JOURNAL OF ZOOLOGICAL SYSTEMATICS AND EVOLUTIONARY RESEARCH

Accepted on 17 August 2011
J Zool Syst Evol Res doi: 10.1111/j.1439-0469.2011.00641.x

[1]*Department of Zoology, University of Graz, Graz, Austria;* [2]*Ichthyology Section, National Museums of Kenya, Nairobi, Kenya*

Genetic distinction of four haplochromine cichlid fish species in a satellite lake of Lake Victoria, East Africa

Elizabeth A. Odhiambo[1,2], Selma I. Mautner[1], Oliver Bock[1] and Christian Sturmbauer[1]

Abstract

Lake Victoria is famous for its in evolutionary terms young but species-rich assemblage of cichlid fishes. This 'superflock' also includes additional species from adjacent water systems. Lake Victoria is surrounded by several smaller lakes that are connected to the main water body of Lake Victoria only through swampy areas. Lake Kanyaboli is one such lake, harbouring a much poorer species diversity, mostly comprised of Lake Victoria endemics, some of which are now considered extirpated from the main lake. The focus of this study was on the modern haplochromine component of the cichlid fauna, represented by *Lipochromis maxillaris*, *Astatotilapia nubila*, *Xystichromis phytophagus* and *Astatotilapia* sp. '*Bigeye*', as well as a number of morphologically distinct haplochromine specimens that could not be assigned to any of the recognized species. We used five microsatellite markers to distinguish these five taxa. Genetically, *L. maxillaris* was clearly differentiated from all other taxa, and *A.* sp. '*Bigeye*' was moderately differentiated from the remaining three. *Astatotilapia nubila*, *X. phytophagus* and the unidentified specimens constituted a partially overlapping cluster. As each of the clusters had several (5–14) private alleles, extremely recent divergence is suggested. As all taxa except for *A.* sp. '*Bigeye*' and the unidentified specimens also occur or at least occurred in Lake Victoria, it is likely that they evolved as part of the Lake Victoria superflock, while *A.* sp. '*Bigeye*' and the unidentified specimens may have currently evolved *in situ*. The observation of slightly distinct albeit overlapping body shapes and the extremely close genetic relationship between three of the five taxa are fully compatible and in support of the hybrid swarm theory of adaptive radiation.

Key words: Lake Kanyaboli – hybrid swarm theory – microsatellites – genetic differentiation – speciation

Introduction

Cichlid fishes of the Eastern African Great Lakes, Victoria, Malawi and Tanganyika, form remarkable and fascinating species flocks representing unique examples of vertebrate explosive speciation and adaptive radiation (Fryer and Iles 1972). Each of the three lakes harbours hundreds of cichlid species, most of which are endemic to their particular lake (Turner et al. 2001). In this regard, haplochromine cichlid fishes of Lake Victoria, consisting of at least 500 endemic species, are considered to have evolved most rapidly and as such provide an extraordinary research target because of their fast speciation rates and tremendous phenotypic variation (Nagl et al. 2000; Seehausen 2000). Until the early 1980s, up to 80% of fish biomass from Lake Victoria was attributable to haplochromine cichlids (Ogutu-Ohwayo 1990; Witte 1992) but following the upsurge of *Lates niloticus* (Nile perch), coupled with other anthropogenic factors that percentage decreased to $<2\%$ (Witte et al. 2000) and several species are now considered extinct (Goudswaard et al. 2006). However, some of these haplochromines that were considered to have been extirpated from Lake Victoria were reported to have survived and even thrived in satellite lakes (Kaufman and Ochumba 1993). Satellite lakes were defined by Greenwood (1965) as 'small water bodies isolated from Lake Victoria by papyrus swamps or sand bars', and they harbour some of the haplochromines from Lake Victoria. Haplochromine cichlids of Lake Victoria have been subjects of various molecular phylogenetic and population genetic studies. Much interest has been focused on the origin and age of the Lake Victoria cichlid superflock (Meyer et al. 1990; Nagl et al. 2000; Seehausen 2002; Verheyen et al. 2003; Elmer et al. 2009), especially following paleolimnological evidence by Johnson et al. (1996) that the lake was desiccated as recently as 12.400 ^{14}C-years ago. Even if the actual age of the flock is currently disputed (Fryer 1997, 2001, 2004; Nagl et al. 2000; Verheyen et al. 2003; Elmer et al. 2009), it is clear that the Lake Victoria cichlid species flock either formed within this enormously short period of time (Seehausen 2002) or was derived from a 'superflock' with its ancestral stocks most likely coming from Lake Kivu (Verheyen et al. 2003, Elmer et al. 2009). Moreover, interspecific hybridization as a mechanism to create biological diversity in the form of a hybrid swarm at an early stage of adaptive radiation was suggested as a likely scenario for Lake Victoria (Seehausen 2004). Only few studies have addressed population structure in Lake Victoria or its satellite lakes (Abila et al. 2004; Maeda et al. 2009; Mzighani et al. 2010).

This study focused on modern haplochromines (Salzburger et al. 2002; Verheyen et al. 2003; Koblmüller et al. 2008) in Lake Kanyaboli, a small satellite lake of Lake Victoria covering an area of approximately 11 km^2 and separated from the Lake Victoria by a dense papyrus swamp (Fig. 1 in Odhiambo et al. 2011). With regard to the cichlid fish, Lake Kanyaboli is home to *Oreochromis esculentus* and *Xystichromis phytophagus*, both of which are currently no longer found in Lake Victoria, as well as at least six haplochromine species including *Astatoreochromis alluaudi* and *Pseudocrenilabrus multicolor victoriae*. The latter two have already been shown to be clearly distinct from the remaining haplochromines, all of which are part of the 'modern' haplochromine superflock (Salzburger et al. 2002; Koblmüller et al. 2008). We focused on the four remaining modern haplochromine species, namely *Lipochromis maxillaris*, *X. phytophagus*, *Astatotilapia nubila*,

Corresponding author: Christian Sturmbauer (christian.sturmbauer@uni-graz.at)
Contributing authors: Elizabeth A. Odhiambo (akinyieo@yahoo.com), Selma I. Mautner (selma.mautner@uni-graz.at), Oliver Bock (oliver.bock@uni-graz.at)

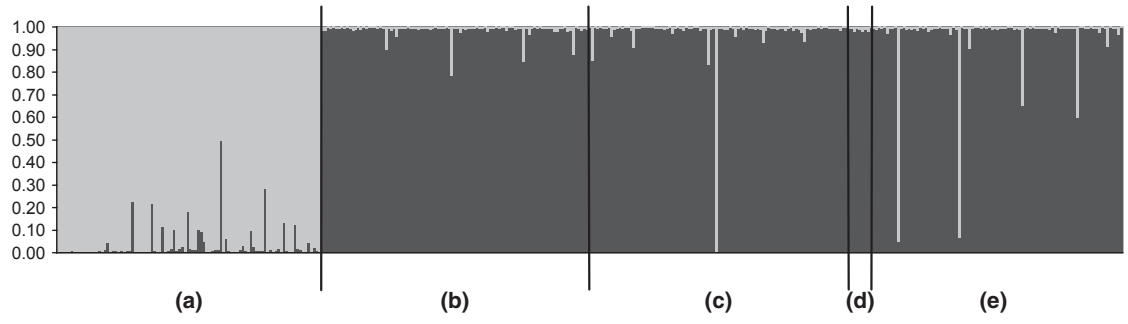

Fig. 1. Results of Bayesian analysis using STRUCTURE. Each individual is represented by a single vertical line and assigned to one of the two clusters. Black bars indicate species identification according to morphology and colour: (a) *Lipochromis maxillaris*, (b) *Astatotilapia nubila*, (c) *Xystichromis phytophagus*, (d) *Astatotilapia* sp. '*Bigeye*', (e) unidentified specimens

the rare and yet undescribed species, *Astatotilapia* sp. '*Bigeye*', plus a group of unidentified specimens that might constitute hybrids or a new haplochromine species. An earlier study using mtDNA sequences of the control region found only low and non-significant F_{ST} values for Lake Kanyaboli haplochromine cichlid species *X. phytophagus* and *A. nubila* (Abila et al. 2008). Previous studies have also used microsatellites in an attempt to study genetic differentiation among Lake Victoria haplochromine species (Abila et al. 2004; Samonte et al. 2007; Elmer et al. 2009; Maeda et al. 2009). In this study, we chose polymorphic microsatellite marker to assess the genetic differentiation of Lake Kanyaboli haplochromines and compare our findings to the morphological differentiation among the same taxa. This study complements a parallel geometric morphometric analysis of the same taxa, which showed clear morphological distinctness of *L. maxillaris* and *A.* sp. '*Bigeye*' but could not clearly separate *X. phytophagus* from *A. nubila* in morphospace, despite their clear differentiation in terms of colour and trophic ecology. Furthermore, the unidentified specimens clustered in close association with *X. phytophagus* and *A. nubila* (Odhiambo et al. 2011). In summary, comparative morphometric analyses showed three morphological clusters, namely *Xystichromis maxillaris*, *A.* sp. 'Bigeye' and the three remaining taxa (*A. nubila*, *Astatotilapia phytophagus* and the unidentified specimens, with a clear tendency of segregation according to the assignment). Moreover, the four recognized species are clearly distinct in male mating colour and in trophic ecology. By complementing the comparative morphological findings with genetic data, we intended to test whether the morphology-based species separation is congruent with genetic differentiation. Specimens that remained unidentified on the basis of morphology were also included to gain further insights into their assignment.

Materials and Methods

Sampling, DNA extraction and amplification

Between May 2007 and April 2008, 291 samples were collected from 33 localities covering the entire lake (approximately 11 km^2) using gill nets, traps and baited hooks. Given the small size of the lake, samples of each species were assumed to comprise a single panmictic population, in agreement with results of Abila et al. (2004). In total, 96 individuals were sampled of *A. nubila* (Boulenger, 1906), 96 of *L. maxillaris* (Trewavas, 1928), 96 of *X. phytophagus* (Greenwood 1965), EIGHT individuals of the rare and yet undescribed species *A.* sp. '*Bigeye*' as well as 91 specimens that could not be assigned to any of the four known species. Fin clips were taken from freshly caught specimens and immediately preserved in absolute ethanol. Voucher specimens were deposited at the National Museums of Kenya, Nairobi, and are listed in Table S1.

DNA extraction, microsatellite PCR and data screening

Total DNA was extracted from the ethanol-preserved fin clips by enzymatic digestion using proteinase K, followed by ammonium acetate and isopropanol precipitation (Sambrook et al. 1989). Samples were screened for variation at each of the five microsatellite loci: TmoM11, TmoM25, TmoM27 (Zardoya et al. 1996), Pzeb3 (van Oppen et al. 1997) and Unh951 (Carleton et al. 2002). PCR was performed under the following conditions: denaturation at 94°C for 180 s, followed by 35 cycles of denaturation at 94°C for 30 s, annealing for 30 s and extension at 72°C for 60 s and a final extension at 72°C for 7 min. The annealing temperature was 50°C for TmoM11, 52°C for TmoM25, 52°C for TmoM27, 55°C for Pzeb3 and 52°C for Unh951. Microsatellites were PCR-amplified in a 10- or 20-μl reaction mix containing 5–10 ng extracted DNA, 0.25 pM of each primer, 0.05 mM dNTP mix, 10 μg μl^{-1} BSA, 1.5 mM $MgCl_2$, 10× Taq polymerase buffer and 0.5 U Taq DNA polymerase. Forward primers were labelled with either FAM or HEX fluorescent dyes. PCR products were analysed on an ABI 3130xl sequencer together with an ABI ROX500 size standard, and allele calling was performed with GENEMAPPER software version 3.7 (Applied Biosystems, Carlsbad, CA, USA).

Data analysis

To detect scoring errors and genotyping artefacts such as null alleles, large allele dropout and stutter bands, we used MICRO-CHECKER (van Oosterhout et al. 2004). ARLEQUIN 3.11 (Excoffier et al. 2005) was used to calculate standard genetic diversity parameters (H_O, H_E, N_A), Hardy–Weinberg equilibrium and linkage disequilibrium, pairwise comparisons among species using allele frequencies (F_{ST}) for populations with high gene flow and sum of squares (R_{ST}) for populations with longer divergence time (Balloux and Lugon-Moulin 2002) and an analysis of molecular variance (AMOVA) among and within populations. Significance levels were determined using the Markov chain method (forecasted chain length = 100 000, dememorization steps = 1000). Rejection levels were calculated with Bonferroni-adjusted p-values. Population structure was inferred with a Bayesian clustering procedure implemented in STRUCTURE 2.2 (Pritchard et al. 2000) to identify the number of clusters of origin for all sampled individuals and to assign individuals to homogeneous clusters. We used an admixture model, assumed correlated allele frequencies and gave no prior species information. We inferred the number of clusters by performing three independent runs for each K (K = 1–7) after a burn-in of 100 000 iterations and a run length of 500 000 iterations followed by the ad hoc statistic ΔK (Evanno et al. 2005). CLUMPP (Jakobsson and Rosenberg 2007) was used to align multiple STRUCTURE runs, and Distruct (Rosenberg 2003) was used to visualize the results. Patterns of genetic differentiation were visualized using a factorial correspondence analysis (FCA) in GENETIX (Belkhir et al. 2004). GENETIX was also used to calculate the number of private alleles

found in each species. To check for a reduction in population size, we used BOTTLENECK (Piry et al. 1999) by comparing expected and observed heterozygosities in each species. Rare alleles are lost more rapidly than common alleles during a population bottleneck causing excess heterozygosity. We used the stepwise mutation model (SMM), the infinite alleles model (IAM) and the two-phased model of mutation (TPM) followed by a one-tailed Wilcoxon signed rank test to assess significance values.

Results

The five microsatellite loci yielded a total of 102 alleles for the 387 samples analysed. Fifty-six alleles were found for *L. maxillaris*, 59 for *A. nubila*, 67 for *X. phytophagus*, 29 in *A.* sp. '*Bigeye*' and 70 in the group of so far unidentified specimens (Table 1). The average number of alleles per species ranged from 5.8 ± 2.3 (*A.* sp. '*Bigeye*' – small sample size) to 13.4 ± 6.1 (Table 1). The average number of private alleles was 14 for *L. maxillaris*, five for *A. nubila*, 10 for *X. phytophagus*, one for *A.* sp. '*Bigeye*' and six for the unidentified specimens (Table 1). For compatibility reasons, the three loci used by Abila et al. (2004) were used (TmoM11, TmoM25 and TmoM27). MICRO-CHECKER indicated the possibility of null alleles at loci TmoM11 and Unh951. All further analyses were repeated excluding these two loci but always yielded similar results to the analyses performed with all five loci as presented in the results.

Discriminating taxa on the basis of genetic signature

When STRUCTURE was run on the total sample (387 individuals at five loci) and $K = 1$–7, with no previous information about species identity, the maximum probability [ln P(X|K)] was found at $K = 2$ which was also supported by the *post hoc* analysis of ΔK, suggesting a split into two distinct genetic clusters. *Lipochromis maxillaris* was assigned to cluster one; all other species were assigned to cluster two (Table 4). Three

Table 1. Levels of genetic diversity for the five sampled taxa

	Lipochromis maxillaris $n = 96$	*Astatotilapia nubila* $n = 96$	*Xystichromis phytophagus* $n = 96$	*Astatotilapia* sp. '*Bigeye*' $n = 8$	Unidentified $n = 91$	Mean N_A per locus
TmoM11 (37)						
N_A	20	17	17	9	20	18.5 ± 1.7
P_A	9	2	3	1	3	
H_O	0.750[1]	0.760[1]	0.740[1]	0.875	0.835[1]	
H_E	0.887	0.873	0.886	0.850	0.881	
Min (bp)	180	193	193	193	178	
Max (bp)	225	229	229	228	229	
TmoM25 (16)						
N_A	12	9	11	5	16	12.0 ± 2.9
P_A	3	1	1	0	1	
H_O	0.615[1]	0.792	0.771	0.750	0.714[1]	
H_E	0.693	0.776	0.741	0.667	0.769	
Min (bp)	364	366	366	368	364	
Max (bp)	386	382	390	382	382	
TmoM27 (13)						
N_A	8	8	10	5	9	8.8 ± 1.0
P_A	1	1	2	0	0	
H_O	0.760	0.594	0.615	0.625	0.593	
H_E	0.793	0.618	0.588	0.858	0.573	
Min (bp)	380	380	380	380	380	
Max (bp)	424	426	436	424	426	
Pzeb3 (8)						
N_A	4	4	7	3	5	5.0 ± 1.4
P_A	1	0	2	0	0	
H_O	0.594	0.573	0.688[1]	0.500	0.604	
H_E	0.626	0.565	0.628	0.808	0.606	
Min (bp)	283	310	310	310	283	
Max (bp)	314	316	320	314	316	
Unh951 (28)						
N_A	12	21	22	7	20	18.8 ± 4.6
P_A	0	1	2	0	2	
H_O	0.604[1]	0.698[1]	0.750[1]	0.500	0.758[1]	
H_E	0.897	0.821	0.849	0.817	0.836	
Min (bp)	209	203	203	229	203	
Max (bp)	257	293	293	257	265	
Mean N_A per spp ± SD	11.2 ± 5.9	11.8 ± 7.0	13.4 ± 6.0	5.8 ± 2.3	14.0 ± 6.8	
Total P_A per spp	14	5	10	1	6	
Mean P_A per spp ± SD	2.8 ± 3.2	1 ± 0.6	2 ± 0.6	n/a	1.2 ± 1.7	

Total number of alleles is included in parentheses for each locus.
Values from *Astatotilapia* sp. '*Bigeye*' were excluded from the mean number of alleles per locus owing to its very small sample number.
N_A, number of alleles; P_A, private alleles; H_O, heterozygosity observed; H_E, heterozygosity expected.
[1]Significant deviation from HWE after Bonferroni correction.

individuals (one *X. phytophagus*, two unidentified specimens) were assigned to the *L. maxillaris* cluster. STRUCTURE assigned the highest posterior probability to a population structure with two ancestral populations. After removing *L. maxillaris* from the analysis, the four remaining entities remained in one cluster owing to the low genetic differentiation among them. Details on the STRUCTURE analysis are given in Table S2. Figure 1 shows the estimated membership coefficients for each individual in each cluster.

Genetic differentiation among taxa

Overall F_{ST} and R_{ST} values among taxa were generally low and ranged from 0 to 0.14 and 0 to 0.30, respectively, among all five sampled entities. The highest pairwise differentiation was always between *L. maxillaris* and the other entities. Despite a low sample size for *A.* sp. '*Bigeye*', this species was also significantly separated from all other taxa. The remaining three taxa turned out to be much more closely related, and the non-significant F_{ST} values corroborated a similarly low degree of genetic separation between *A. nubila* and *X. phytophagus* and the unidentified specimens (Table 2).

AMOVA results showed that 93% of the genetic variation was within taxa and the variation among them was low (7%) but still significant. Following the results obtained from STRUCTURE, AMOVA was repeated taking into account the two clusters found by STRUCTURE. Although 13% of genetic variation was found among the two clusters, this proportion of variance was not significant (Table 3).

FCA analysis showed a clear distinction of *L. maxillaris* and *A.* sp. 'Bigeye'. After removing these species, the remaining three taxa were still overlapping, but a fairly clear distinction of the unidentified individuals was possible (Fig. 2).

Analyses of population bottlenecks found significant evidence for a recent bottleneck using the IAM for *L. maxillaris* ($p = 0.031$) which was no longer significant after removing the two loci that showed a tendency for null alleles. All tests using the other two models SMM and TPM, however, did not reveal an excess of heterozygotes for any of the entities and thus no evidence for a population bottleneck in any of the species. Using only three loci that showed no tendency for null alleles resulted in no indication for a bottleneck in any population under any model.

Discussion

Three of the five taxa analysed here, *L. maxillaris* (Greenwood, 1980), *A. nubila* (Boulenger, 1906) and *X. phytophagus* (Greenwood, 1966), are officially described and valid species on the basis of male nuptial coloration and other morphological and ecological characters. *A.* sp. '*Bigeye*' (Kaufman, 1996) is valid on the same basis but not yet officially described. The fifth taxon referred to as 'unidentified' consisted of specimens that could not be grouped with any of the described and already mentioned four species on the basis of morphology. The recently published geometric morphometric analysis of the same individuals showed clear morphological differences between *L. maxillaris* and *A.* sp. '*Bigeye*' but could not clearly separate *X. phytophagus*, *A. nubila* and the unidentified specimens in morphospace, despite their clear differentiation in terms of colour and feeding ecology (Odhiambo et al. 2011). These findings highlighted the young evolutionary age of the Lake Victoria cichlid species flock and are relevant in the light of Seehausen's hybrid swarm theory (Seehausen 2004).

Intraspecific and interspecific differentiation among the entities

Pairwise genetic differences revealed high levels of differentiation between *L. maxillaris* and all other species as well as moderate levels of differentiation between *A.* sp. '*Bigeye*' and all other entities. The remaining three entities, however, turned out to be extremely closely related with very low levels of genetic differentiation. R_{ST} values were constantly higher than F_{ST} values for the differentiation between *L. maxillaris* and all other species, indicating larger allele size differences. Results from FCA analysis corroborated F_{ST} and R_{ST} levels of differentiation, also clearly separating *L. maxillaris* from all other species and also separating *A.* sp. '*Bigeye*' very clearly from the remaining entities. *Astatotilapia nubila*, *X. phytophagus* and the unidentified individuals appeared as a closely related cluster with a tendency to separate into three subunits. The clear separation of *L. maxillaris* from all other taxa was also found with the clustering analysis in STRUCTURE, whereas the remaining four taxa were all assigned to the second cluster. The assignment of one *X. phytophagus* and two unidentified specimens to the *L. maxillaris* cluster might be due to a morphological assignment error, although introgression cannot be ruled out. Results from FCA show *A.* sp. '*Bigeye*' to be quite different from the other species, but STRUCTURE did not recognize it as a different cluster, most likely due to low sample size of only eight specimens for this species. Results from AMOVA revealed that most of the genetic differentiation (93%) was within the taxa and that there was only very little (7%) differentiation among the taxa.

The observed low level of differentiation among four of the five taxa is in agreement with the finding of Maeda et al. (2009) who found no population (and species) differentiation in two sympatric Lake Victoria species (*Haplochromis pyrrhocephalus* and *Haplochromis laparogramma*) on the basis of a STRUCTURE analysis and on R_{ST} values. Samonte et al. (2007) also found levels of differentiation within species to be similar to those between species. Mzighani et al. (2010) using SINEs also did not find any significant differentiation between or within the species of Lake Victoria haplochromines that they studied. Other studies that used mtDNA (Meyer et al. 1990; Nagl et al. 2000) and SINEs (Terai et al. 2004; Mzighani et al. 2010) have also not been successful in genetically differentiating between the haplochromine species of Lake Victoria. The low level of genetic diversification and incomplete lineage sorting has been repeatedly attributed to the young age of these haplochromines, many of which are still at an incipient stage of evolution (Meyer 1993). In the light of all published data, the clear genetic differentiation of *L. maxillaris* to the remaining entities seems quite remarkable, especially as our analysis is based on five microsatellite loci only.

A second aspect of the close evolutionary relationship between Lake Victoria haplochromines is that they are likely to be still genetically compatible and have not yet crossed the border of postzygotic isolation (Crapon de Caprona and Fritzsch 1984). In fact, most species, if not all, are only reproductively isolated by behaviour (Seehausen and van Alphen 1998), or by allopatry. It has been shown that speciation can be reversed by a breakdown of reproductive isolation among sympatric species – the result of high water

Table 2. Pairwise R_{ST} (below diagonal) and pairwise F_{ST} (above diagonal)

	Lipochromis maxillaris	*Astatotilapia nubila*	*Xystichromis phytophagus*	*Astatotilapia* sp. '*Bigeye*'	Unidentified
L. maxillaris	–	0.140*	0.135*	0.103*	0.136*
A. nubila	0.276*	–	−0.0001	0.054*	0.001
X. phytophagus	0.272*	−0.005	–	0.051*	−0.0002
A. sp. '*Bigeye*'	0.129*	−0.009	−0.011	–	0.053*
Unidentified	0.303*	−0.002	−0.002	0.017	–

*$p < 0.001$.

Table 3. AMOVA results after test for variation (a) among all taxa (b) between the two clusters where *Lipochromis maxillaris* forms one cluster and the other specimens the second cluster (1023 permutations)

Model	Source of variation	df	Sum of squares	Variance components	Percentage variation	p-value
a	Among species	4	93.40	0.15	7.44	< 0.001
	Within species	767	1400.98	1.83	92.56	< 0.001
b	Among clusters	1	84.87	0.28	13.44	ns
	Among species–within clusters	3	19.572	0.01	0.36	< 0.01
	Within species	767	1400.95	1.83	86.2	< 0.001

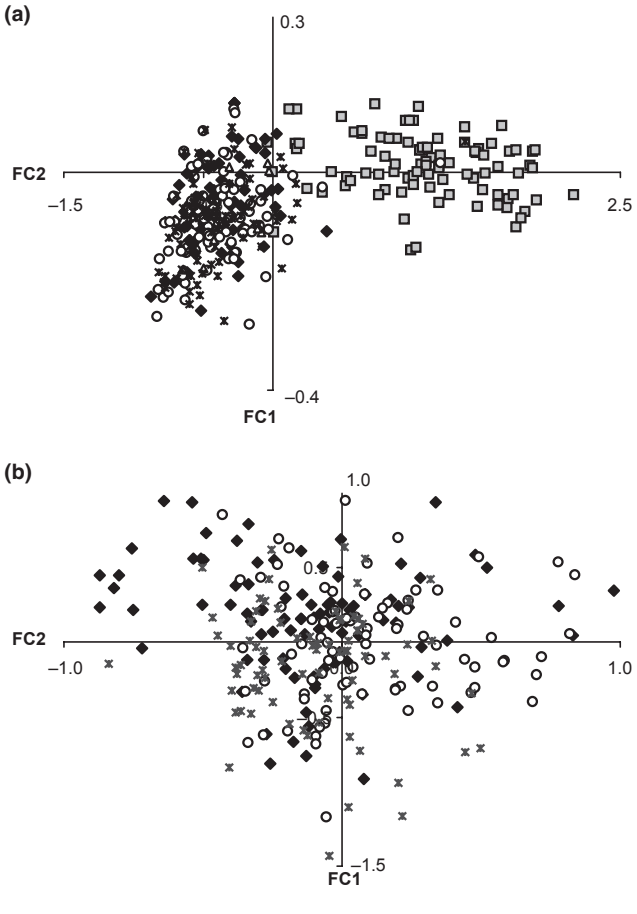

Fig. 2. Factorial correspondence analysis based upon five microsatellite loci for identified and unidentified taxa. (a) Four species plus unidentified specimens (b) three closely related taxa only (*Astatotilapia nubila*, *Xystichromis phytophagus* and the unidentified specimens). (▣) *Lipochromis maxillaris*, (♦) *A. nubila*, (○) *X. phytophagus*, (Δ) *Astatotilapia* sp. '*Bigeye*', (⁎) unidentified specimens

turbidity interfering with the vision-based mate selection (Seehausen et al. 1997). Meanwhile, viable offspring have readily been obtained from crossings of several haplochromine cichlid species under laboratory conditions (Seehausen et al. 1997; van der Sluijs et al. 2008), and the existence of hybrids was also demonstrated in the wild (Seehausen et al. 2008). Hybridization, however, has a second important implication concerning the mode of evolution at such an early stage of adaptive radiation. Seehausen (2004) listed several cases in which evolutionarily young species underwent secondary admixis after an initial phase of diversification. Admixis could have been triggered by climatic events affecting the lake level or perturbating lacustrine habitat (Sturmbauer and Meyer 1992; Johnson et al. 1996). Hybridization events produce broad variation in ecologically relevant traits, or even novel (so-called transgressive) phenotypes, on which divergent selection can act as soon as ecological conditions have normalized. Seehausen (2004) suggested that the instantaneous and simultaneous generation of variation in several functional traits can boost speciation rates to a level impossible to reach via non-hybrid speciation. The observation of slightly distinct albeit overlapping body shapes (Odhiambo et al. 2011) and an extremely close genetic relationship between three of the five taxa are fully compatible and in support of the hybrid swarm theory of adaptive radiation.

Effect of potential bottlenecks

While the desiccation of Lake Victoria during the Pleistocene has been well documented by Johnson et al. (1996, 2000), recent findings by Elmer et al. (2009) have confirmed that desiccation did not extirpate the haplochromines but only bottlenecked them. Maeda et al. (2009) also found indications for recent population growth in two closely related Lake Victoria haplochromines. In the most recent past, haplochromines in the main Lake Victoria have suffered severe reductions in population size owing to upsurge of the introduced *Lates niloticus* – Nile perch (Witte 1992; Goudswaard et al. 2006). According to Luikart et al. (1998), the distortion that is caused by a severe population reduction, when rare alleles are lost more rapidly than less rare ones, is only transient and likely to be detectable only for a few dozens of generations. Cornuet and Luikart (1996) also observed that population substructure might obscure the heterozygosity excess expected

Table 4. Proportion of membership of each predefined population in each of the two clusters

Given population	Inferred clusters		Number of individuals
	1	2	
Lipochromis maxillaris	0.967	0.033	96
Astatotilapia nubila	0.014	0.986	96
Xystichromis phytophagus	0.024	0.976	96
Astatotilapia sp. '*Bigeye*'	0.013	0.987	8
Unidentified	0.039	0.961	91

in a bottleneck population. Therefore, owing to the fact that the Pleistocene desiccation happened in the quite distant past and Lake Kanyaboli might as well have been a substructure of the general Lake Victoria population, it is no surprise that no bottleneck was detected. The latter reduction due to Nile perch upsurge did not apply to the haplochromines in Lake Kanyaboli because the lake has so far not been populated by Nile perch. Even though Lake Kanyaboli is still connected to Lake Victoria through the Yala swamp, low dissolved oxygen levels that characterize such swamps are believed to have prevented the Nile perch from gaining entry into some satellite lakes including Lake Kanyaboli.

Origin of Lake Kanyaboli haplochromine species and species number

Compared to Lake Victoria, Lake Kanyaboli has an extremely species-poor haplochromine species diversity. Three of the five surveyed taxa also occur or at least occurred in Lake Victoria, so that they most likely invaded Lake Kanyaboli. Only *A*. sp. 'Bigeye' is endemic to Lake Kanyaboli, as it has not been found in any other lakes yet. The results of our comparative morphological analysis (Odhiambo et al. 2011) corroborated the original strategy to collect and analyse all unidentifiable specimens, in that almost all collected adult specimens that did not fall within any of the already recognized species clustered in one homogenous cloud separated somewhat from *A*. *nubila* and *X*. *phytophagus* (figure 6b in Odhiambo et al. 2011). These specimens should have been spread over or in-between the recognized species, when they had been misidentified or hybrid specimens were abundant. Indeed, most specimens constitute a distinct morphogroup, albeit with considerable overlap with the other two species. The same picture emerged from our FCA analysis based upon five microsatellite loci (Fig. 2b), so that it seems likely that a fifth species of modern haplochromines in *status nascendi* exists. The number of alleles observed in the so far unidentified entity was also similar to that observed in the other four species. Also, the so far unidentified group of specimens had six private alleles and was genetically homogenous. One should also mention that *A*. *nubila* and *X*. *phytophagus* have very similarly low F_{ST} in relation to their F_{ST} to the unidentified individuals but are clearly differentiated in terms of morphology, feeding ecology and colour. This means that selectively neutral (or near-neutral) genetic markers are obviously not yet segregating, even if traits under natural selection are already distinct. This is the picture one would expect for extremely young or incipient species, especially when they emerge from a hybrid swarm (syngameon). Potential introgression of another Lake Victoria species, even in combination with the hybrid swarm theory, would be consistent with the existence of private alleles in the group of unidentified specimens. Whether the candidate incipient species evolved *in situ* or is also present in the main Lake Victoria cannot be decided at the present stage. Further clarification concerning the relationship between the unidentified specimens in relation to other Lake Victoria species can only come from an extended multilocus approach with a rigorous phylogenetic framework. For species that still occur in the main Lake Victoria (*L*. *maxillaris* and *A*. *nubila*), it would be interesting to extend the data set to include the main lake so as to assess the connection between the satellite and the main lake.

Acknowledgements

We would like to thank the administration and the members of the Ichthyology Section of the National Museums of Kenya for facilitation and assistance with the fieldwork. The Fisheries Department of Kenya, Siaya District Office, provided valuable support. We also thank E. Were of Dominion Farms for valuable help in logistics. S. Koblmüller and two anonymous referees gave valuable comments to this manuscript. This work was supported by grants P17680 and P20994 of the Austrian Science Fund. E. Odhiambo was supported by a PhD fellowship by the Austrian Academic Exchange Service (OeAD).

Zusammenfassung

Genetische Differenzierung von vier Arten moderner Haplochromis*-artiger Buntbarsche in einem Satellitensee des Viktoriasees*

Der ostafrikanische Viktoriasee ist berühmt für seine extrem junge und artenreiche Buntbarsch-Fauna. Zu diesem 'Super-Artenschwarm' zählen auch viele Arten aus den umliegenden Gewässersystemen. Der Viktoriasee ist umgeben von einer Reihe kleinerer Satelliten-Seen, die oft nur durch seichte Sumpfgebiete abgetrennt sind. Im Kanyabolisee, der einer der Satelliten-Seen ist, findet man eine wesentlich geringere Anzahl von Arten, von denen die meisten auch im Viktoriasee vorkommen. Diese Studie konzentrierte sich auf die modernen *Haplochromis*-artigen Buntbarsche *Lipochromis maxillaris*, *Astatotilapia nubila*, *Xystichromis phytophagus* und *Astatotilapia* sp. '*Bigeye*', sowie eine erhebliche Zahl nicht identifizierbarer Individuen, die sich phänotypisch deutlich von den bereits beschriebenen Arten unterschieden. Um die Kongruenz von phänotypischer und genetischer Divergenz zu testen, wurden fünf Mikrosatellitenloci analysiert. Genetisch unterschied sich *L*. *maxillaris* am deutlichsten von allen anderen Einheiten. Auch *A*. sp. '*Bigeye*' unterschied sich klar von den verbleibenden drei Einheiten. Diese erwiesen sich als genetisch sehr nahe verwandt, hatten jedoch alle einige (5–14) private Allele. Sie haben sich vermutlich erst vor kurzer Zeit gebildet. Die nicht identifizierten Individuen bilden eine homogene Gruppe innerhalb der drei genetisch nah verwandten Einheiten und könnten daher als junge oder erst in Entstehung begriffene Art gesehen werden. Da drei der fünf Taxa auch im Viktoriasee vorkommen, ist die Evolution dieser Taxa in die adaptive Radiation des Victoria-Superflocks eingebunden. Die große morphometrische und genetische Überlappung dreier Taxa ist im Einklang mit der Hybridschwarm-Theorie der adaptiven Radiation. Dabei erzeugt eine umweltinduzierte sekundäre Admixis junger Arten jenen Pool an morphologischen Varianten, der dann sehr schnell unter der Wirkung von disruptiver und sexueller Selektion in eigenständige Arten evolvieren kann.

References

Abila R, Barluenga M, Engelken J, Meyer A, Salzburger W (2004) Population structure of a haplochromine cichlid fish from a satellite lake of Lake Victoria. Mol Ecol **13:**2589–2602.

Abila R, Salzburger W, Ndonga MF, Owiti DO, Barluenga M, Meyer A (2008) The role of the Yala swamp lakes in the conservation of Lake Victoria region haplochromine cichlids: evidence from genetic and trophy ecology studies. Lakes Reservoirs: Res. Manage. **13:**95–104.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. Mol Ecol **11:**155–165.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) Genetix 4.05, logiciel sous Windows TM pour la genetique des populations. Laboratoire Genome, Populations, Interactions, CNRS UMR 5000, Universite de Montpellier II, Montpellier, France.

Carleton KL, Streelman JT, Lee BY, Garnhart N, Kidd M, Kocher TD (2002) Rapid isolation of CA microsatellites from the *Tilapia* genome. Anim Genet **33:**140–144.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics **1444:**2001–2014.

Crapon de Caprona MD, Fritzsch B (1984) Interspecific fertile hybrids of haplochromine Cichlidae (Teleostei) and their possible impossible importance for speciation. Neth J Zool **34:**503–538.

Elmer KR, Reggio C, Wirth T, Verheyen E, Salzburger W, Meyer A (2009) Pleistocene desiccation in East Africa bottlenecked but did not extirpate the adaptive radiation of Lake Victoria haplochromine cichlid fishes. PNAS **106:**13404–13409.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol **14:**2611–2620.

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online **1:**47–50.

Fryer G (1997) Biological implications of a suggested Late Pleistocene desiccation of Lake Victoria. Hydrobiologia **354:**177–182.

Fryer G (2001) On the age and origin of the species flock of haplochromine cichlid fishes of Lake Victoria. Proc Biol Sci **268:**1147–1152.

Fryer G (2004) Speciation rates in lakes and the enigma of Lake Victoria. Hydrobiologia **519:**167–183.

Fryer G, Iles TD (1972) The Cichlid Fishes of the Great Lakes of Africa: Their Biology and Evolution. Neptune City, TFH Publications.

Goudswaard KPC, Witte F, Katunzi EFB (2006) The invasion of an introduced predator, Nile perch (*Lates niloticus*) in Lake Victoria (East Africa): chronology and causes. Environ Biol Fish **81:**127–139.

Greenwood PH (1965) The cichlid fishes of Lake Nabugabo, Uganda. Bull Br Mus **12:**313–357.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics **23:**1801–1806.

Johnson TC, Scholz CA, Talbot MR (1996) Late Pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. Science **273:**1091–1093.

Johnson TC, Kelts K, Odada E (2000) The Holocene History of Lake Victoria. Ambio **29:**2–11.

Kaufman LS, Ochumba P (1993) Evolutionary and conservation biology of cichlid fishes as revealed by faunal remnants in Northern Lake Victoria. Conserv Biol **7:**719–730.

Koblmüller S, Schliewen UK, Duftner N, Sefc KM, Katongo C, Sturmbauer C (2008) Age and spread of the haplochromine cichlid fishes in Africa. Mol Phylogenet Evol **49:**153–169.

Luikart G, Allendorf FW, Cornuet J-M, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered **89:**238–247.

Maeda K, Takeda M, Kamiya K, Aibara M, Mzighani SI, Nishida M, Mizoiri S, Sato T, Terai Y, Okada N, Tachida H (2009) Population structure of two closely related pelagic cichlids in Lake Victoria, *Haplochromis pyrrhocephalus* and *H. laparogramma*. Gene **441:**67–73.

Meyer A (1993) Phylogenetic relationships and evolutionary processes in East African cichlid fishes. Trends Ecol Evol **8:**279–284.

Meyer A, Kocher TD, Basasibwaki P, Wilson AC (1990) Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. Nature **347:**550–553.

Mzighani SI, Nikaido M, Takeda M, Seehausen O, Budeba YL, Ngatunga BP, Katunzi EF, Aibara M, Mizoiri S, Sato T, Tachida H, Okada N (2010) Genetic variation and demographic history of the *Haplochromis laparogramma* group of Lake Victoria - An analysis based on SINEs and mitochondrial DNA. Gene **450:**39–47.

Nagl S, Tichy H, Mayer WE, Takezaki N, Takahata N, Klein J (2000) The origin and age of haplochromine fishes in Lake Victoria, East Africa. Proc Biol Sci **267:**1049–1061.

Odhiambo EA, Kerschbaumer M, Postl L, Sturmbauer C (2011) Morphometric differentiation among haplochromine cichlid fish species of a satellite lake of Lake Victoria. J Zool Syst Evol Res **49:**216–223.

Ogutu-Ohwayo R (1990) The decline in the native fishes of Lakes Victoria and Kyoga (East Africa) and the impact of the introduced species, especially the Nile perch, *Lates niloticus*, and Nile tilapia, *Oreochromis niloticus*. Environ Biol Fish **27:**81–96.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes **4:** 535–538.

van Oppen MJH, Rico C, Deutsch JC, Turner GF, Hewitt GM (1997) Isolation and characterization of microsatellite loci in the cichlid fish *Pseudotropheus zebra*. Mol Ecol **6:**387–388.

Piry S, Luikart G, Cornuet JM (1999) Bottleneck: a computer program for detecting recent reductions in effective population size from allele frequency data. J Hered **90:**502–503.

Pritchard JK, Stephen M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics **155:**945–959.

Rosenberg NA (2003) Distruct: a program for the graphical display of population structure. Mol Ecol Notes **4:**137–138.

Salzburger W, Meyer A, Baric S, Verheyen E, Sturmbauer C (2002) Phylogeny of the Lake Tanganyika cichlid species flock and its relationship to the Central and East African haplochromine cichlid fish faunas. Syst Biol **51:**113–135.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbour, New York.

Samonte IE, Satta Y, Sato A, Tichy H, Takahata N, Klein J (2007) Gene flow between species of Lake Victoria haplochromine fishes. Mol Biol Evol **24:**2069–2080.

Seehausen O (2000) Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: effects of sexual selection. Adv Ecol Res **31:**237–274.

Seehausen O (2002) Patterns in fish radiation are compatible with Pleistocene desiccation of Lake Victoria and 14600 year history for its cichlid species flock. Proc Biol Sci **269:**491–497.

Seehausen O (2004) Hybridization and adaptive radiation. Trends Ecol Evol **19:**198–207.

Seehausen O, van Alphen JJM (1998) The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex). Behav Ecol Sociobiol **42:**1–8.

Seehausen O, van Alphen JJM, Witte F (1997) Cichlid fish diversity threatened by eutrophication that curbs sexual selection. Science **277:**1808–1811.

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:**620–626.

van der Sluijs I, van Dooren TJM, Seehausen O, van Alphen JJM (2008) A test of fitness consequences of hybridization in sibling species of Lake Victoria cichlid fish. J Evol Biol **21:**480–491.

Sturmbauer C, Meyer A (1992) Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. Nature **358:**578–581.

Terai Y, Takezaki N, Mayer WE, Tichy H, Takahata N, Klein J, Okada N (2004) Phylogenetic relationships among East African haplochromine fish as revealed by short interspersed elements (SINEs). J Mol Evol **58:**64–78.

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.

Verheyen E, Salzburger W, Snoeks J, Meyer A (2003) Origin of the superflock of cichlid fishes from lake Victoria, East Africa. Science **300:**325–329.

Witte F (1992) The destruction of an endemic species flock: quantitative data on the decline of the haplochromine cichlids of Lake Victoria. Environ Biol Fishes **34:**1–28.

Witte F, Msuku BS, Wanink JH, Seehausen O, Katunzi EFB, Goudswaard PC, Goldschmidt T (2000) Recovery of cichlid species in L. Victoria: an examination of factors leading to differential extinction. Rev Fish Biol Fish **10:**233–241.

Zardoya R, Vollmer DM, Craddock C, Streelman JT, Karl S, Meyer A (1996) Evolutionary conservation of microsatellite flanking regions and their use in resolving the phylogeny of cichlid fishes (Pisces: Perciformes). Proc Biol Sci **263:**1589–1598.

Supporting Information

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

Table S1. Identification of the analysed specimens and their catalogue number at the Ichthyology Department of the National Museums of Kenya, where the specimens were deposited.

Table S2. STRUCTURE and ΔK results for $K = 1–7$ (burn-in: 100 000, run length: 500 000).

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