Molecular Ecology (2004) **13**, 2589–2602 doi: 10.1111/j.1365-294X.2004.02270.x

Population-structure and genetic diversity in a haplochromine fish cichlid of a satellite lake of Lake Victoria

ROMULUS ABILA,*† MARTA BARLUENGA,* JOHANNES ENGELKEN,* AXEL MEYER* and WALTER SALZBURGER*

**Department of Biology, University Konstanz, Universitaetsstrasse 10, 78457 Konstanz, Germany, †Department of Zoology, Maseno University, PO Box 333, Maseno, Kenya*

Abstract

The ~500 species of the cichlid fish species flock of Lake Victoria, East Africa, have evolved in a record-setting 100 000 years and represent one of the largest adaptive radiations. We examined the population structure of the endangered cichlid species *Xystichromis phytophagus* from Lake Kanyaboli, a satellite lake to Lake Victoria in the Kenyan Yala wetlands. Two sets of molecular markers were analysed — sequences of the mitochondrial control region as well as six microsatellite loci — and revealed surprisingly high levels of genetic variability in this species. Mitochondrial DNA sequences failed to detect population structuring among the three sample populations. A model-based population assignment test based on microsatellite data revealed that the three populations most probably aggregate into a larger panmictic population. However, values of population pairwise F_{ST} indicated moderate levels of genetic differentiation for one population. Eleven distinct mitochondrial haplotypes were found among 205 specimens of *X. phytophagus*, a relatively high number compared to the total number of 54 haplotypes that were recovered from hundreds of specimens of the entire cichlid species flock of Lake Victoria. Most of the *X. phytophagus* mitochondrial DNA haplotypes were absent from the main Lake Victoria, corroborating the putative importance of satellite lakes as refugia for haplochromine cichlids that went extinct from the main lake in the last decades and possibly during the Late Pleistocene desiccation of Lake Victoria.

Keywords: cichlid species flock, Lake Victoria, microsatellites, mitochondrial DNA haplotype network, population structure, satellite lake

Received 7 February 2004; revision received 22 April 2004; accepted 28 May 2004

Introduction

With a surface area of about 69 000 km², equivalent to the size of Ireland, Lake Victoria is the largest tropical freshwater lake in the world. However, the shallow lake basin is only 80 m deep at its maximum, in contrast to the other two large, but much deeper, East African lakes, the Rift Valley lakes Malawi and Tanganyika. Since its scientific discovery in 1858, when John Hanning Speke identified Lake Victoria as the long quested source of the Nile, the lake has attracted intense interest from many types of biologists. At the centre of the scientific studies of Lake Victoria is the adaptive radiation of about 500 species of haplochromine cichlid fishes (Fryer & Iles 1972; Greenwood 1981; Seehausen *et al*. 1997a; Stiassny & Meyer 1999; Kornfield & Smith 2000; Verheyen *et al*. 2003). Unique cichlid species flocks also formed in lakes Tanganyika and Malawi (Kornfield & Smith 2000) — in Lake Tanganyika several cichlid lineages even radiated in parallel (Nishida 1991; Salzburger *et al*. 2002) and in Lake Malawi more endemic species have been described than from any other lake (Turner *et al*. 2001). However, with an age of a mere 100 000 years (Verheyen *et al*. 2003), the Lake Victoria species flock of haplochromines represents the fastest known large-scale explosive speciation in vertebrates.

The cichlid species flock of Lake Victoria was among the first endemic faunas to be studied by means of modern molecular phylogenetic tools (Meyer *et al*. 1990). The most recent study proposed an origin of the 'superflock' (including the cichlid faunas from lakes Victoria, Albert, Edward, George, Kgoga and Kivu) from Lake Kivu endemics

Correspondence: Axel Meyer. Fax: + 49-7531-883018; E-mail: axel.meyer@uni-konstanz.de

(Verheyen *et al*. 2003). Also, the age of the species flock has been discussed in detail (Meyer *et al*. 1990; Booton *et al*. 1999; Nagl *et al*. 2000; Fryer 2001; Sturmbauer *et al*. 2001; Seehausen 2002; Seehausen *et al*. 2003; Verheyen *et al*. 2003), particularly after geological data suggested that the entire lake basin dried out completely between 15 600 and 14 700 years ago (Johnson *et al*. 1996). Despite the many studies on the origin and age of the lake's endemic cichlid fauna, the population structure within Lake Victoria cichlid species has not been studied extensively.

Most alarming is the ecological disaster (Goldschmidt 1996) that afflicts Lake Victoria as a result of the introduction of the Nile Perch *Lates niloticus* in the 1950s to improve fisheries (Ogutu-Ohwayo 1990, 1993; Witte *et al*. 1992; Goldschmidt *et al*. 1993) (see also Seehausen *et al*. 1997b; Witte *et al*. 1999, 2000; for information on partial recovery of Lake Victoria's haplochromine fauna in the 1990s). The observed drastic decline of cichlids from Lake Victoria is, however, not only the result of direct predation by the Nile Perch upon cichlids. Other anthropogenic causes, such as deforestation coupled with eutrophication and turbidization of the lake habitat, which was hypothesized to leading to hybridization of many cichlid species (Seehausen *et al*. 1997a), and the introduction of the nonindigenous water hyacinth overgrowing the lake's surface, have contributed as well to the eradication of approximately 200 endemic cichlid species (Goldschmidt *et al*. 1993; Goldschmidt 1996). The deforestation around Lake Victoria was also increased because the oily meat of the large Nile Perch had to be smoked for preservation instead of using the traditional method of sun-drying, which was used with the smaller haplochromine cichlids. This eradication of large numbers of Lake Victoria cichlids is possibly the largest extinction event of vertebrate species in modern human history (Goldschmidt *et al*. 1993).

Satellite lakes, i.e. small water bodies isolated from Lake Victoria by sand bars or papyrus swamps, have not only provided excellent opportunities for evolutionary research (Greenwood 1965; Fryer & Iles 1972), but have also been postulated to be refugia for cichlid species that have already disappeared from the main lake (Kaufman & Ochumba 1993). The best-known example for a satellite lake of Lake Victoria is probably Lake Nabugabo in Uganda. It is only about 24 km^2 in area, has an average depth of 4.5 m, and is about 4000 years old. Five out of its eight haplochromine species are endemic and seem to have evolved within the confines of the small lake (Greenwood 1965; Kaufman & Ochumba 1993). From Lake Kanyaboli (size 10.5 km^2; average depth 2.5 m) in the Kenyan Yala swamps two endemic species have been described, of which one species (a small, large-eyed algae-scraper) seems to have evolved within the lake and thus is likely to represent a true endemic, while the second species, *Xystichromis phytophagus*, was suggested to have disappeared from the main water body of Lake Victoria in the last decade (Kaufman & Ochumba 1993). Some satellite lakes also provide the opportunity to study cichlid communities in the absence of the Nile Perch. For example, piscivore cichlids are still abundant in some of these lakes, such as the paedophage *Lipochromis maxillaris* in Lake Kanyaboli (Kaufman & Ochumba 1993; Ogutu-Ohwayo 1993). This is not true for Lake Nabugabo, where the Nile Perch was introduced in the 1960s (Ogutu-Ohwayo 1993).

Here, we present a population-genetic survey of three populations of the endangered cichlid species *Xystichromis phytophagus* from Lake Kanyaboli, Kenya. We have chosen this species from a small and shallow satellite lake of Lake Victoria for four main reasons. First, we wanted to assess the population structure of a haplochromine species of the 'Lake Victoria superflock' in a habitat where no obvious barriers for gene flow exist. Second, we wanted to study the population structure of such a haplochromine species in the absence of the Nile Perch, which massively affected the cichlid communities of Lake Victoria. Third, we intended to compare the mitochondrial haplotype diversity within *X. phytophagus* to those of the entire cichlid species flock of Lake Victoria (Nagl *et al*. 2000; Verheyen *et al*. 2003). Finally, the recent decline of haplochromine cichlids in Lake Victoria has called for urgent research and the conservation of the remaining species and populations where they still exist.

We compared the population structure of three sample populations of *X. phytophagus* from Lake Kanyaboli using two kinds of markers, mitochondrial control region sequences and six microsatellite loci. We also constructed a haplotype network including all available mitochondrial control region sequences from cichlids from Lake Victoria and compared the haplotype diversity within *X. phytophagus* to other cichlids of the 'Lake Victoria superflock'.

Materials and methods

Study area

The samples were collected in Lake Kanyaboli, a small (10.5 km^2) and shallow (average depth 2.5 m; maximum depth 4.5 m) freshwater lake situated in the Yala wetlands in Western Kenya (Fig. 1). The Yala swamp is Kenya's largest wetland and covers about 175 km^2 along the northern shores of Lake Victoria. Recent satellite imaging surveys suggest that the swamp area might be much larger (M. Otieno, personal communication). It is bordered to the North by the Nzoia River and to the South by the Yala River. Three main lakes exist in the Yala wetlands (Kanyaboli, Namboyo and Sare), of which Lake Kanyaboli is the largest and most remote from Lake Victoria. Lake Kanyaboli is separated from Lake Victoria by massive papyrus swamps that at present inhibit faunal exchanges

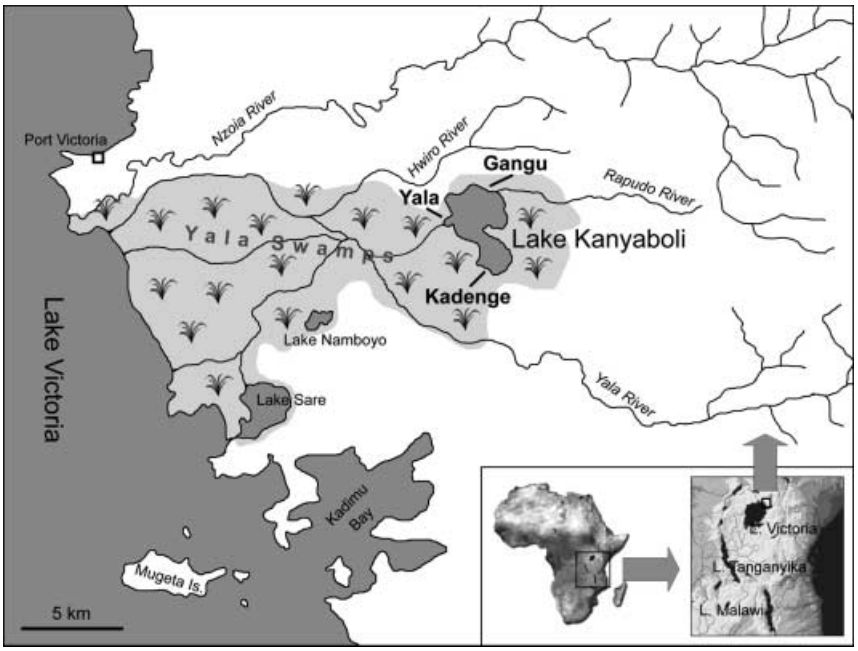

Fig. 1 Map showing the Yala Swamp area at the Kenyan shore of Lake Victoria and the three sampling populations (Gangu, Kadenge and Yala) in Lake Kanyaboli.

between the two lakes. No Nile Perch has ever been observed in Lake Kanyaboli, corroborating that it has been isolated from Lake Victoria since at least the 1950s. The fish fauna of the Yala swamp lakes is dominated by cichlids — three species of tilapia (*Oreochromis esculentus*, *O. niloticus* and *O. leucostictus*) and eight haplochromine cichlid species (Kaufman & Ochumba 1993; Odhiambo 2002). Besides its cichlid fauna, the Yala swamp is home to a rich and complex community of animals including the endangered sitatunga antelope (*Tragecephalus spekei*) as well as endemic papyrus birds. Unfortunately, past reclamation endeavours to improve food security in this area have greatly altered the wetland habitat and its fauna, and threaten the future survival of its biodiversity.

Three populations of *Xystichromis phytophagus* from Lake Kanyaboli were sampled: 'Gangu' in the north, close to the mouth of River Rapudo, 'Kadenge' on the southern shore, and 'Yala' in a bay on the western shoreline of the lake. The distance between the sampling sites ranged between 2 and 2.5 km (see Fig. 1).

Sampling and DNA extraction

A total of 205 adult specimens of *X. phytophagus* were collected from three localities in Lake Kanyaboli (Fig. 1, Table 1), with an equal proportion of males and females in our sample. Sampling was performed between February and July 2002. Each population was sampled at least twice to account for potential temporal fluctuations during the rainy season from April to June. Voucher specimens have been deposited at the Royal Museum for Central Africa, Tervuren, Belgium. Muscle tissue from preserved specimens in 90% ethanol was used as a source of DNA. Total DNA was extracted by sodium chloride extraction and ethanol precipitation after initial proteinase K digestion (Bruford *et al*. 1998).

Table 1 Number of individuals per haplotype (frequencies) in *Xystichromis phytophagus* in the sampled populations Gangu, Kadenge and Yala

Haplotype	Individuals/ haplotype (%)	Gangu	Kadenge	Yala
1	99 (48.29)	33	32	34
2	43 (20.98)	13	19	11
3	30 (14.63)	7	10	13
4	6 (2.93)	3	2	1
5	17 (8.30)	9	5	3
6	2 (0.97)	—	2	—
7	2 (0.97)	1	—	1
8	1 (0.49)	1	—	—
9	3 (1.46)	1	2	—
10	1 (0.49)	1	—	—
11	1 (0.49)	1	—	—
Total	205 (100)	70	72	63

Mitochondrial DNA (mtDNA) amplification and sequencing

For polymerase chain reaction (PCR) amplification of the first section of the mitochondrial control region, the fastest evolving segment of the mitochondrial genome, the published primers L-Pro-F (Meyer *et al*. 1994) and TDK-D

(Kocher *et al*. 1989) were used. PCR amplification was performed in a reaction volume of 21.1 μL [9.9 μL high-performance liquid chromatography (HPLC) water, 2 μL buffer, 1.6 μL 10 mM dNTPs, 1.4 μL 10 mM $MgCl_2$, 2 μL of each primer/2 nM, 0.2 μL *Taq* DNA polymerase and 2 μL diluted DNA] under the following conditions: 35 cycles with a denaturation phase at 94 °C for 30 s, an annealing phase at 52 °C for 30 s, and an extension phase at 72 °C for 90 s. PCR products were visualized by mini-gel electrophoresis using ethidium bromide staining and 1% agarose gels.

Two microlitres of purified PCR product were used as the template in the cycle sequencing reaction. The reaction mixture for cycle sequencing was made up of 1 μL of 10 μM L-Pro-F primer, 1.5 μL BigDye termination reaction mix (Applied Biosystems) and 5.5 μL HPLC water. The annealing temperature for cycle sequencing was adjusted to 50 °C. The cycle-sequenced products were purified with an ethanol–sodium acetate precipitation, re-suspended in 15 μL HPLC water and analysed on an ABI 3100 capillary DNA sequencer (Applied Biosystems).

Microsatellites

A total of six microsatellite loci developed for other cichlid species [*Copadichromis cyclicos* (UNH001, UNH002; Kellogg *et al*. 1995); *Tropheus moorii* (TmoM5, TmoM11, TmoM27; Zardoya *et al*. 1996); and *Astatoreochromis alluaudi* (OSU20D; Wu *et al*. 1999)], were chosen for population analysis. The loci were selected to have at least 10 uninterrupted tandem repeats (based on the information available from other cichlid species) and to show variation in *X. phytophagus*. For PCR amplification, the same reaction mix composition as above was used (total volume: 21.1 μL), the forward primers being end-labelled with fluorescent dyes FAM or HEX. The PCR conditions were as described above, except for a final extension phase after 35 cycles at 72 °C for 30 min. The PCR products were diluted 1 : 10 and 1 μL of this dilution was added to 0.125 μL of ROX 500 size standard (Applied Biosystems) in 9 μL HPLC water. The samples were then denatured for 4 min at 94 °C and immediately placed on ice. The microsatellite markers were analysed on an ABI 3100 capillary sequencer (Applied Biosystems), the loci were scored with GENESCAN and GENOTYPER software (Applied Biosystems).

Data analysis I: mitochondrial control region

The obtained DNA sequences of the first segment of the mitochondrial control region were edited and aligned by eye using the computer programs SEQUENCE NAVIGATOR (Applied Biosystems) and BIOEDIT (Hall 2003), resulting in an alignment of 359 base pairs. Eleven different haplotypes were detected in the 205 specimens of *X. phytophagus* (Table 1). To compare the genetic diversity in the *X. phytophagus* populations of Lake Kanyaboli to those of the entire species flock of Lake Victoria, we added 42 haplotypes from Lake Victoria and five haplotypes found in Lake Kivu and other smaller lakes (Meyer *et al*. 1990; Nagl *et al*. 2000; Verheyen *et al*. 2003) to these 11 haplotypes and constructed a minimum spanning network with the computer program TCS (Clement *et al*. 2000). Note that we did not include haplotypes 75 and 76 of Verheyen *et al*. (2003), since these haplotypes belong to a separate lineage. Thus, a total of 58 haplotypes representing 340 sequences were used for the construction of the haplotype network. A detailed list of all studied taxa, their haplotype assignment and sampling localities, the haplotype frequency if available, and the GenBank accession numbers are given in Appendix I. We also provide original haplotype numbers for the sequences already analysed by Verheyen *et al*. (2003). Alternative branching orders in the TCS-generated network were assessed by a maximum likelihood search with PAUP* 4.0b10 (Swofford 2003) and only those connections between haplotypes favoured by the maximum likelihood method were depicted leaving the general topology of the TCS-generated network unaffected. We also calculated the consistency index (CI) (Kluge & Farris 1969) for each mutation to identify homoplasious mutations (CI < 1) and depicted these mutations in our haplotype network.

In a second step, the genetic differences between the sampled populations of *X. phytophagus* were tested using *F*-statistics (Weir & Cockerham 1984) as calculated by ARLEQUIN v.2.1 (Schneider *et al*. 1999). Critical significance levels for multiple testing were corrected following the sequential Bonferroni procedure (Rice 1989). In addition, we analysed the extent of geographical heterogeneity in population frequency distribution through a Monte Carlo simulation (for a detailed description see Roff & Bentzen 1989) as implemented in the MONTE program of the REAP 4.1 package (McElroy *et al*. 1992), carrying out 10 000 randomization procedures.

Data analysis II: microsatellite markers

A total of 191 specimens of *X. phytophagus* were successfully amplified for at least five of six microsatellite loci. Genetic variability at these loci was estimated for each population as the number of alleles per locus (N_A), observed (H_O) and expected (H_E) heterozygosity (Table 2). Departure from Hardy–Weinberg expectations for every locus was calculated within and across populations using a test analogous to Fisher's exact tests (Guo & Thompson 1992) estimated with a 100 000 step/1000 iteration Markov Chain Monte Carlo series of permutations, as implemented in ARLEQUIN. Genetic differences between populations were measured with both Wright's *F*-statistics (F_{ST}) (Weir & Cockerham 1984) based on differences in allele frequencies, and *R*-statistics (R_{ST}) (Slatkin 1995) based on

Table 2 Population pairwise F_{ST} (mtDNA sequences and microsatellite loci) and R_{ST} (microsatellite loci)

Population comparison	mtDNA F_{ST}	Microsatellite F_{ST}	Microsatellite R_{ST}
Yala–Gangu	0.008	0.022*	0.000
Yala–Kadenge	0.004	0.010*	0.000
Gangu–Kadenge	0.000	0.002	0.000

*$P < 0.001$; all others were not significant, i.e. $P \geq 0.05$.

differences in the allele size, as implemented in ARLEQUIN (Schneider *et al*. 1999). Both statistics were used because there is no clear consensus over their relative accuracy (see Balloux & Lugon-Moulin 2002).

Linkage disequilibrium for pairs of loci was tested for all possible pairs of loci in each population and globally for each pair of loci across populations with ARLEQUIN. Critical significance levels for multiple testing were corrected following the sequential Bonferroni procedure (Rice 1989).

We then performed a model-based population assignment test as implemented in the computer program STRUCTURE 2.1 (Pritchard *et al*. 2000). Markov Chain Monte Carlo simulations were run with 500 000 replicates and a burn-in of 50 000 replicates for K (number of populations) = 1, 2, 3, 4 and 5, and applying the admixture model, in which individuals may share portions of the genome assigned to more than one population as a result of mixed ancestry (Pritchard *et al*. 2000).

Results

Population structure inferred from mtDNA sequences

Eleven distinct haplotypes were observed in the 205 specimens of *Xystichromis phytophagus*, differing by one to six mutations in the 359-bp alignment. Private haplotypes were only found in two populations, 'Gangu' and 'Kadenge', although in low abundance. The haplotype frequency distribution — especially that of the three main haplotypes occurring in 83.9% of all specimens — was similar in all three populations (Table 1).

Pairwise population F_{ST} comparisons failed to find any genetic structuring among the three studied populations based on mtDNA sequences ($P > 0.05$) (Table 2). Similarly, the assessment of the degree of geographical heterogeneity with REAP 4.1 revealed no significant heterogeneity in mtDNA haplotype frequency distribution among the three populations ($\chi^2 = 20.88$, $P = 0.40$).

Population genetic structure inferred from microsatellites

The microsatellite markers exhibited high polymorphism in *X. phytophagus* (Table 3). A total of 160 alleles had been found in the 191 individuals. The total numbers of alleles per population were 22 ('Yala'), 23 ('Kadenge') and 22 ('Gangu') in UNH001 (total number of alleles: 30); 11, 14 and 12 in UNH002 (16); 21, 24 and 22 in TmoM5 (27); 12, 15 and 19 in TmoM11 (20); 10, 12 and 10 in TmoM27 (13); and 29, 40 and 41 in OSU20D (54).

Table 3 Number of alleles (N_A), observed (H_O) and expected (H_E) heterocygosity, as well as size range of the alleles in the microsatellite loci

Microsatellite locus	N_A	H_O	H_E	Size range
TmoM5	27	0.904	0.923	315–369
TmoM11	20	0.846	0.882	195–233
TmoM27	13	0.623	0.786	380–428
UNH001	30	0.953	0.935	162–228
UNH002	16	0.904	0.895	197–235
Osu20D	54	0.899	0.958	140–280

Significant departures from Hardy–Weinberg equilibrium were detected for 10 out of 24 comparisons. When pooling across populations substantial heterozygote deficit was found in four out of the six analysed loci (UNH002, TmoM11, TmoM27, OSU20D). Deviations from Hardy–Weinberg equilibrium for the same four loci were found in the population of Kadenge, two loci (UNH002, TmoM27) were deviating from Hardy–Weinberg equilibrium in the population from Yala, and no heterozygote deficit was found in the population from Gangu.

The pairwise F_{ST} comparisons based on microsatellites yielded significant differences between the population from Yala and the populations from Kadenge and Gangu, but no population genetic differentiation between the two latter populations was detected. R-statistics, which are thought to reflect more accurately the mutation pattern of microsatellites (Slatkin 1995), failed to find any significant population structuring (Table 2). R-statistics have the drawback of a very high variance, for which reason F_{ST} often outperforms its R_{ST} counterpart (see Gaggiotti *et al*. 1999). Disequilibrium between pairs of loci was nonsignificant in every comparison ($P > 0.05$).

The population assignment test performed with STRUCTURE 2.1 (Pritchard *et al*. 2000) revealed that the three sample populations most probably belong to a single panmictic population ($K = 1$; estimated -ln probability of data = 6070; $P > 0.99$).

Genetic diversity inferred from the haplotype network

In our haplotype network, all haplotypes from Lake Kanyaboli were placed — as expected — within the 'main Lake Victoria clade' of Verheyen *et al*. (2003) (corresponding to linage VC in Nagl *et al*. 2000), and not in the 'rift valley clade' represented by haplotypes 1–5 in Fig. 2 (corresponding to lineage VB in Nagl *et al*. 2000). This

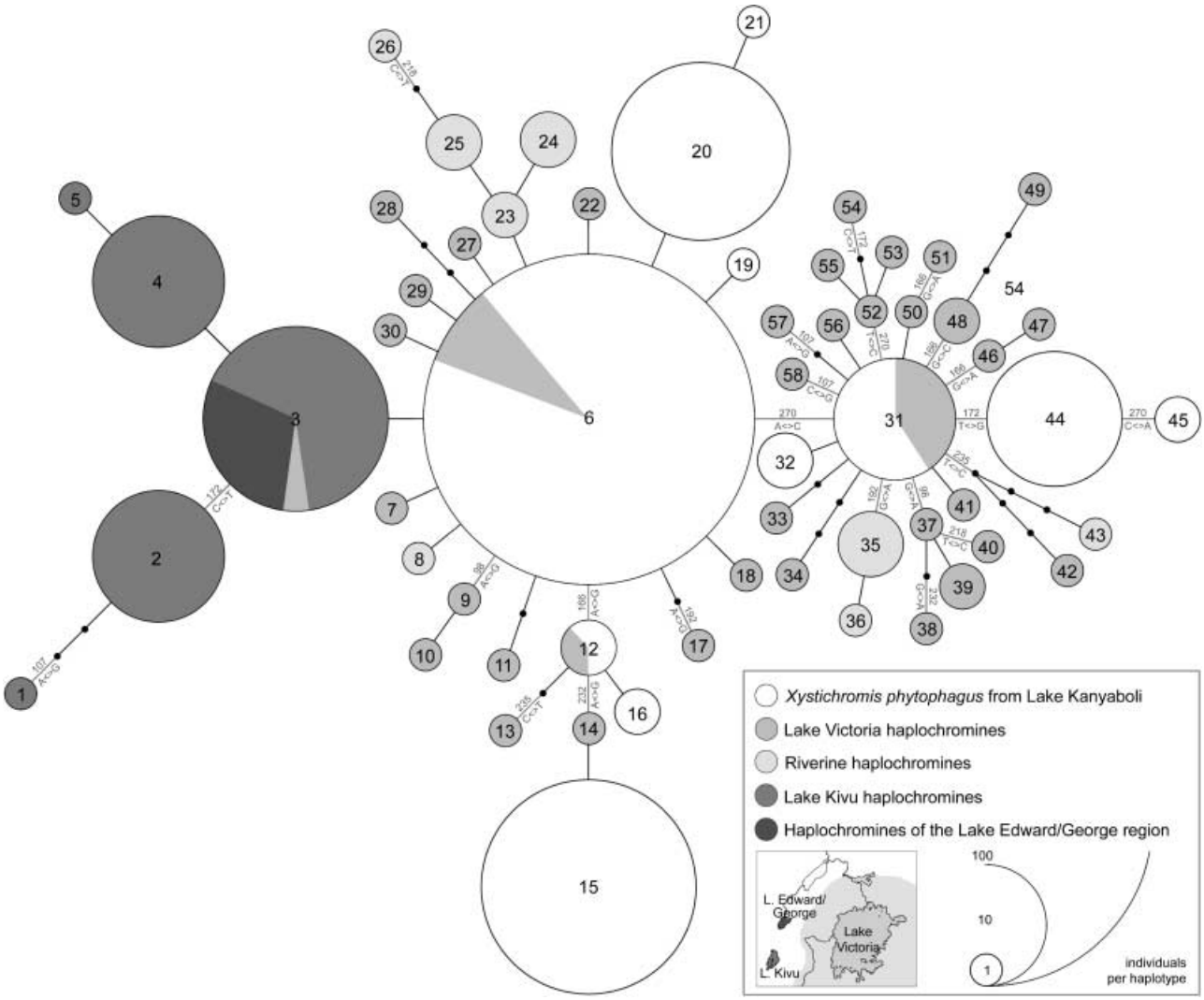

Fig. 2 Unrooted haplotype network of haplochromines from the Lake Victoria cichlid species flock with special emphasis on *Xystichromis phytophagus* from Lake Kanyaboli, including five haplotypes from Lake Kivu and the Lake Edward/George region (Verheyen *et al*. 2003). The most abundant haplotype in *X. phytophagus* was haplotype 6; haplotypes 6, 12 and 31 were shared between *X. phytophagus* and other species of the Lake Victoria cichlid species flock. As an evaluation of the network, and as an indication of where alternative connections would be possible, we depicted the consistency index (CI) (Kluge & Farris 1969) and the specific mutation for all those branches in the network where C < 1. These branches were shaded in grey. The size of the haplotypes refers to the number of specimens that shared that particular haplotype. Note that this information was available only for *X. phytophagus* and some of the remaining haplotypes (Appendix 1).

pattern is further corroborated by the finding that all *X. phytophagus* sequences showed the diagnostic character state 'C' in position 269 of the alignment (Verheyen *et al*. 2003). Also, the sequences determined by Meyer *et al*. (1990), which consist of the first section of the control region and were therefore not included in Verheyen *et al*. (2003), were resolved in that clade — a finding that contradicts that of Nagl *et al*. (2000) who suggested that two sequences determined by Meyer *et al*. (1990) would not belong to their VC clade. Of the 11 distinct haplotypes found in the 205 specimens of *X. phytophagus*, only three were shared with specimens from the main lake. These three haplotypes (6, 12, and 31 in Fig. 2) represented a total of 108 Lake Kanyaboli specimens, so that 97 specimens (47% of all *X. phytophagus* samples) had haplotypes so far not known from Lake Victoria. Haplotype 6 was the most abundant one in *X. phytophagus*. The maximum distance between different Lake Victoria haplotypes was nine mutations (2.5%), that between haplotypes of *X. phytophagus* was six mutations (1.7%).

Discussion

Mitochondrial DNA haplotype diversity

The first molecular phylogenetic study of haplochromine cichlids from the Lake Victoria region based on mtDNA sequences established the monophyly of the species flock (Meyer *et al*. 1990). Several hundred mitochondrial control region sequences of specimens of Lake Victoria haplochromines have been analysed since then (Nagl *et al*. 2000; Verheyen *et al*. 2003). All studies agree on the extremely young age, < 200 000 years, for the entire species assemblage, which is reflected by the small number of haplotypes found as well as the small genetic distances between the haplotypes. In our analysis, a total of 54 haplotypes were detected in the specimens from Lake Victoria and its immediate surroundings [including *Xystichromis phytophagus* from Lake Kanyaboli; but not including haplotypes from Lake Kivu and other 'Rift valley haplotypes' (Verheyen *et al*. 2003)]. Since our haplotype network analysis (Fig. 2) was based on the first section of the control region only, while Verheyen *et al*. (2003) analysed the entire control region, the two resulting networks are not completely comparable. The general topology, however, is almost identical in both networks with a clear separation of the 'Rift Valley haplotypes' found in lakes Kivu, Edward, George, Albert and surroundings (represented by haplotypes 1–5 in Fig. 2) from the haplotypes of main Lake Victoria and its surrounding bodies of water. The shorter data set used here mainly led to the fusion of certain haplotypes and to shorter branch lengths in the network.

In the 205 specimens of *X. phytophagus*, 11 distinct mitochondrial haplotypes were identified. Compared to the total number of 54 different haplotypes for the entire Lake Victoria species assemblage, the finding of such a relatively high number of haplotypes in a single species seems astounding. In addition, the *X. phytophagus* haplotypes differ from each other by up to six mutations, which is close to the amount of the maximum distance between haplotypes in the entire assemblage (nine mutations). Similar haplotype numbers have, however, been observed in two haplochromine species in the crater lakes Nshere (10 haplotypes) and Lutoto (seven haplotypes) in Uganda (Sato *et al*. 2003). Also, some Lake Kivu species show a comparable number of haplotypes (Verheyen *et al*. 2003).

More than half of the *X. phytophagus* individuals from Lake Kanyaboli (but only three haplotypes; see Fig. 2 and Appendix I) shared their mitochondrial control region sequence with other species of the Lake Victoria superflock analysed so far. Shared mtDNA haplotypes between species have repeatedly been found in closely related haplochromines from Lake Malawi (Moran & Kornfield 1993; Parker & Kornfield 1997), in haplochromines from Lake Kivu (Verheyen *et al*. 2003), as well as in haplochromines of the Lake Victoria region (Meyer *et al*. 1990; Nagl *et al*. 2000; Verheyen *et al*. 2003). This incomplete mitochondrial lineage sorting is most likely the result of the extremely young age of these species assemblages coupled with rapid speciation events. Also in noncoding regions of the nuclear DNA the persistence of polymorphisms was detected in the Lake Victoria cichlid assemblage (Nagl *et al*. 1998). However, in the Tropheini, the closest but older sister group in Lake Tanganyika of the haplochromine cichlid species flocks of lakes Malawi and Victoria (Salzburger *et al*. 2002), incomplete mitochondrial lineage sorting could not be detected and — in contrast to the Malawi and Victoria situation — all morphologically distinguishable species were also genetically distinct (Sturmbauer *et al*. 2003).

The DNA sequences of eight of the 11 haplotypes, which represent almost 50% of all studied individuals of *X. phytophagus*, had so far not been found in any other representative of the Lake Victoria superflock. One explanation for this observation might be a sampling bias and the underrepresentation of Lake Victoria haplochromines. Although we compared the 11 haplotypes to those of previous studies, which together included about 900 specimens of haplochromine cichlids from the Lake Victoria region (Meyer *et al*. 1990; Nagl *et al*. 2000; Verheyen *et al*. 2003), it seems that increased sampling might lead to the recovery of new, previously missing, haplotypes. When analysing a large combined data set, Verheyen *et al*. (2003) noted that there are intermediate haplotypes missing in the Lake Victoria sample. They interpreted this as the likely signature of a massive extinction event, which could have been related to the desiccation of Lake Victoria between 15 600 and 14 700 years ago (Johnson *et al*. 1996). This would, however, not explain the existence of the eight private haplotypes in *X. phytophagus*. Because of its much younger age compared to the main Lake Victoria basin and its relative instability because of its maximum depth of only 4.5 m, Lake Kanyaboli could not have acted as a reservoir during this dry period. Also, it is quite unlikely that these eight haplotypes arose *in situ* in Lake Kanyaboli (partly because Lake Kanyaboli is too young to allow for the evolution of so many new haplotypes, and partly because these eight haplotypes have not only one ancestor but have evolved from at least four different Lake Victoria haplotypes). Instead, it might be possible that most of these eight haplotypes may have existed in Lake Victoria but recently became extinct, perhaps as a result of the recent decline of haplochromines caused by the introduction of the Nile Perch.

The population structure of X. phytophagus *in Lake Kanyaboli*

The analyses of the mtDNA sequences revealed nonsignificant levels of genetic structuring between the

three populations of *X. phytophagus* from Lake Kanyaboli. Population pairwise *F*-statistics were nonsignificant in all comparisons (Table 2) and no geographical heterogeneity in population frequency distributions has been observed in the Monte Carlo simulations with REAP 4.1. Also, the model-based population assignment test (STRUCTURE 2.1) indicated that *X. phytophagus* still forms a panmictic population in Lake Kanyaboli, and population pairwise R-statistics based on the microsatellite data were nonsignificant. Conversely, population pairwise comparisons with *F*-statistics based on the microsatellite data showed significant levels of population structuring. The Yala population was found to be isolated from both Gangu and Kadenge (Table 2), however, with a relatively low F_{ST} values of 0.022 (for the pairwise comparison with the Gangu population) and 0.01 (for the pairwise comparison with the Kadenge population).

Taken as a whole, our results revealed that *X. phytophagus* most likely aggregate into a single large panmictic population in Lake Kanyaboli. Only microsatellite-based *F*-statistics show moderate levels of population structuring for the Yala population. The sampling site at Yala differs from those at Gangu and Kadenge in that at Yala Lake Kanyaboli drains into its only outlet forming a shallow bay (Fig. 1). However, this does not result in different environmental characteristics such as pH, conductivity, temperature or visibility (data not shown). As a result of the similarity of the habitats in all three populations, we do not believe that local adaptations account for the observed level of population structuring.

In a review on (marine) fish species with high gene flow, Waples (1998) discussed the risk that statistical tests may reject the null hypothesis of no population differentiation when it is true. Besides the type I error — i.e. rejection of the null hypothesis by chance when it is true — Waples (1998) finds that sampling errors can cause the false rejection of the null hypothesis (he also argues that statistically significant differences may sometimes be too small to be biologically meaningful). According to Waples (1998), random sampling error is a major source of noise in estimating allele frequencies. This seems to be particularly problematic in species with high gene flow, where the signal obtained by *F*-statistics is generally very low. While our taxon sampling, with more than 60 individuals per population, seems sufficient to overcome an intralocus signal : noise ratio problem, the number of loci analysed might be too small to account for a potential interlocus sampling error. The analysis of additional loci would be necessary to avoid such problems.

Both sets of molecular markers revealed a surprisingly high level of genetic variability within *X. phytophagus* in Lake Kanyaboli, implying a historically large effective population size and the lack of population bottlenecks in the past. This suggests that, in Lake Kanyaboli, a relatively large number of specimens have originally been isolated from their ancestors in Lake Victoria. However, the isolation of the Lake Kanyaboli populations from a much larger original population in Lake Victoria might have been accompanied by genetic drift and the random loss of alleles in the colonizing population, which might explain deviations from Hardy–Weinberg equilibrium (given that our sample sizes per population were relatively large it seems unlikely that this pattern is the result of sampling bias).

The population genetic structure of haplochromines from Lake Victoria has so far not been studied in sufficient detail. However, several studies of haplochromines and other cichlids have been undertaken in lakes Malawi and Tanganyika (Arnegard *et al*. 1999; Markert *et al*. 1999; Danley *et al*. 2000; Rüber *et al*. 2001; Rico & Turner 2002; Baric *et al*. 2003). These studies mostly involved rock-dwelling cichlids because the patchy habitat distribution of rocky shores coupled with philopatric behaviour more obviously implies population structuring as a result of isolation. In rock-dwelling Lake Malawi mbuna cichlids, deep waters and sandy stretches have been identified as strong barriers to gene flow (van Oppen *et al*. 1997; Arnegard *et al*. 1999; Markert *et al*. 1999) — sometimes a habitat discontinuity of only about 35 m seems sufficient to isolate two populations (Rico & Turner 2002). However, in pelagic cichlids from Lake Malawi population structuring has not been observed and it has been suggested that at least some species form lake-wide panmictic populations (Shaw *et al*. 2000).

The Lake Kanyaboli *X. phytophagus* are, however, not pelagic, and neither do deep-water barriers exist nor are there habitat discontinuities along the shores. In fact, the lake is very shallow (on average 2.5 m deep) and the bottom is uniformly sandy/muddy with layers of reed debris throughout the lake. This habitat homogeneity might explain why *X. phytophagus* forms a single panmictic population in Lake Kanyaboli, while in other — mostly rock-dwelling — cichlid population structuring has been observed over distances that are far beyond the ~2.5 km that separate the populations investigated here (see e.g. van Oppen *et al*. 1997; Rico & Turner 2002).

Conservation genetic implications

After the recent decline of Lake Victoria's cichlid fauna, satellite lakes have been recognized as one out of four possible major refugia for indigenous fishes (Kaufman & Ochumba 1993). Other possible refugia for haplochromines are schools of the silver cyprinid (*Rastrinoebola argentea*), microbial mats at depths of 30–40 m and the water column near the oxycline [according to Kaufman & Ochumba (1993)]. Although some endemic species from satellite lakes seem to represent new species that evolved in geographical isolation (Greenwood 1965), other satellite

lake species were suggested to have once been common in the main lake but to have disappeared only recently from there (Kaufman & Ochumba 1993; Ogutu-Ohwayo 1993). Here we show that not only are endemic species conserved in such habitats but that also a relatively high level of the genetic diversity is maintained in these isolated water bodies. This is reflected by the relatively large number of mtDNA haplotypes in *X phytophagus* compared to the entire cichlid assemblage of Lake Victoria, as well as by the presence of the large number of alleles in six microsatellite loci. Kaufman & Ochumba (1993) suggested a 'two-fold option' for species conservation in the Lake Victoria area, the preservation of satellite lakes with intact faunas and the use of other (empty) satellite lakes as repositories for indigenous Lake Victoria species. There may be some practical difficulties with the feasibility of the latter option, yet, our genetic data agree with the importance of satellite lakes that Kaufman & Ochumba (1993) recognized for conservation efforts (in particular, of course, for the Yala swamp system). Such conservation efforts should be accompanied by more detailed population genetic studies as well as ecological and limnological research. Regrettably, conflicting demands for greater food production threaten the future existence of this tropical biodiversity hot spot.

It has been suggested that the entire lake basin of Lake Victoria dried out completely during the Late Pleistocene (Johnson *et al*. 1996; but see Fryer 1997, 2001, 2004), raising the possibility that the entire Lake Victoria haplochromine cichlid species flock became extinct and has re-evolved at a record-setting pace since the refilling of the Lake Victoria basin. The finding of a relatively high number of mitochondrial haplotypes in a single species in Lake Kanyaboli (11 haplotypes) compared to the entire Lake Victoria assemblage (54 haplotypes; 38 haplotypes occur in Lake Victoria itself, the rest in surrounding bodies of water; see Appendix I) shows that a relatively large portion of the genetic diversity of Lake Victoria's entire haplochromine cichlid assemblage can be maintained in a small lake. This observation would provide empirical evidence for the theoretical possibility that — if refugia persisted during the Late Pleistocene arid period in the form of small, shallow lakes in the confines of the present Lake Victoria basin (see e.g. Fryer 2001) — such refugial populations could have retained a good portion of the genetic (and possibly morphological) variation of the entire species flock (Verheyen *et al*. 2003, 2004). The loss of genetic variation during the Late Pleistocene arid period, during which an almost complete extinction of most Lake Victoria cichlid species occurred (Verheyen *et al*. 2003), might then have been rather minor permitting the rapid re-colonization and diversification of the Lake Victoria haplochromine cichlid species flock from a small number of shallow refugial lake populations (Verheyen *et al*. 2004).

Acknowledgements

We would like to thank J. Snoeks of the Royal Museum for Central Africa, Tervuren, Belgium, for help during species identification. We would also like to acknowledge the assistance of C. Clabaut, E. Hespeler, A. Merkel and U. Topel; and L. Kaufman, J. Luo, M. Sanetra, H. Waidbacher as well as L. Rüber, the second anonymous reviewer and the Subject Editor Bernatchez for discussion and valuable comments on the manuscript. This study was supported by a DAAD fellowship to R.A., the Consejería de Educación y Cultura del Principado de Asturias, Plan de Investigación Desarrollo Tecnológico e Innovación de Asturias 2000–04 to M.B., grants of the Deutsche Forschungsgemeinschaft to A.M., and the Landesstiftung Baden-Württemberg, the Center for Junior Research Fellows (University Konstanz) and the EU (Marie Curie fellowship) to W.S.

References

Arnegard ME, Markert JA, Danley PD, Stauffer JR, Ambali AJ, Kocher TD (1999) Population structure and colour variation of the cichlid fish *Labeotropheus fuelleborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. *Proceedings of the Royal Society London Series B*, **266**, 119–130.

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

Baric S, Salzburger W, Sturmbauer C (2003) Phylogeography and evolution of the Tanganyikan cichlid genus *Tropheus* based upon mitochondrial DNA sequences. *Journal of Molecular Evolution*, **56**, 54–68.

Booton GC, Kaufman L, Chandler M, Ogutu-Ohwayo R, Duan W, Fuerst PA (1999) Evolution of the ribosomal RNA internal transcribed spacer one (ITS-1) in cichlid fishes of the Lake Victoria region. *Molecular Phylogenetics and Evolution*, **11**, 273–282.

Bruford MW, Hanotte O, Brookfield JFY, Burke T (1998) Multi-locus and single-locus DNA fingerprinting. In: *Molecular Analysis of Populations* (ed. Hoelzel AR), pp. 283–336. Oxford University Press, New York.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Danley PD, Markert JA, Arnegard ME, Kocher TD (2000) Divergence with gene flow in the rock-dwelling cichlids of Lake Malawi. *Evolution*, **54**, 1725–1737.

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. *Proceedings of the Royal Society London Series B*, **268**, 1147–1152.

Fryer G (2004) Speciation rates in lakes and the enigma of Lake Victoria. *Hydrobiologia*, **519**, 167–184.

Fryer G, Iles TD (1972) *The Cichlid Fishes of the Great Lakes of Africa: Their Biology and Evolution* Oliver & Boyd, Edinburgh.

Gaggiotti OE, Lange O, Rassmann K, Gliddon C (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology*, **8**, 1513–1520.

Goldschmidt T (1996) *Darwin's Dreampond. Drama in Lake Victoria*. MIT Press, Cambridge, MA.

Goldschmidt T, Witte F, Wanink J (1993) Cascading effects of the introduced Nile Perch on the detritivorous/phytoplanktivorous species in the sublittoral areas of Lake Victoria. *Conservation Biology*, **7**, 686–700.

Greenwood PH (1965) The cichlid fishes of Lake Nabugabo, Uganda. *Bulletin of the Brittish Museum of Natural History (Zoology)*, **12**, 315–357.

Greenwood PH (1981) Species-flocks and explosive evolution. In: *The Evolving Biosphere* (ed. Forey PL), pp. 61–73. Cambridge University Press, Cambridge.

Guo S, Thompson E (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hall T (2003) *BioEdit — Biological Sequence Alignment Editor for Windows*. North Carolina State University, Raleigh.

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

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

Kellogg KA, Markert JA, Stauffer JR, Kocher TD (1995) Microsatellite variation demonstrates multiple paternity in lekking cichlid fishes from Lake Malawi, Africa. *Proceedings of the Royal Society London Series B*, **260**, 79–84.

Kluge AG, Farris JS (1969) Quantitative phyletics and the evolution of anurans. *Systematic Zoology*, **18**, 1–32.

Kocher TD, Thomas WK, Meyer A *et al.* (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences USA*, **86**, 6196–6200.

Kornfield I, Smith PF (2000) African cichlid fishes: model systems for evolutionary biology. *Annual Reviews in Ecology and Systematics*, **31**, 163–196.

Markert JA, Arnegard ME, Danley PD, Kocher TD (1999) Biogeography and population genetics of the Lake Malawi cichlid *Melanochromis auratus*: habitat transience, philopatry and speciation. *Molecular Ecology*, **8**, 1013–1026.

McElroy DM, Moran P, Birmingham E, Kornfield I (1992) REAP: an integrated program for the analysis of restriction fragment data. *Journal of Heredity*, **83**, 157–158.

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.

Meyer A, Morrissey JM, Schartl M (1994) Recurrent origin of a sexually selected trait in *Xiphophorus* fishes inferred from a molecular phylogeny. *Nature*, **368**, 539–542.

Moran P, Kornfield I (1993) Retention of an ancestral polymorphism in the mbuna species flock (Teleostei: Cichlidae) of Lake Malawi. *Molecular Biology and Evolution*, **10**, 1015–1029.

Nagl S, Tichy H, Mayer WE, Takahata N, Klein J (1998) Persistence of neutral polymorphisms in Lake Victoria cichlid fish. *Proceedings of the National Academy of Sciences USA*, **95**, 14238–14243.

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. *Proceedings of the Royal Society London Series B*, **267**, 1049–1061.

Nishida M (1991) Lake Tanganyika as an evolutionary reservoir of old lineages of East African cichlid fishes: inferences from allozyme data. *Experientia*, **47**, 974–979.

Odhiambo EA (2002) *A taxonomic study of Lake Kanyaboli cichlid fishes using morphological methods*. Masters Thesis. International Institute for Infrastructural Hydraulic and Environmental Engineering, Delft, The Netherlands.

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*. *Environmental Biology of Fishes*, **27**, 81–96.

Ogutu-Ohwayo R (1993) The effects of predation by Nile Perch, *Lates niloticus* L., on the fish of Lake Nabugabo, with suggestions for conservation of endangered endmic cichlids. *Conservation Biology*, **7**, 701–711.

van Oppen MJ, Turner GF, Rico C *et al.* (1997) Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. *Proceedings of the Royal Society London Series B*, **264**, 1803–1812.

Parker A, Kornfield I (1997) Evolution of the mitochondrial DNA control region in the *mbuna* (Cichlidae) species flock of Lake Malawi, East Africa. *Journal of Molecular Evolution*, **45**, 70–83.

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

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Rico C, Turner GF (2002) Extreme microallopatric divergence in a cichlid species from Lake Malawi. *Molecular Ecology*, **11**, 1585–1590.

Roff DA, Bentzen P (1989) The statistical analysis of mitochondrial DNA polymorphisms: chi 2 and the problem of small samples. *Molecular Biology and Evolution*, **6**, 539–545.

Rüber L, Meyer A, Sturmbauer C, Verheyen E (2001) Population structure in two sympatric species of the Lake Tanganyika cichlid tribe Eretmodini: evidence for introgression. *Molecular Ecology*, **10**, 1207–1225.

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

Sato A, Takezaki N, Tichy H, Figueroa F, Mayer WE, Klein J (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.

Schneider S, Roessli D, Excoffier L (1999) *Arlequin, Version 2.0: a Software for Genetic Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Geneva.

Seehausen O (2002) Patterns in fish radiation are compatible with Pleistocene desiccation of Lake Victoria and 14,600 year history for its cichlid species flock. *Proceedings of the Royal Society London Series B*, **269**, 491–470.

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

Seehausen O, Witte F, Katunzi EF, Smits J, Bouton N (1997b) Patterns of the remnant cichlid fauna in Southern Lake Victoria. *Conservation Biology*, **11**, 890–904.

Seehausen O, Koetsier E, Schneider MV *et al.* (2003) Nuclear markers reveal unexpected genetic variation and a Congolese-Nilotic origin of the Lake Victoria cichlid species flock. *Proceedings of the Royal Society London Series B*, **270**, 129–137.

Shaw PW, Turner GF, Idid MR, Robinson RL, Carvalho GR (2000) Genetic population structure indicates sympatric speciation of Lake Malawi pelagic cichlids. *Proceedings of the Royal Society London Series B*, **267**, 2273–2280.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Stiassny MLJ, Meyer A (1999) Cichlids of the Rift Lakes. *Scientific American*, **February**, 64–69.

Sturmbauer C, Baric S, Salzburger W, Rüber L, Verheyen E (2001) Lake level fluctuations synchronize genetic divergences of cichlid fishes in African lakes. *Molecular Biology and Evolution*, **18**, 144–154.

Sturmbauer C, Hainz U, Baric S, Verheyen E, Salzburger W (2003) Evolution of the tribe Tropheini from Lake Tanganyika: synchronized explosive speciation producing multiple evolutionary parallelism. *Hydrobiologia*, **500**, 51–64.

Swofford DL (2003) *PAUP* — Phylogenetic Analyses Using Parsinomy and Other Methods*, Version 4.0. Sinauer, Sunderland, MA.

Turner GF, Seehausen O, Knight ME, Allender CJ, Robinson RL (2001) How many species of cichlid fishes are there in African lakes? *Molecular Ecology*, **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.

Verheyen E, Salzburger W, Snoeks J, Meyer A (2004) Response to comment on 'Origin of the superflock of cichlid fishes from lake Victoria, East Africa'. *Science*, **304**, 963c.

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**, 438–450.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Witte F, Goldschmidt T, Wanink J *et al.* (1992) The destruction of an endemic species flock: quantitative data on the decline of the haplochromine cichlids of Lake Victoria. *Environmental Biology of Fishes*, **34**, 1–28.

Witte F, Goudswaard PC, Katunzi EF, Mkumbo OC, Seehausen O, Wanink JH (1999) Lake Victoria's ecological changes and their relationships with the riparian societies. In: *Ancient Lakes: Their Cultural and Biology Diversity* (eds Kawanabe H, Coulter GW, Roosevelt AC), pp. 189–202. Kenobi Publications, Ghent, Belgium.

Witte F, Msuku BS, Wanink JH *et al.* (2000) Recovery of cichlid species in L. Victoria: an examination of factors leading to differential extinction. *Reviews in Fish Biology and Fisheries*, **10**, 233–241.

Wu L, Kaufman L, Fuerst PA (1999) Isolation of microsatellite markers in *Astatoreochromis alluaudi* and their cross-species amplifications in other African cichlids. *Molecular Ecology*, **8**, 895–897.

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). *Proceedings of the Royal Society London Series B*, **263**, 1589–1598.

This work is part of Romulus Abila's PhD project investigating the niche differentiation, local adaptation and phylogenetic status of the haplochromine cichlid fauna of Lake Kanyaboli, Kenya. The laboratory work for this study has been carried out in the laboratory of Axel Meyer at the University of Konstanz, Germany. The research of Marta Barluenga and Walter Salzburger focuses on the understanding of speciation and adaptive radiation in cichlids. Johannes Engelken was doing his bachelor thesis on haplochromine cichlids in the Axel Meyer laboratory.

Appendix 1

Specimen information, GenBank accession number, sample locality and source of DNA sequence of all mtDNA sequences used in this study

Haplotype (Fig. 2)	NH	Taxon	NS	GenBank Acc. Nr.	Locality	Source	Haplotype in Verheyen *et al.* (2003)
1	1	*Haplochromis* sp. *cbrebridens/olivaceus*	1	AY226646	Lake Kivu	Verheyen *et al.* 2003	17
2	16	*Haplochromis astatodon*	8	AY226611–AY226618	Lake Kivu	Verheyen *et al.* 2003	7
		Haplochromis nigroides	2	AY226619–AY226620	Lake Kivu	Verheyen *et al.* 2003	7
		Haplochromis rubescens	1	AY226621	Lake Kivu	Verheyen *et al.* 2003	7
		Haplochromis sp. *cbrebridens/olivaceus*	5	AY226622–AY226626	Lake Kivu	Verheyen *et al.* 2003	7
3	32	*Haplochromis astatodon*	1	AY226655	Lake Kivu	Verheyen *et al.* 2003	25
		Haplochromis graueri	1	AY226656	Lake Kivu	Verheyen *et al.* 2003	25
		Haplochromis insidiae	2	AY226657, AY226658	Lake Kivu	Verheyen *et al.* 2003	25
		Haplochromis olivaceus	1	AY226659	Lake Kivu	Verheyen *et al.* 2003	25
		Haplochromis paucidens	1	AY226660	Lake Kivu	Verheyen *et al.* 2003	25
		Haplochromis scheffersi	1	AY226661	Lake Kivu	Verheyen *et al.* 2003	25
		Haplochromis sp. *cbrebridens/olivaceus*	4	AY226662–AY226665	Lake Kivu	Verheyen *et al.* 2003	25
		Haplochromis microchrysomelas	2	AY226699, AY226700	Lake Kivu	Verheyen *et al.* 2003	56
		Haplochromis olivaceus	1	AY226701	Lake Kivu	Verheyen *et al.* 2003	56
		Haplochromis paucidens	1	AY226702	Lake Kivu	Verheyen *et al.* 2003	56
		Haplochromis sp. *cbrebridens/olivaceus*	8	AY226703–AY226710	Lake Kivu	Verheyen *et al.* 2003	56
		Haplochromis vittatus	1	AY226711	Lake Kivu	Verheyen *et al.* 2003	56
		Haplochromis sp.	1	AY226728	Chambura	Verheyen *et al.* 2003	25
		Haplochromis sp.	4	AY226729–AY226732	Mugogo	Verheyen *et al.* 2003	25
		Haplochromis sp.	1	AY226733	Victoria Nile	Verheyen *et al.* 2003	25
		Haplochromis sp.	1	AF213565	Lake Edward	Nagl *et al.* 2000	25
		Yssichromis laparogramma	1	AF213522	Lake Victoria	Nagl *et al.* 2000	25
4	16	*Haplochromis astatodon*	3	AY226671–AY226673	Lake Kivu	Verheyen *et al.* 2003	47
		Haplochromis insidiae	3	AY226674–AY226676	Lake Kivu	Verheyen *et al.* 2003	47
		Haplochromis nigroides	2	AY226677, AY226678	Lake Kivu	Verheyen *et al.* 2003	47
		Haplochromis paucidens	5	AY226679–AY226683	Lake Kivu	Verheyen *et al.* 2003	47
		Haplochromis sp. *cbrebridens/olivaceus*	2	AY226684, AY226685	Lake Kivu	Verheyen *et al.* 2003	47
		Haplochromis sp.	1	AY226686	Lake Kivu	Verheyen *et al.* 2003	47
5	1	*Haplochromis paucidens*	1	AY226687	Lake Kivu	Verheyen *et al.* 2003	48
6	110	*Xystichromis phytophagus* HT01	99	AY624744–AY624842	Lake Kanyaboli	this study	n/a
		Haplochromis sp.	2	AY226759, AY226760	Lake Victoria	Verheyen *et al.* 2003	77
		Gaurochromis simpsoni	1	AF213518	Lake Nabugabo	Nagl *et al.* 2000	77
		Paralabidochromis baedlei	1	AF213519	Lake Nabugabo	Nagl *et al.* 2000	77
		Yssichromis laparogramma	1	AF213521	Lake Victoria	Nagl *et al.* 2000	80
		Enterochromis cinctus	1	AF213526	Lake Victoria	Nagl *et al.* 2000	77
		Prognathochromis venator	1	AF213537	Lake Victoria	Nagl *et al.* 2000	81
		Haplochromis sp.	1	AF213588	Lake Victoria	Nagl *et al.* 2000	87
		Ptyochromis ishmaeli 3	1	AY629399	Lake Victoria	Meyer *et al.* 1990	n/a
		Neochromis nigricans 4	1	AY629400	Lake Victoria	Meyer *et al.* 1990	n/a
		Platytaeniodus degeni 3	1	AY629401	Lake Victoria	Meyer *et al.* 1990	n/a

Appendix 1 *Continued*

Haplotype (Fig. 2)	NH	Taxon	NS	GenBank Acc. Nr.	Locality	Source	Haplotype in Verheyen *et al.* (2003)
7	1	*Ptyochromis ishmaeli* b	1	AY629402	Lake Victoria	Meyer *et al.* 1990	n/a
8	1	*Haplochromis* sp.	1	AY226761	Nawampasa	Verheyen *et al.* 2003	78
9	1	*Ptyochromis xenognathus*	1	AY629403	Lake Victoria	Meyer *et al.* 1990	n/a
10	1	*Lipochromis obesus*	1	AY629405	Lake Victoria	Meyer *et al.* 1990	n/a
11	1	*Prognathochromis longirostris*	1	AY629406	Lake Victoria	Meyer *et al.* 1990	n/a
12	2	*Xystichromis phytophagus* HT11	1	AY624948	Lake Kanyaboli	this study	n/a
		Paralabidochromis chilotes	1	AF213540	Lake Victoria	Nagl *et al.* 2000	90
13	1	*Paralabidochromis plagiodon*	1	AF213547	Lake Victoria	Nagl *et al.* 2000	91
14	1	*Yssichromis laparogramma*	1	AF213520	Lake Victoria	Nagl *et al.* 2000	89
15	43	*Xystichromis phytophagus* HT02	43	AY624843–AY624885	Lake Kanyaboli	this study	n/a
16	2	*Xystichromis phytophagus* HT07	2	AY624941, AY624942	Lake Kanyaboli	this study	n/a
17	1	*Paralabidochromis chilotes*	1	AF213539	Lake Victoria	Nagl *et al.* 2000	79
18	1	*Prognathochromis dentex*	1	AY629407	Lake Victoria	Meyer *et al.* 1990	n/a
19	1	*Xystichromis phytophagus* HT08	1	AY624943	Lake Kanyaboli	this study	n/a
20	30	*Xystichromis phytophagus* HT03	30	AY624886–AY624915	Lake Kanyaboli	this study	n/a
21	1	*Xystichromis phytophagus* HT10	1	AY624947	Lake Kanyaboli	this study	n/a
22	1	*Astatotilapia piceatus*	1	AY629408	Lake Victoria	Meyer *et al.* 1990	n/a
23	2	*Haplochromis* sp.	2	AY226718, AY226719	Rwihindi and Cohoha	Verheyen *et al.* 2003	82
24	3	*Haplochromis* sp.	3	AY226720–AY226722	Cohoha	Verheyen *et al.* 2003	83
25	3	*Haplochromis* sp.	3	AY226723–AY226725	Rweru and Cohoa	Verheyen *et al.* 2003	84
26	1	*Haplochromis* sp.	1	AY226726	Cohoha	Verheyen *et al.* 2003	85
27	1	*Prognathochromis venator*	1	AF213538	Lake Victoria	Nagl *et al.* 2000	81
28	1	*Prognathochromis paraguiarti*	1	AY629409	Lake Victoria	Meyer *et al.* 1990	n/a
29	1	*Harparogramma guiarti*	1	AY629410	Lake Victoria	Meyer *et al.* 1990	n/a
30	1	*Astatotilapia velifer*	1	AF213550	Lake Victoria	Nagl *et al.* 2000	88
31	11	*Xystichromis phytophagus* HT04	6	AY624916–AY624921	Lake Kanyaboli	this study	n/a
		Astatotilapia nubila	1	AF213524	Lake Victoria	Nagl *et al.* 2000	92
		Paralabidochromis chilotes	1	AF213525	Lake Victoria	Nagl *et al.* 2000	98
		Lipochromis melanopterus	1	AF213527	Lake Victoria	Nagl *et al.* 2000	95
		Neochromis nigricans	1	AF213528	Lake Victoria	Nagl *et al.* 2000	121
		Astatotilapia velifer	1	AF213551	Lake Victoria	Nagl *et al.* 2000	94
32	3	*Xystichromis phytophagus* HT09	3	AY624944–AY624946	Lake Kanyaboli	this study	n/a
33	1	*Paralabidochromis labiatus*	1	AF213548	Lake Victoria	Nagl *et al.* 2000	92
34	1	*Neochromis nigricans*	1	AF213544	Lake Victoria	Nagl *et al.* 2000	99
35	4	*Haplochromis* sp.	4	AY226780–AY226783	Lake Bunyoni/Mburo	Verheyen *et al.* 2003	119
36	1	*Haplochromis* sp.	1	AY226784	Lake Bunyoni	Verheyen *et al.* 2003	120
37	1	*Ptyochromis xenognathus*	1	AF213533	Lake Victoria	Nagl *et al.* 2000	110
38	1	*Ptyochromis xenognathus*	1	AF213534	Lake Victoria	Nagl *et al.* 2000	109
39	2	*Haplochromis* sp.	1	AF213581	Lake Wamala/Lake Victoria	Nagl *et al.* 2000	111
		Haplochromis sp.	1	AF213583	Lake Wamala/Lake Victoria	Nagl *et al.* 2000	112

Appendix 1 *Continued*

Haplotype (Fig. 2)	NH	Taxon	NS	GenBank Acc. Nr.	Locality	Source	Haplotype in Verheyen *et al.* (2003)
40	1	*Ptyochromis xenognathus*	1	AF213532	Lake Victoria	Nagl *et al.* 2000	113
41	1	*Ptyochromis sauvagei*	1	AF213549	Lake Victoria	Nagl *et al.* 2000	106
42	1	*Neochromis nigricans*	1	AF213545	Lake Victoria	Nagl *et al.* 2000	96
43	1	*Haplochromis* sp.	1	AY266766	Nawampasa	Verheyen *et al.* 2003	97
44	17	*Xystichromis phytophagus* HT05	17	AY624922–AY624938	Lake Kanyaboli	this study	n/a
45	2	*Xystichromis phytophagus* HT06	2	AY624939, AY624940	Lake Kanyaboli	this study	n/a
46	1	*Haplochromis* sp. *'rock'*	1	AF213541	Lake Victoria	Nagl *et al.* 2000	108
47	1	*Astatotilapia velifer*	1	AF213553	Lake Victoria	Nagl *et al.* 2000	107
48	2	*Astatotilapia nubila*	1	AF213536	Lake Victoria	Nagl *et al.* 2000	117
		Haplochromis sp.	1	AF213582	Lake Wamala/Lake Victoria	Nagl *et al.* 2000	116
49	1	*Ptyochromis xenognathus*	1	AF213535	Lake Victoria	Nagl *et al.* 2000	118
50	1	*Astatotilapia velifer*	1	AF213552	Lake Victoria	Nagl *et al.* 2000	114
51	1	*Haplochromis velvet black*	1	AF213543	Lake Victoria	Nagl *et al.* 2000	115
52	1	*Astatotilapia nubila*	1	AY629411	Lake Victoria	Meyer *et al.* 1990	n/a
53	1	*Psammochromis riponianus*	1	AF213530	Lake Victoria	Nagl *et al.* 2000	102
54	1	*Paralabidochromis baedlei*	1	AF213546	Lake Victoria	Nagl *et al.* 2000	104
55	1	*Paralabidochromis plagiodon*	1	AF213529	Lake Victoria	Nagl *et al.* 2000	105
56	1	*Ptyochromis sauvagei*	1	AF213531	Lake Victoria	Nagl *et al.* 2000	122
57	1	*Ptyochromis xenognathus*	1	AY629404	Lake Victoria	Meyer *et al.* 1990	n/a
58	1	*Haplochromis lividus*	1	AF213523	Lake Victoria	Nagl *et al.* 2000	93
	340		340				

NH, number of individuals per haplotype.
NS, number of specimens.

