $F_{10, 134}$ = 339.40, P<0.001; SL: $F_{10, 142}$ = 37.10, P<0.001; colour morph: $F_{10, 142}$ = 1.00, P = 0.44). Two eco-morphological distances, body depth and eye length, as well as the size of the second egg dummy differed between colour morphs after accounting for differences in SL (Table 1): blue males tended to have larger eyes, deeper bodies, and a larger second egg dummy than yellow-red chest males. However, none of these differences remained significant after applying a sequential Bonferroni correction.

Genetic population structure – Both colour morphs had similar levels of genetic diversity at microsatellite markers (Fig. 3C; $AR_{\text{Blue males}}$ = 4.55 and $AR_{\text{Yellow-red chested males}}$ = 4.62, P = 0.40 based on 35 individuals; $GD_{\text{Blue males}}$ = 0.665 and $GD_{\text{Yellow-red chested males}}$ = 0.682, P = 0.43). There was no significant linkage disequilibrium between pairs of loci after sequential Bonferroni correction. Colour morphs were differentiated at microsatellite markers (F_{ST} = 0.007, P = 0.02) but not at the *LWS* opsin gene (F_{ST} = -0.006, P = 0.55). Global genetic differentiation was statistically significant but subtle, i.e. we did not see a clear differentiation in the factorial analysis (Fig. 3C). The subtle global genetic differentiation may have resulted from differentiation at two out of nine microsatellite loci; yet on their own these loci are not significant after sequential Bonferroni correction. Finally the frequency of the *LWS* class I (H-type) allele was exactly 0.5 in blue males, it was 0.41 in yellow-red chested males. There was a trend for this to differ between the morphs (Fig. 3D, One side χ^2_1 = 1.89, P = 0.08). Both colour morphs were in Hardy-Weinberg equilibrium at the *LWS* locus (P = 0.34, P = 0.71 for blue and yellow-red chested males, respectively).

Spawning seasonality – The proportions of reproductively active and quiescent fish did not differ between colour morphs over the year (χ^2_6 = 10.42, P = 0.11). The proportion of mature males among blue males caught was lower in September than in other months (all z-values > 2.10, all P < 0.04) while in yellow-red chest males, the proportion of mature males in September was significantly lower only compared to that in March (March: 0.82 vs September: 0.33, z-value = 2.11, P = 0.035). February was the only month where we found the proportion of mature males to differ significantly between colour morphs (Blue morph: 0.80 vs Yellow-red chest morph: 0.40, z-value = 2.10, P = 0.04). A similar pattern was found for mean gonad stage: mean gonad stage in blue males was significantly lower in September than in other months (all t-tests > 4.96, all P < 0.001) while mean gonad stage in yellow-red chest males did not differ significantly between any of our monthly samples (all t-tests <1.78, all P > 0.07). Overall mean gonad stage between colour morphs differed through time ($F_{6,171}$ = 3.02, P = 0.008) but this was due to an effect of the low proportion of mature males among the blue

morph in September and after exclusion of fish caught in September, there was no longer any difference in breeding seasonality between the morphs ($F_{5,148}$ = 1.66, P = 0.15).

DISCUSSION

Two modes of speciation

The most extensive recent species radiations of animals have occurred in one evolutionary lineage of the cichlid fish family, the haplochromines. Indeed haplochromines account for more than 80% of all species of African lake cichlids (Seehausen 2006). The rich literature on speciation and adaptive radiation in haplochromines in the Great Lakes of Africa contrasts with a relative paucity of studies of speciation of the same lineage in small geographically well-confined lakes. We are aware of only four studies dealing with haplochromines in crater lakes (Sato et al. 2003, Samonte et al. 2007, Machado-Schiaffino et al. 2015, Malinsky et al. 2015). With a single exception (Malinsky et al. 2015), no evidence of speciation was reported, and all previous studies that found evidence of sympatric speciation of cichlids within crater lakes dealt with lineages that did not radiate much in the Great Lakes. This raises important questions. The difference between the lineages in their evidence for speciation in crater lakes could be due to study bias or different frequencies of occurrence in crater lakes between these lineages. Alternatively, it is possible that the speciation mechanisms that are important in the haplochromine species radiations require spatial population structure and therefore do not operate in narrow sense sympatry in crater lakes (Kisel and Barraclough 2010).

Here, we studied two modes of divergence in haplochromines in the same young crater lake system: allopatric divergence between the crater lake and the river from which the lake was colonized and sympatric speciation involving male nuptial colour polymorphism within the lake. Given the support for monophyly of the Mpanga River/Lake Saka clade in our data presented here, and the additional support from published genetic and genomic data (Supplementary Figure 1 in Meier et al. 2017a, Verheyen et al. 2003), divergence both between and within these populations can only have begun after Lake Saka formed between 12’000 and 4’000 years ago (Nixon & Hornung, 1973; Vinogradov et al., 1978). Major droughts affected the crater lakes in the region as recently as 1500 and 1750 years ago (Russell et al., 2007). Given the shallow bathymetry of Lake Saka and the tiny size of the deeper explosion crater, it is entirely possible that the lake was dry at this time. Speciation is hence very unlikely to be older than 10’000 years and may have started as recently as just 1500-1700 years ago.

It is impossible to infer whether, and to what extent, speciation has happened when populations are completely allopatric as is the case for the divergence between the Mpanga River and the Lake Saka populations. On the other hand, we can infer with confidence that speciation has progressed to an advanced stage between the colour morphs within Lake Saka. Both morphs are

found all around the lake in full sympatry. Yet we find significant genetic differentiation at neutral markers of gene flow as well as subtle, but significant, ecological and size differences between them. Given the extent of phenotypic and genetic differentiation between the river population and both lake morphs, including traits that matter to mate choice, it seems reasonable to suggest that the allopatric divergence process observed qualifies as incipient speciation too. Laboratory mate choice experiments could confirm or reject this hypothesis in the future (Selz et al. 2016).

Sympatric speciation from an ancient hybrid stock

A prerequisite to infer sympatric speciation is the demonstration of origin from a common ancestor as opposed to sympatric character displacement following secondary contact between populations that had previously diverged in allopatry. Consistent with an origin from a single ancestral population, we found that Lake Saka haplochromines were fixed for a single mitochondrial haplotype (based on a 1678bp segment from D-loop and ND2 regions combined), which was shared between both colour morphs (Fig. 2A). Although this haplotype was otherwise shared exclusively with cichlids of the nearby Mpanga River, our reconstruction of haplotype networks from D-loop alone for which larger sample sizes from other lakes were available revealed that this haplotype is shared with all lakes in the region (supplementary figure S1A). Interestingly, it is the only haplotype known that occurs in all lakes in the region and it takes a central position in the haplotype network of the regional cichlid fish superflock, connecting the haplotype radiations in the western rift lakes with that in Lake Victoria but isolated from the latter by several mutations (Fig. 2A and supplementary figure S1). The cichlids of the Mpanga River were more variable than those in Lake Saka, having four different mitochondrial haplotypes. Given that populations in crater Lake Saka and the Mpanga River are fixed or nearly fixed respectively for a mitochondrial haplotype that is shared between all lakes in the region and central in the haplotype network implies that the cichlids of Mpanga/Saka may represent a population close to the ancestral population of the Lake Victoria Region Superflock. The LVRS has evolved from an ancient hybrid population but fixed just one of the parental mitochondrial lineages, the Congolese lineage (Meier et al. 2017a). Consistent with an origin of the Lake Saka/Mpanga cichlids from that same hybrid population, and much like in the radiations of lakes Victoria and Edward, we find two anciently divergent haplotypes at the long wavelength sensitive opsin gene (*LWS*) within Lake Saka/Mpanga, each of which is close to one of the haplotypes of either the Congolese or the Upper Nile lineage (Fig. 2C). Earlier work on the LVRS had already shown that the cichlids of Lake Saka share the same genomic admixture proportions between these two lineages as the cichlids of all Albertine rift lakes (Albert, Edward, Kivu; Meier et al., 2017a). That earlier study had also revealed that Lake Saka/Mpanga forms a genomically monophyletic group within the superflock, most closely related to some Lake Edward species.

Our microsatellite data showed little overlap of alleles between Lake Saka or Mpanga River and species from Lake Edward (Fig. 3A). We had chosen to include individuals of those Lake Edward species that in our earlier study (Meier et al. 2017a) appeared the most closely related to Saka/Mpanga haplochromines. The clear differentiation that we observed suggests that Lake Saka and the Mpanga River did not receive much recent gene flow from the larger lakes in the region. Several lines of evidence support a recent colonization of Lake Saka and subsequent isolation from the nearby Mpanga River: i) The fixation of a single mitochondrial haplotype in Lake Saka that takes a central position in the haplotype network of the entire LVRS, and is also the most common haplotype in the Mpanga River. ii) The low genetic diversity at microsatellite markers (Fig. 3A). iii) The close phylogenetic relationship to Mpanga River haplochromines, based on microsatellites. iv) And lastly, the monophyly of the Lake Saka morphs to the exclusion of Mpanga River populations supported by our microsatellite allele frequency-based neighbour-joining tree (Fig. 3D).

Different modes of speciation are associated with different phenotypic dimensions of divergence

Clear differences in eco-morphology and diet support the hypothesis that divergent ecological selection initiated or is driving divergence between Lake Saka and Mpanga River fish, as has also been suggested for other crater lakes in the Uganda region (Machado-Schiaffino et al. 2015). Interestingly the extent of sexual dimorphism in size and colouration was very different between Lake Saka and Mpanga River suggesting that sexual selection may play a role in the divergence between lake and stream populations, but more importantly that sexual selection is much stronger in the lake than the river. Lake males grow larger than females and show dramatic bright nuptial coloration, whereas females are cryptic light brownish. River males in contrast tend to stay smaller than river females and show only muted nuptial coloration (Figure 1). This is the first case we know of, where different prevalence of evidence for sexual selection has been shown between direct sister taxa of cichlids that occupy riverine versus lacustrine habitat.

The difference in the hue of male nuptial coloration was the most striking difference between the sympatric incipient species within Lake Saka, suggesting an important role of divergent sexual selection in this intra-lacustrine diversification. In the large Lake Victoria cichlid radiation, those closely related species that have fully sympatric distribution ranges differ more often than others dramatically in male nuptial coloration with either yellow-red or blue males. Sexual selection on yellow-red/blue male colour variation had, therefore been proposed to be involved in sympatric speciation in Lake Victoria (Seehausen & van Alphen, 1999; Seehausen & Schluter, 2004; Seehausen et al., 2008; Meier et al., 2017a). However, in a large lake it is very difficult to rule out past periods of spatial isolation. Because such periods between the morphs in tiny and recent crater Lake Saka can

effectively be ruled out, our Lake Saka data are consistent with the hypothesis of truly sympatric speciation involving divergent sexual selection on yellow-red/blue male breeding colouration.

Sexual selection often interacts with ecology either because divergent sexual selection is mediated by differences in habitats (Endler and Basolo 1998; Boughman 2002; Seehausen et al. 2008), because both sexual and ecological selection tend to be divergent between the same habitats (Boughman 2001; van Rijssel et al. 2018), or because sexual selection targets different indicator traits of ecological performance in different habitats (Maan and Seehausen 2011). The differences in habitat use and the subtle morphological differences between males of the two colour morphs in Lake Saka suggest that this is the case in this system. The blue morph was significantly associated with more open water and slightly deeper habitat, whereas the yellow-red morph dominated shallower habitat with macrophyte cover. It further appeared that morphs were differentiated in the diurnal feeding rhythm, with yellow-red morphs having freshly filled stomachs during mid-day at a time when most blue males had empty stomachs. Blue males tended to have larger eyes, deeper bodies and a larger second egg dummy than yellow males (Table 1). Although these differences were subtle, and significance was lost after sequential Bonferroni correction, the direction of differences would be consistent with a pattern of adaptation. Larger eyes and larger egg dummies are typical adaptations to living in deeper waters in Lake Victoria (Goldschmidt et al., 1990).

Speciation by selection on polymorphic male nuptial coloration

The kind of colour polymorphism that characterizes the Lake Saka cichlids is widespread among the cichlids of Lake Victoria (Seehausen et al. 1999 a&b), and is likely to have a relatively simple genetic basis but is not a simple Mendelian trait (Magalhaes and Seehausen 2010). It is often divergently resolved and fixed between sister species during speciation (Seehausen and Schluter 2004). It is also widespread among the cichlids of lakes Edward and Kivu (Seehausen pers. obs). Future population genomic work will need to address the question if the nature of the yellow-red/blue polymorphism is due to recurrent mutation or an ancient genetic polymorphism.

In Lake Victoria, species divergence into a species with yellow-red and one with blue male breeding dress is often associated with correlated divergence at the LWS opsin gene and both are often associated with divergence between habitats with alternative light conditions (Seehausen et al., 2008; Carleton et al. 2005). LWS class II haplotypes are generally associated with relatively more red shifted, turbid and/or deep water conditions, whereas class I haplotypes can be found in a range of different light environments (Meier et al. 2017a). In Lake Saka we found both colour morphs to be polymorphic for both haplotype classes, a situation that is uncommon among Lake Victoria cichlids where most populations have been shown to be fixed for one or the other (Terai et al. 2002; Terai et

al. 2006; Seehausen et al. 2008). Both haplotype classes were present also in the Mpanga River cichlids. It seems therefore likely that the LWS polymorphism was present in the founding population of Lake Saka.

Contrary to work on speciation in the Lake Victoria cichlid genus *Pundamilia* (Seehausen et al., 2008), we found no associations between LWS haplotype class and blue versus yellow-red nuptial colouration in Lake Saka. We take this as suggesting that incipient speciation between yellow-red and blue colour morphs is possible without LWS-sequence divergence. Reproductive isolation and neutral genetic differentiation did not seem to be explained by sampling site, and hence likely not by spawning site, nor by spawning time segregation. It seems therefore likely that behavioural mating preferences are present but are not mediated by sequence variation at the LWS opsin gene. Such mating preferences may be mediated by differences in the sequences of other opsin genes, in the expression of opsin genes, and/or by divergence at other mating preference genes. In theory, reproductive isolation between yellow-red and blue male nuptial colour morphs could also be due to strong disruptive ecological selection without strongly divergent female mate preferences (vanDoorn et al. 2009), but this seems very unlikely given that only very subtle ecological differences between the morphs were found.

Besides female mate choice, yellow-red/blue male nuptial colour polymorphisms in haplochromine cichlids can be under disruptive sexual selection by male-male competition. The fitness consequences of the latter can be negatively frequency dependent (Dijkstra et al., 2007), thereby promoting the maintenance of colour variation (Seehausen & Schluter 2004; Dijkstra et al. 2010). Aggression biases towards the most frequent male type may facilitate the initial establishment of novel colour phenotypes and aggression bias towards their own phenotype may promote intraspecific polymorphism as well as coexistence among reproductively isolated species (Seehausen & Schluter, 2004; van Doorn et al., 2004; Dijkstra et al., 2007).

Albeit some of our sample sizes are small, we describe a promising new model system for sympatric speciation in haplochromines. Importantly, the subtle genetic, morphological and ecological differentiation between the Lake Saka incipient species would not have been detectable without good record of live coloration of each individual and a priori knowledge of the colour morphs. Further work is now needed to provide genomic insights into the demography and genome-wide signature of speciation in this system as well as behavioural experiments to determine the degree of assortative mating by direct mate choice and the phenotypic basis for divergent mating preferences.

Conclusions

In conclusion, we described a case of likely sympatric speciation involving divergent sexual selection on male nuptial coloration in a population of haplochromine cichlid fish in a small crater lake. We also described allopatric divergence between the crater lake species and the closely related river population. Both divergence events are age-constrained by the geological history of the crater lake and may be as recent as 1'500 years, but are very unlikely to be older than 10'000 years old. The phenotypic dimensions of divergence are completely different. Not constrained by gene flow, allopatric divergence involves many different morphological traits and the degree of expression of nuptial coloration (muted in the river, but dramatic in the lake cichlids), in addition to strong neutral genetic differentiation. On the contrary, sympatric divergence within the crater lake is associated with dramatic differences in male breeding colouration, but only subtle differences in ecology and morphology and shallow neutral genetic differentiation.

ACKNOWLEDGMENTS

We thank Jackson Efitre and the dedicated field assistants of the Kibale Fish Project for help with fieldwork, the students of the Tropical Biology Association Kibale Course of 2003 for help with stomach content analyses, and Rachel Tongue for help with morphometric distances. The genetic lab work was supported by Swiss National Science Foundation grants 3100A0-118293/1 and 31003A_163338 to OS. Field work conducted by LJC was supported by Funding for this research was provided from the National Science Foundation and the Wildlife Conservation Society. Permission to conduct research in Uganda was acquired from the Uganda National Council for Science and Technology.

621 **References**

622 Allender, C. J., O. Seehausen, M. E. Knight, G. F. Turner, & N. Maclean. 2003. Divergent selection
623 during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration.
624 Proceedings of the National Academy of Sciences of the United States of America 100:14074-14079.

625 Barel, C. D. N., M. van Oijen, F. Witte, & E. Witte-Maas, 1977. An introduction to taxonomy and
626 morphology of haplochromine Cichlidae from Lake Victoria. Netherlands Journal of Zoology 27: 333-
627 389.

628 Barluenga, M., K. N. Stölting, W. Salzburger, M. Muschick, & A. Meyer, 2006. Sympatric speciation in
629 Nicaraguan crater lake cichlid fish. Nature 439: 719–723.

630 Bolnick, D. I., & B. M. Fitzpatrick, 2007. Sympatric speciation: Models and empirical evidence. Annual
631 Review of Ecology Evolution and Systematics 38: 459–487.

632 Boughman, J. W. 2001. Divergent sexual selection enhances reproductive isolation in sticklebacks.
633 Nature 411:944-948.

634 Boughman, J. W. 2002. How sensory drive can promote speciation. Trends in Ecology & Evolution
635 17:571-577.

636 Butlin, R. K., J. Galindo, & J. W. Grahame, 2008. Sympatric, parapatric or allopatric: the most
637 important way to classify speciation? Philosophical Transactions of the Royal Society of London Series
638 B, Biological Sciences 363: 2997–3007.

639 Carleton, K. L., & T. D. Kocher, 2001. Cone opsin genes of african cichlid fishes: tuning spectral
640 sensitivity by differential gene expression. Molecular Biology and Evolution 18: 1540–1550.

641 Carleton, K. L., J. W. L. Parry, J. K. Bowmaker, D. M. Hunt, & O. Seehausen, 2005. Colour vision and
642 speciation in Lake Victoria cichlids of the genus *Pundamilia*. Molecular Ecology 14: 4341–4353.

643 Clement, M., D. Posada, & K. A. Crandall, 2000. TCS: a computer program to estimate gene
644 genealogies. Molecular Ecology 9: 1657–1659.

645 Coyne, J. A., & H. A. Orr, 2004. Speciation. Sinauer Associates, Sunderland, MA, USA.

646 Crow, K. D., H. Munehara, & G. Bernardi, 2010. Sympatric speciation in a genus of marine reef fishes.
647 Molecular Ecology 19: 2089–2105.

648 Dijkstra, P. D., O. Seehausen, M. E. R. Pierotti, & T. G. G. Groothuis, 2007. Male-male competition and
649 speciation: aggression bias towards differently coloured rivals varies between stages of speciation in
650 a Lake Victoria cichlid species complex. Journal of Evolutionary Biology 20: 496–502.

651 Dijkstra, P. D., J. Lindstroem, N. B. Metcalfe, C. K. Hemelrijk, M. Brendel, O. Seehausen, and T. G. G.
652 Groothuis. 2010. Frequency-dependent social dominance in a color polymorphic cichlid fish.
653 Evolution 64:2797-2807.

654 Elmer, K. R., T. K. Lehtonen, A. F. Kautt, C. Harrod & A. Meyer, 2010. Rapid sympatric ecological
655 differentiation of crater lake cichlid fishes within historic times. BMC Biology 8: 60.

656 Endler, J. A. and A. L. Basolo. 1998. Sensory ecology, receiver biases and sexual selection. Trends in
657 Ecology & Evolution 13:415-420.

658 Excoffier, L., & H. E. L. Lischer, 2010. Arlequin suite ver 3.5: a new series of programs to perform
659 population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564–567.

660 Feder, J. L., S. M. Flaxman, Egan, S. P., Comeault, A. A. & P. Nosil. 2013. Geographic mode of
661 speciation and genomic divergence. Annual Review of Ecology Evolution and Systematics 44: 4.1-
662 4.25.

663 Foll, M. & O. Gaggiotti, 2008. A genome-scan method to identify selected loci appropriate for both

664 dominant and codominant markers: A Bayesian perspective. Genetics 180: 977-993.

665 Futuyma, D. J., & G. C. Mayer, 1980. Non-allopatric speciation in animals. Systematic Biology 29: 254–
666 271.

667 Goldschmidt, T., F. Witte, & J. de Visser, 1990. Ecological segregation in zooplanktivorous
668 Haplochromine species (Pisces: Cichlidae) from Lake Victoria. Oikos 58: 343–355.

669 Goudet, J., (2002). FSTAT, a program to estimate and test gene diversities and fixation indices
670 (version 2.9.3). http://www.unil.ch/izea/softwares/fstat.html.

671 Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for
672 Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95-98.

673 Hyslop, E. J., 1980. Stomach contents analysis-a review of methods and their application. Journal of
674 Fish Biology 17: 411–429.

675 Kautt, A. F., G. Machado-Schiaffino, J. Torres-Dowdall, & A. Meyer, 2016a. Incipient sympatric
676 speciation in Midas cichlid fish from the youngest and one of the smallest crater lakes in Nicaragua
677 due to differential use of the benthic and limnetic habitats? Ecology and Evolution 6: 5342–5357.

678 Kautt, A. F., G. Machado-Schiaffino, and A. Meyer. 2016b. Multispecies outcomes of sympatric
679 speciation after admixture with the source population in two radiations of Nicaraguan crater lake
680 Cichlids. Plos Genetics 12: e1006157.

681 Kisel, Y., & T. G. Barraclough, 2010. Speciation has a spatial scale that depends on levels of gene flow.
682 American Naturalist 175: 316–334.

683 Kocher, T. D., 2004. Adaptive evolution and explosive speciation: the cichlid fish model. Nature
684 Reviews Genetics 5: 288–298.

685 Kocher, T. D., J. A. Conroy, K. R. McKaye, J. R. Stauffer, & S. F. Lockwood, 1995. Evolution of NADH
686 dehydrogenase subunit 2 in east African cichlid fish. Molecular Phylogenetics and Evolution 4: 420–
687 432.

688 Kocher, T. D., W. K. Thomas, A. Meyer, S. V. Edwards, S. Pääbo, F. X. Villablanca, & A. C. Wilson, 1989.
689 Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved
690 primers. Proceedings of the National Academy of Sciences of the United States of America 86: 6196–
691 6200.

692 Maan, M. E. and O. Seehausen. 2011. Ecology, sexual selection and speciation. Ecology Letters
693 14:591-602.

694 Machado-Schiaffino, G., Kautt A. F., Kusche, H. & A. Meyer, 2015. Parallel evolution in Ugandan
695 crater lakes: repeated evolution of limnetic body shapes in haplochromine cichlid fish. BMC
696 Evolutionary Biology 15:9.

697 Magalhaes, I. S. and O. Seehausen. 2010. Genetics of male nuptial colour divergence between
698 sympatric sister species of a Lake Victoria cichlid fish. Journal of Evolutionary Biology 23:914–924.

699 Magalhaes, I. S., S. Mwaiko, & O. Seehausen, 2010. Sympatric colour polymorphisms associated with
700 nonrandom gene flow in cichlid fish of Lake Victoria. Molecular Ecology 19: 3285–3300.

701 Martin, C. H., J. S. Cutler, J. P. Friel, C. D. Touokong, G. Coop, and P. C. Wainwright. 2015. Complex
702 histories of repeated gene flow in Cameroon crater lake cichlids cast doubt on one of the clearest
703 examples of sympatric speciation. Evolution 69:1406-1422.

704 Malinsky, M., R. J. Challis, A. M. Tyers, S. Schiffels, Y. Terai, B. P. Ngatunga, E. A. Miska, R. Durbin, M.
705 J. Genner, & G. F. Turner, 2015. Genomic islands of speciation separate cichlid ecomorphs in an East
706 African crater lake. Science 350: 1493–1498.

707 Meier, J. I., D. A. Marques, S. Mwaiko, C. E. Wagner, L. Excoffier, & O. Seehausen, 2017a. Ancient

hybridization fuels rapid cichlid fish adaptive radiations. Nature Communications 8: 14363.

Meier, J. I., V. C. Sousa, D. A. Marques, O. M. Selz, C. E. Wagner, L. Excoffier and O. Seehausen, 2017b. Demographic modelling with whole-genome data reveals parallel origin of similar *Pundamilia* cichlid species after hybridization. Molecular Ecology 26: 123-141.

Mills, K., 2009. Ugandan crater lakes: limnology, palaeolimnology and palaeoenvironmental history. PhD Thesis, Loughborough University, UK.

Moser, F. N., J. C. van Rijssel, S. Mwaiko, J. I. Meier, B. Ngatunga and O. Seehausen, 2018. The onset of ecological diversification 50 years after colonization of a crater lake by haplochromine cichlid fishes. Proceedings of the Royal Society B-Biological Sciences 285: 20180171.

Muschick, M., A. Indermaur, & W. Salzburger, 2012. Convergent evolution within an adaptive radiation of cichlid fishes. Current Biology 22: 2362–2368.

Nixon, P. H., & G. Hornung, 1973. The carbonatite lavas and tuffs near Fort Portal, Western Uganda. Overseas Geological and Mineral Research 41: 168–179.

Nosil, P., 2008. Speciation with gene flow could be common. Molecular Ecology 17: 2103–2106.

Nosil, P., 2012. Ecological Speciation. Oxford University Press, Oxford, UK.

Rice, W. R., 1989. Analyzing tables of statistical tests. Evolution 43: 223–225.

Russell, J. M., D. Verschuren, & H. Eggermont, 2007. Spatial complexity of “Little Ice Age”climate in East Africa: sedimentary recordsfrom two crater lake basins in western Uganda. The Holocene 17: 183–193.

Samonte, I.E., Satta, Y., Sato, A., Tichy, H., Takahata, N. & J. Klein, 2007. Gene flow between species of Lake Victoria Haplochromine fishes. Molecular Biology and Evolution 24: 2069-2080.

Santini, F., M. P. Miglietta, & A. Faucci, 2012. Speciation: where are we now? Evolutionary Biology 39: 141–147.

Sato, A., Takezaki, N., Tichy, H., Figueroa, F., Mayer, W. E. & J. Klein, 2003. Origin and speciation of haplochromine fishes in East African crater lakes investigated by the analysis of their mtDNA, Mhc Genes, and SINEs. Molecular Biology and Evolution 20: 1448-1462.

Schliewen, U. K., D. Tautz, & S. Pääbo, 1994. Sympatric speciation suggested by monophyly of crater lake cichlids. Nature 368: 629–632.

Schliewen, U., K. Rassmann, M. Markmann, J. Markert, T. Kocher, and D. Tautz. 2001. Genetic and ecological divergence of a monophyletic cichlid species pair under fully sympatric conditions in Lake Ejagham, Cameroon. Molecular Ecology 10:1471-1488.

Seehausen, O., 2006. African cichlid fish: a model system in adaptive radiation research. Proceedings of the Royal Society B-Biological Sciences 273: 1987–1998.

Seehausen, O., & D. Schluter, 2004. Male-male competition and nuptial-colour displacement as a diversifying force in Lake Victoria cichlid fishes. Proceedings of the Royal Society B-Biological Sciences 271: 1345–1353.

Seehausen, O., & J. J. M. van Alphen, 1999a. Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? Ecology Letters 2: 262–271.

Seehausen, O., J. J. M. Van Alpen, and F. Witte. 1999b. Can ancient colour polymorphisms explain why some cichlid lineages speciate rapidly under disruptive sexual selection? Belgian Journal of Zoology 129:43-60.

Seehausen, O., F. Witte, J. J. M. van Alphen, & N. Bouton, 1998. Direct mate choice maintains diversity among sympatric cichlids in Lake Victoria. Journal of Fish Biology 53: 37–55.

751 Seehausen, O., R. K. Butlin, I. Keller, C. E. Wagner, J. W. Boughman, P. A. Hohenlohe, C. L. Peichel, G.-
752 P. Saetre, C. Bank, Å. Brännström, A. Brelsford, C. S. Clarkson, F. Eroukhmanoff, J. L. Feder, M. C.
753 Fischer, A. D. Foote, P. Franchini, C. D. Jiggins, F. C. Jones, A. K. Lindholm, K. Lucek, M. E. Maan, D. A.
754 Marques, S. H. Martin, B. Matthews, J. I. Meier, M. Möst, M. W. Nachman, E. Nonaka, D. J. Rennison,
755 J. Schwarzer, E. T. Watson, A. M. Westram, & A. Widmer, 2014. Genomics and the origin of species.
756 Nature Reviews Genetics 15: 176–192.

757 Seehausen, O., Y. Terai, I. S. Magalhaes, K. L. Carleton, H. D. J. Mrosso, R. Miyagi, I. van der Sluijs, M.
758 V. Schneider, M. E. Maan, H. Tachida, H. Imai, & N. Okada, 2008. Speciation through sensory drive in
759 cichlid fish. Nature 455: 620–626.

760 Selz, O. M., R. Thommen, M. E. R. Pierotti, J. M. Anaya-Rojas, and O. Seehausen. 2016. Differences in
761 male coloration are predicted by divergent sexual selection between populations of a cichlid fish.
762 Proceedings of the Royal Society B-Biological Sciences 283.

763 Sugawara, T., Y. Terai, and N. Okada. 2002. Natural selection of the rhodopsin gene during the
764 adaptive radiation of East African Great Lakes Cichlid fishes. Molecular Biology and Evolution
765 19:1807-1811.

766 Terai, Y., O. Seehausen, T. Sasaki, K. Takahashi, S. Mizoiri, T. Sugawara, T. Sato, M. Watanabe, N.
767 Konijnendijk, H. D. J. Mrosso, H. Tachida, H. Imai, Y. Shichida, & N. Okada, 2006. Divergent selection
768 on opsins drives incipient speciation in Lake Victoria cichlids. PLoS Biology 4: e433.

769 Terai, Y., W. E. Mayer, J. Klein, H. Tichy, & N. Okada, 2002. The effect of selection on a long
770 wavelength-sensitive (LWS) opsin gene of Lake Victoria cichlid fishes. Proceedings of the National
771 Academy of Sciences of the United States of America 99: 15501–15506.

772 van Doorn, G. S., U. Dieckmann, & F. J. Weissing, 2004. Sympatric speciation by sexual selection: a
773 critical reevaluation. American Naturalist 163: 709–725.

774 van Doorn, G.S., Edelaar, P. & F. J. Weissing, 2009. On the origin of species by natural and sexual
775 selection. Science 326: 1704-1707.

776 van Rijssel, J. C., F. N. Moser, D. Frei, & O. Seehausen. 2018. Prevalence of disruptive selection
777 predicts extent of species differentiation in Lake Victoria cichlids. Proceedings of the Royal Society B-
778 Biological Sciences 285:20172630.

779 van Oosterhout, C., W. F. Hutchinson, D. Wills, & P. Shipley, 2004. MICRO-CHECKER: software for
780 identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–
781 538.

782 Verheyen, E., W. Salzburger, J. Snoeks, & A. Meyer, 2003. Origin of the superflock of Cichlid fishes
783 from Lake Victoria, East Africa. Science 300: 325–329.

784 Vinogradov, V. I., A. A. Krasnov, V. N. Kuleshov, & L. D. Sulerzhitskiy, 1978. $^{13}C/^{12}C$, $^{18}O/^{16}O$, and ^{14}C
785 concentrations in the carbonatites of the Kalyango Volcano (East Africa). Izvestiya Akademii Nauk,
786 Seriya Geologicheskaya 6: 33–41.

787 Wagner, C. E., L. J. Harmon, and O. Seehausen. 2012. Ecological opportunity and sexual selection
788 together predict adaptive radiation. Nature 487:366-370.

789 Wagner, C. E., L. J. Harmon, & O. Seehausen, 2014. Cichlid species-area relationships are shaped by
790 adaptive radiations that scale with area. Ecology Letters 17: 583–592.

791

Table 1: Morphological differences between lake and river fish, and between colour morphs. Comparisons between lake and river fish were based on fish caught in 2003 and comparisons between colour morphs were based on samples from 1998-2003 (see main text for details). Sample sizes (N) are indicated for each comparison. Morphological comparisons included measurements of size (SL - standard length), eco-morphological distances (HL – head length, HW – head width, BD – body depth, LJL – lower jaw length, LJW – lower jaw width, SnL – snout length, SnW – snout width, CD – cheek depth, IOW – inter orbital width, EyD – eye diameter, EyL – eye length, POD – preorbital depth; absolute average and standard residuals in brackets; see materials and methods) and egg dummies (number (N eggD), size of the first egg dummy (EggD S1), size of the second egg dummy (EggD S2)). Indicated in bold are significant effects. In bold and italics are effects that remained significant after Bonferroni correction.

	Lake vs River[a]						Male nuptial colour morphs					
	N River	N Lake	River	Lake	*F*	*P*	N Blue	N Yellow	Blue	Yellow	*F*	*P*
SL	11	145	56.48	61.43	0.39	0.53	83	103	63.44	62.13	***11.63***	***<0.001***
HL	11	145	20.71 (0.035)	19.13 (-0.003)	0.01	0.9	83	103	20.58 (-0.006)	20.26 (0.008)	0.01	0.92
HW	11	145	10.33 (-0.793)	9.22 (0.058)	**7.67**	***0.006***	63	84	9.57 (0.089)	9.63 (-0.064)	0.88	0.35
BD	11	145	20.92 (-1.535)	17.92 (0.113)	**33.14**	***<0.001***	83	103	21.65 (0.171)	20.92 (-0.134)	**4.47**	**0.04**
LJL	11	145	7.48 (0.419)	6.99 (-0.031)	2.07	0.15	83	103	7.56 (0.051)	7.35 (-0.040)	0.4	0.52
LJW	11	145	5.19 (-0.562)	4.48 (0.041)	**3.77**	**0.05**	83	103	4.15 (-0.089)	4.37 (0.073)	1.23	0.27
SnL	11	145	6.19 (0.121)	5.71 (-0.01)	0.17	0.68	63	84	5.46 (-0.107)	5.65 (0.079)	1.28	0.26
SnW	11	145	7.19 (-0.619)	6.33 (0.045)	**4.57**	**0.03**	63	84	6.05 (-0.001)	6.29 (-0.001)	0	0.99

CD	11	145	3.77 (1.035)	3.66 (-0.078)	**13.62**	***<0.001***	63	84	4.29 (-0.051)	4.11 (0.04)	0.3	0.53
IOW	11	145	4.92 (1.105)	4.82 (-0.081)	**15.53**	***<0.001***	83	103	4.73 (-0.033)	4.70 (0.026)	0.16	0.69
EyD[b]	-	-	-	-	-	-	45	43	6.79 (0.171)	6.58 (-0.179)	2.73	0.1
EyL	11	145	6.38 (-1.368)	5.18 (0.100)	**25.14**	***<0.001***	83	103	6.74 (0.181)	6.41 (-0.144)	**5.02**	**0.03**
POD[b]	-	-	-	-	-	-	45	43	2.36 (-0.038)	2.35 (0.038)	0.12	0.73
N eggD[c]	-	-	-	-	-	-	20	18	2.1	1.972	1.14	0.28
EggD S1[b]	-	-	-	-	-	-	19	17	0.26	-0.282	2.69	0.11
EggD S2[b]	-	-	-	-	-	-	17	17	0.398	-0.402	**6.19**	**0.02**

a: comparison including males and females

b: comparison based on 1 sampling event.

c: variables analysed with a Kruskall Wallis test due to the violation of normality (N eggD) or homoscedasticity (SL).

Table 2: F_{ST} statistics of the locus-by-locus AMOVA between lake and river fishes and male nuptial colour morphs of haplochromine cichlids of Lake Saka.

	Lake vs. River		Colour morphs	
Locus	F_{ST}	*P*-value	F_{ST}	*P*-value
Tmo5	***0.027***	***0.037***	-0.008	0.802
Osu20	***0.066***	***0.001***	0.009	0.160
Osu16	***0.031***	***0.027***	0.001	0.381
Osu19	0.010	0.166	***0.021***	***0.018***
Ppun32	***0.168***	***0.001***	0.013	0.165
Ppun21	***0.054***	***0.002***	0.005	0.283
Ppun17	0.011	0.154	***0.038***	***0.006***
Ppun7	-0.002	0.485	-0.001	0.489
Ppun5	-0.007	0.684	-0.009	0.874

In bold, significant F_{ST} at $\alpha = 0.05$

Figures

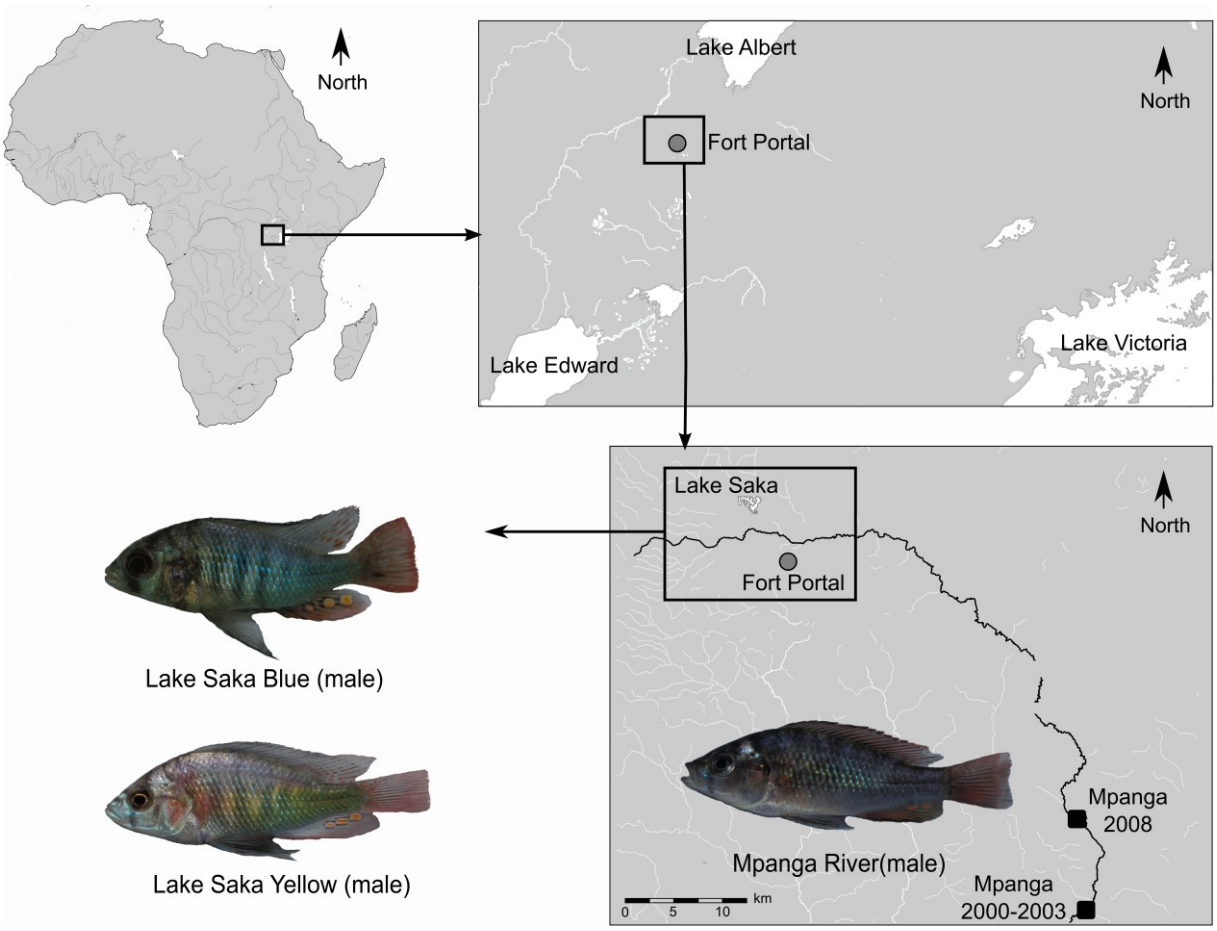

Figure 1. Map of Africa with inlets showing on the location of Lake Saka in relation to the major lakes of the Lake Victoria region. Sampling locations within the Mpanga River are indicated. Two common nuptial colour morphs of males are found in Lake Saka (blue and yellow morphs) whereas river fish did not express colour polymorphism.

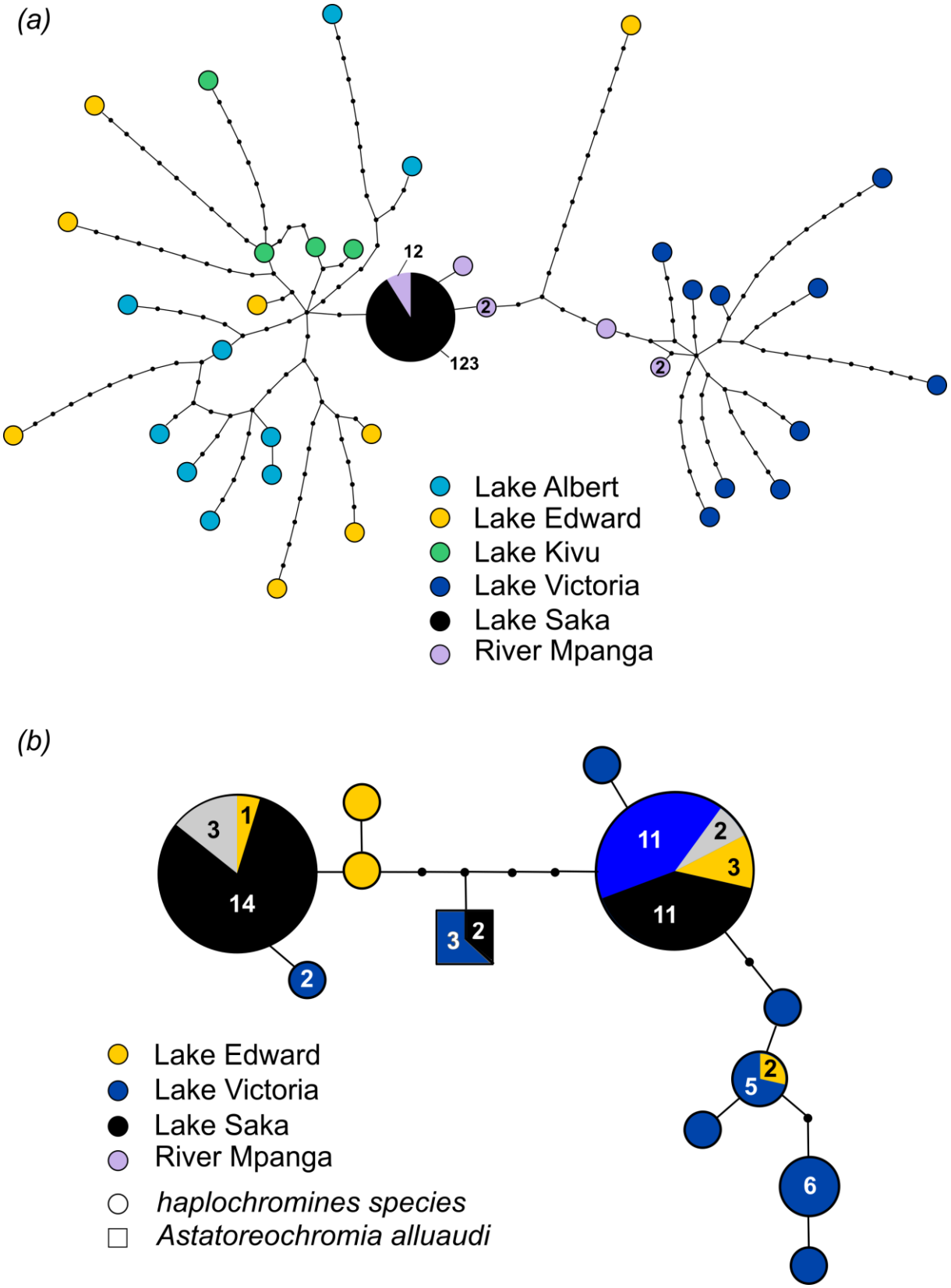

Figure 2. Haplotype networks based on (a) mtDNA (collapsed ND2 and D loop) sequences and (b) LWS opsin gene in the Lake Victoria region. Each dot represents 1 individual except for haplotypes where the number of individuals is indicated.

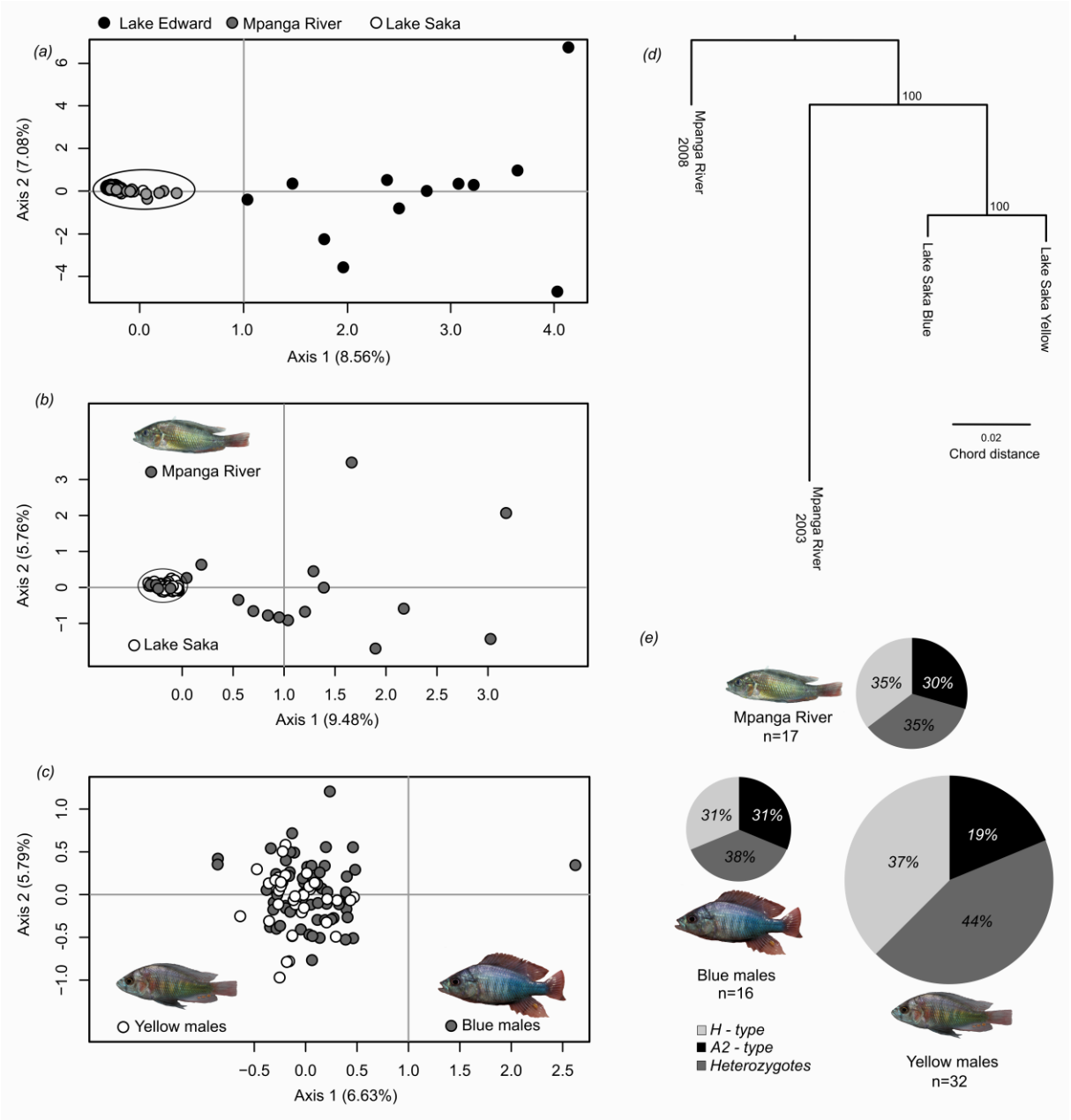

Figure 3. Genetic diversity and differentiation within Lake Saka region. Results of a factorial correspondence analysis of microsatellite diversity for (a) populations from Lake Edward (n = 12), Lake Saka (n = 119) and Mpanga River (n = 19). (b) populations from Lake Saka and Mpanga River. (c) Lake Saka color morphs (35 blue males and 62 yellow males). Circles indicate region of the maps which would be zoomed in. (d) Phylogram showing the genetic relationship among populations based on Cavalli-Sforza Chord distances. Numbers indicate statistical support based on 1000 bootstrap replicates. (e) Allele frequency and heterozygote proportion at the LWS opsin gene in Mpanga River and within Lake Saka by colour morphs.

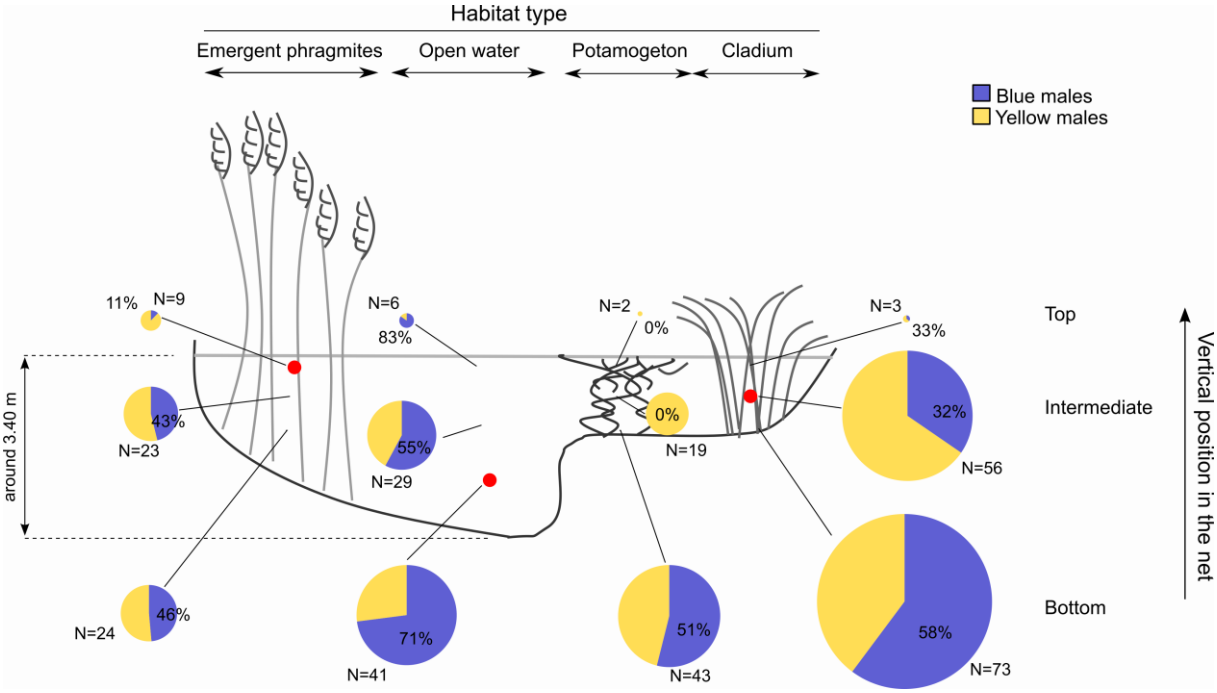

Figure 4. Habitat and depth association between colour morphs. Pies represent the proportions of blue and yellow males in each habitat by vertical position. Proportions are corrected for the total number of blue and yellow males caught (respectively N = 155 and 173). Red dots indicate significant differences (P<0.05) between blues and yellows males.