African Journal of Tropical Hydrobiology and Fisheries 16: 24-32 (2018)
© Lake Victoria Fisheries Organization

DNA Evidence for a Population Bottleneck in Lake Victoria Nile perch

DORCAS LUSWETI [1 *], JOEL W. OCHIENG[2,4], JOYCE G. MAINA[2] and PETER KINYANJUI [3]

[1] *School of Biological & Physical Sciences, Moi University, P.O. Box 3900-30100 Eldoret*
[2] *Department of Animal Science, University of Nairobi, P.O. Box 30197 Nairobi*
[3] *Department of Biochemistry, University of Nairobi, P.O. Box 30197 Nairobi*
[4] *Centre for Biotechnology & Bioinformatics, University of Nairobi P.O. Box 30197 Nairobi*

[*] Corresponding author: Email, dlusweti@yahoo.co.uk

Abstract
A recent decline in Nile perch abundance in Lake Victoria is of concern to the fishing industry. To determine if this decline threatens the persistence of Nile perch, we quantified the level of genetic variation and the structure of its populations. Samples of Nile perch from lakes Victoria and Turkana were characterized by a single haplotype, while those from Lake Albert exhibited two haplotypes varying in a single T-C transition. Variation across the seven microsatellites was low with allelic richness ranging from 2.6 to 3.3. Bayesian analysis of genetic structure showed all samples belong to one population, consistent with a low genetic differentiation ($F_{ST} = 0.054$; $p \geq 0.05$). Patterns of allelic distribution and heterozygote excess suggested a recent population bottleneck. Such species are either threatened or endangered based on their low genetic diversity, lack of genetic differentiation among populations, molecular signatures of recent bottleneck, and reduced census.

Keywords: Genetic bottleneck, Genetic variation and structure, Genotyping Sequence analysis

Introduction
The extinction rates of aquatic species are amongst the highest known for any group with approximately 34% of fish species threatened with extinction over the next few decades (IUCN, 2008). Nile perch, *Lates niloticus,* was introduced into Lake Victoria to utilise the haplochromines, which were seen as fish with little economic potential (Pringle, 2005). Its population has fluctuated extensively (Taabu-Munyaho *et al*., 2016) and it was reported to have declined sharply in 2016 as a result of intensive fishing and a reduction in their average size. A cumulative decline of 43% between 1999 and 2003 was reported from Kenya (Kenya Fisheries Department, 2003) and measures to protect Nile perch have been implemented through Fisheries Regulations that prohibit catching fish below 50 cm and above 80 cm in length (Kenya Gazette, 2003).

The genetic structure of Nile perch in Lake Victoria is unclear and morphometric investigations (Hauser *et al*., 1998; Ward *et al.,* 2005) have failed to identify the species unambiguously, making the taxonomic status of the stock uncertain. Allozyme variation revealed the presence of one species, *L. niloticus,* but with a possible presence of *L. macrophthalmus* (Hauser *et al*., 1998). The most recent study of Nile perch genetics in Victoria (Mwanja and Mwanja, 2008) suggested some genetic structure within a single species in the lake, based on only two microsatellite loci. Information on the population genetic structure and the amount and distribution of genetic variation in a species thought to be declining can help predict its survival and evolutionary potential because these generally correlate positively with genetic diversity (Frankham *et al,*. 1999). These indices may also provide insights into the mechanisms of establishment and spread, and in designing conservation management systems.

This study quantified the genetic variation and spatial genetic structure from different localities using highly variable microsatellite loci. Genetic diversity indices (number of alleles and heterozygosities) as well as genetic differentiation were estimated using Bayesian methods on microsatellite loci. Species were

also diagnosed using a universal 'barcode'. The sequence of mitochondrial Cytochrome *c* Oxidase subunit I (CO 1) has been proposed as a global bioidentification system for animals (Hebert *et al.,* 2003) but some controversy surrounds this approach (Lipscomb *et al.,* 2003; Moritz and Cicero, 2004). The number of species that can be inferred from species-specific haplotypes obtained by sequencing the CO 1 locus was investigated to give additional insights into the genetic structure, focusing on the possible higher-level divergence.

Methods

Population sampling and DNA extraction

This work was carried out in the Kenyan waters of Lake Victoria, which include the Nyanza Gulf, a semi-enclosed body of water with an average depth of 8 m (maximum = 45 m). Fish specimens were collected from five points, one of them (Dunga) at the head of the gulf, one near the Ugandan border to the north (Sio Port) and the remaining three (Mbita, Mfangano Island and Karungu Bay) near the southern entrance to the gulf (Figure 1).

Figure 1. Lake Victoria showing the Nyanza (Winam) Gulf and the areas from which samples were obtained for this study (see text).

Between 30 and 40 individuals were collected at each site using nylon monofilament gill nets of various mesh sizes (6.2 to 25 cm, knot to knot). In addition to these, 11 and 10 individuals were collected from Lakes Albert (Uganda) and Turkana (Kenya) respectively. A sample of muscle tissue for molecular analyses was excised from each fish using a new set of scalpels, and stored in absolute ethanol, after which the fish were archived at the National Museums of Kenya. A portion of the tissue sample required for DNA extraction were transferred into a lysis solution and digested with proteinase *K* at 55°C for 4 hours. After digestion, DNA was extracted following standard procedures (Sambrook *et al.,* 1989) and the pellet dissolved in 300 µl of sterile water.

Sequencing of CO 1 region

Two primers were designed following multiple alignments of available CO 1 sequences for haplochromine cichlids held in public databases, National Centre for Biotechnology Information (NCBI). The following primers were designed from a best fit consensus sequence flanking the target region: 5'-TCAACCAACCACAAAGACATCGGCAC-3' and 5'-ACTTCAGGGTGTCCAAAGAATCAGAA-3'.

Polymerase Chain Reaction (PCR) for the CO 1 region was performed in 20 µL volumes comprising 1x PCR buffer (10 mM Tris–HCl pH 8·3, 50 mM KCl, 0·001% gelatin (Sigma, St Louis, MO, USA), 0·25% Nonidet P40 (BHD, Poole, UK) and 2 mM $MgCl_2$) and contained approximately 1 ng of genomic DNA, 0.125 mM of each dNTPs, 0.3 µM of each primer, and 0.5 Units of Platinum Taq (Invitrogen). All amplifications were carried out on a GeneAmp 9700 with an initial denaturation of 7 minutes at 95°C, followed by 32 repeated cycles of denaturation at 95°C, annealing at 55°C and an extension of 1 minute at 72°C. A final extension of 72°C for 10 minutes was applied to all reactions.

Sequencing reactions

Following confirmatory agarose gel electrophoresis using small quantities, the remaining PCR products were purified on a QuiAquick PCR kit (Quiagen, Germany) according to the manufacturer's protocol. Purified PCR samples were sequenced using BigDye Treminator v3 Cycle Sequencing Ready Reaction Kit (Applied Biosystems) and primers previously used in PCR reactions: 5'-TCAACCAACCACAAAGACATCGGCAC-3' and 5'-ACTTCAGGGTGTCCAAAGAATCAGAA-3'.

The sequenced products were separated on a 3730 DNA analyser (Applied Biosystems) at the sequencing

facility of the International Livestock Research Institute, Nairobi, and raw data were imported into Chromas v 1.5 (Technelysium Pty. Ltd) for editing.

Sequence data analysis

Raw data were edited using Chromas V1.5 (Technelysium) and sequences aligned using ClustalX (EMBL). The aligned sequences in this study had a length of 600 characters with no indels (insertion or deletion events). Sequences obtained were deposited in NCBI under accession Nos. KC907307 – KC907316. There being no variation at all among samples, except at a single nucleotide position among samples from Lake Albert, we aligned Lake Victoria sequences against those of other *Lates* species, *L. calcarifer* from Australia and *L. uwisara* from Myanmar, from NCBI to give insights into the level of variation expected at this locus in the *Lates* group.

Genotyping at SSR loci

Nine fluorescent microsatellites (SSRs) characterised in haplochromine cichlids and another eight characterised in gilthead seabream were amplified in this study. However, only seven, OSU13d, OSU16d, OSU19t, OSU20d, OSU22d (Wu *et al*., 1996) and SaGT26, SaGT32 (Batargias *et al*., 1999) were suitable for analysis, as assessed from data completeness, visual inspection, null allele tests and linkage disequilibrium test results.

PCR was performed as described for CO 1, except that the reaction volume for microsatellite amplifications was 10 µL, and annealing temperatures varied with primer. For each sample, one microlitre of the PCR products were separated on a 3730 DNA analyser at the sequencing facility of the International Livestock Research Institute, Nairobi, and raw data were imported into ABI Prism GeneMapper Software v 3.0 (Applied Biosystems) for size calling. Duly size-called diploid allele size data from SSRs were exported to an Excel spreadsheet for statistical analyses.

The Lewontin-Krakauer (LK) test (Lewontin and Krakauer, 1973) was used to examine whether the SSR loci used evolved neutrally by testing for selection at each SSR locus. Lewontin and Krakauer (1973) argued that the variance in F_{ST} among loci under neutral expectations should be approximately

$$\sigma^2_{exp} = kF_{ST}^2/(n-1)$$

where σ^2_{exp} is the expected variance in F_{ST}, n is the number of subpopulations sampled, and k is a specific constant for the underlying allele frequency distribution among subpopulations. Using simulations, they concluded that if $k \leq 2$ for loci governed by drift alone, i.e., if $k > 2$, then there is evidence of selection. Due to reservations regarding the LK test, we also used a pairwise test between two allopatric geographic populations (Turkana and Albert; where the number of sampled populations n were = 2) to complement the global LK test in which n was > 2. Tests were conducted by substituting the observed variance with the expected variance in the equation above and calculating for k. These tests for neutrality were done *a posteriori*, having obtained other results that required further tests to ascertain this neutrality.

Statistical analysis of genetic variation and structure

Gene diversity indices, quantified as heterozygosities, number of alleles and allelic richness per population (Nei, 1987), and the inbreeding coefficient (F_{IS}) among Nile perch geographic populations were calculated using GENEPOP on the web, version 3.3 (Raymond and Rousset, 1995). Levels of population genetic differentiation, with F_{ST} being the infinite allele model of mutation (Weir and Cockerham, 1984) was estimated in GENEPOP and their significance assessed by the log-likelihood exact *G*-test on genotypes (Goudet, 2001). Since the rate of mutations and the evolution model at the SSRs used here was not known, the stepwise mutational model, R_{ST} (Slatkin, 1995), was also used in estimating the levels of differentiation. In each analysis, possible non-random allelic associations between pairs of loci were assessed within each population, and globally, using Fisher's exact test on contingency tables. The significance level for each test was adjusted using sequential Bonferroni procedure (Rice, 1989). The confidence level for the null hypothesis was set at 99%.

Genetically homogeneous clusters of individuals were determined using a Bayesian model-based method implemented in the programme STRUCTURE Version 2.1 (Pritchard *et al*., 2000), which apportions population ancestry to individuals based on the model chosen, i.e., *K* value. This was done without assuming predefined structure. To simulate a more likely number of clusters (*K*), an alternative optimization criterion-delta *K* was used; it relates to the second order rate change in the log probability of the data (Evanno *et al.,* 2005). A hierarchical approach was used to

successively analyze structure: structure was initially examined for the entire dataset and genetically homogeneous subsets identified, then further substructure was probed using STRUCTURE. The following model parameters were adopted: 10,000 iterations each for the "burn-in" and Markov Chain Monte Carlo (MCMC) stages and a "correlated allele frequency model" for ancestry. Both "admixture" and "no admixture" models were investigated when testing for structure within populations. Ten runs for each K (for K = 1-10) were used to calculate means and standard deviation (SD) for posterior probability of the data for a given K (L(K)) and delta-K. These parameters provided generally conservative optimization criteria for replicate runs when detecting populations. Presence of structure was assumed when plots of delta-K provided evidence of a clear peak for K values (Evanno *et al*., 2005).

To validate a lack of structure, assignments of individuals at the most-likely K values were investigated further. Symmetry in the proportion of samples assigned to each population and an apparent admixture of most individuals was interpreted to mean an absence of genetic structure, following Pritchard and Wen (2004). The proportion of each population ancestry for the number of populations set by the K value was used to classify individuals. Individuals were considered well assigned if they possessed >90% of ancestry from a single population, otherwise they were considered admixed. An incorrect assignment was concluded when the predefined source population showed the lowest likelihood value of being the source.

Results

Results of CO 1 sequence variation

The amplified region of CO 1 had a length of 600 bp with no indels. Comparison among samples from Lake Victoria showed no variation; all individuals were characterized by the same haplotype (Figure 2). Interestingly, Nile perch from Lake Turkana were also fixed for the same haplotype found in Victoria but fish from Lake Albert exhibited two haplotypes, one unique while the other found in Lake Victoria. The two haplotypes varied at a single T-C transition. Comparison showed that sequence of Nile perch sampled at an undisclosed location in Tanzania (DQ108019) was fixed for the same haplotype observed in Lake Victoria samples. This specimen was almost certainly from Lake Victoria since Nile perch does not occur anywhere else in Tanzania.

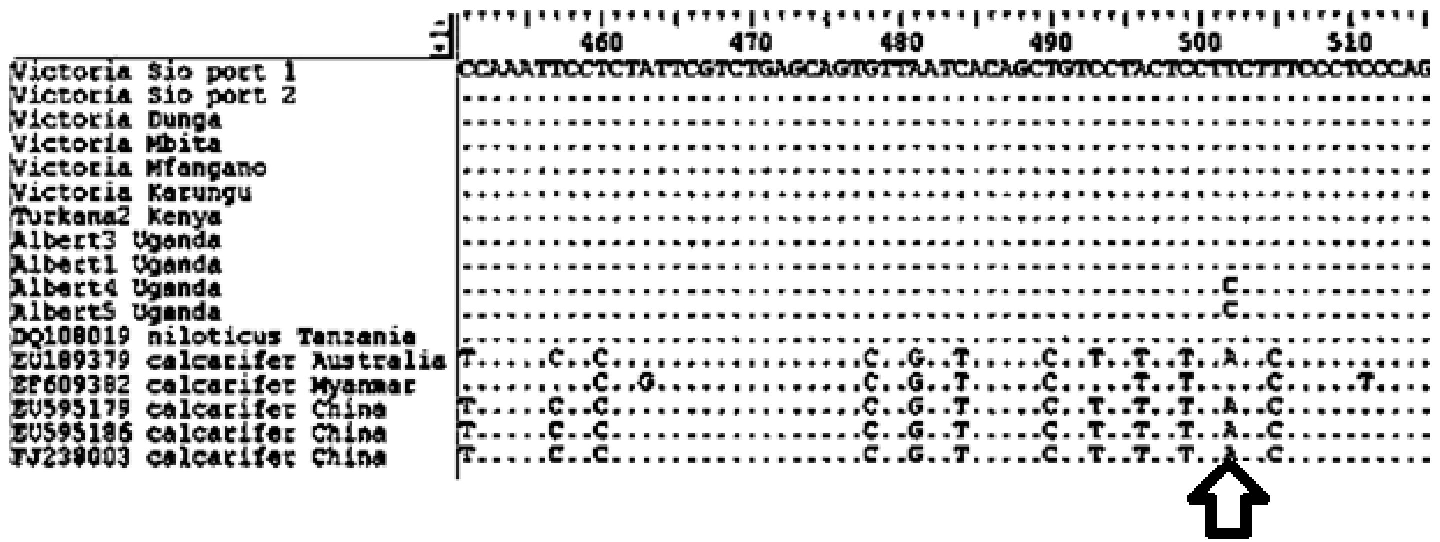

Figure 2. CO 1 Sequence positions 455-515 of Nile perch samples showing no variation among samples of Lake Victoria. All samples from Lake Victoria are fixed for the same haplotype (characterized by a T), including that from Tanzania. Lake Albert has two haplotypes separated by a T-C transition (shown by arrow). Victoria perch differ from Asian and Australian Barramundi by over 90 synapormophies.

A comparison between sequences of Nile perch to those of barramundi *L. calcarifer* available in NCBI showed 94 synapomorphies differentiating the two species. However, there was low variation within barramundi, with only two transitions differentiating a sample from Western Australia from those of South China. At the

two characters where the Australian sample showed variation from those of China, the nucleotides of Australian barramundi were similar to those observed in fish from Lake Victoria. The barramundi sample from Myanmar differed from all other barramundis at 53 nucleotides, and was also the most differentiated from Lake Victoria samples, differing at 99 nucleotides. This seems to support the view that the Myanmar population belongs to the recently-described species *L. uwisara* (Pethiyagoda and Gill, 2012). The diagnostic nucleotide position at which Nile perch of Lake Albert varied had an interesting result; the T-C transition clustered three Lake Albert samples on one hand and the rest of the Lake Albert, Lake Victoria (including the Tanzanian specimen), Lake Turkana, and Myanmar barramundi on the other. Australian and Chinese barramundi had an "A" nucleotide.

Results of Microsatellite variation

The variance of F_{ST} among loci in an n = 2 model (Lakes Albert and Turkana only) was very low ($\sigma^2 = 0.0001$), while the mean (multilocus) F_{ST} was 0.02 (hence F_{ST}^2 was = 0.0004). In the two-population model, from the equation $\sigma^2_{exp} = kF_{ST}^2/(n-1)$, the expected variance (σ^2_{exp}) = kF_{ST}^2, because (n-1) = 1. Accordingly, k was = 0.25 for the loci analysed. Neutral expectation is that $k < 2$ for loci governed by drift alone. Repeating the test for global F_{ST} assuming up to seven subpopulations ($2 \geq n \leq 7$) all returned values of $k < 2$ ($k = 0.25$ when n = 2; $k = 1.75$, n = 7). Therefore, the loci apparently evolved neutrally in the populations being analysed.

Low number of alleles with excess of heterozygotes

Genetic diversity across the seven microsatellites was considerably low with the average number of alleles per locus = 5 (Table 1). H_O was consistently and significantly higher than H_E for all populations. Tests for non-random segregation of alleles showed that the markers used largely segregated independently and were representative of multiple genomic regions. The average F_{IS} across loci per population ranged from -0.872 to -0.469 (negative values) and were all significantly different from zero (Table 1). Only 0.18 % of total variation was partitioned between populations and 2.3% within a population and most variation (95.9%) resided within individuals.

Table 1. Genetic diversity indices (quantified as heterozygosities and allelic richness) and inbreeding coefficients among Nile perch geographic populations. The observed HE values are consistently higher than those expected under Hardy-Weinberg, showing a transient heterozygote excess.

Population	Population size	Loci	Exp. Hz	Obs. Hz.	No of Alleles	Allelic richness	F_{IS}
Karungu	29	7	0.530	0.837	3.29	3.0	-0.60
Sio Port	33	7	0.520	0.861	3.14	2.7	-0.67
Mfangano	27	7	0.570	0.857	3.86	3.3	-0.52
Dunga	24	7	0.507	0.863	3.00	2.6	-0.73
Mbita	29	7	0.503	0.852	3.14	2.7	-0.72
Turkana	11	7	0.477	0.857	2.29	2.3	-0.87
Albert	10	7	0.518	0.743	2.71	2.7	-0.47
All Populations	163	7	0.538	0.848	5.0	N/A	N/A

Patterns of allelic distribution in all the seven loci studied suggested an eroded genetic variation, with all loci showing fewer alleles relative to a large allele size range. For example, locus OSU20d had four alleles in Karungu population, but the range of allele sizes was 60 (smallest allele = 145; largest allele = 205) (Figure 3). Most loci were deficient in the proportion of rare alleles, having frequencies above 5%.

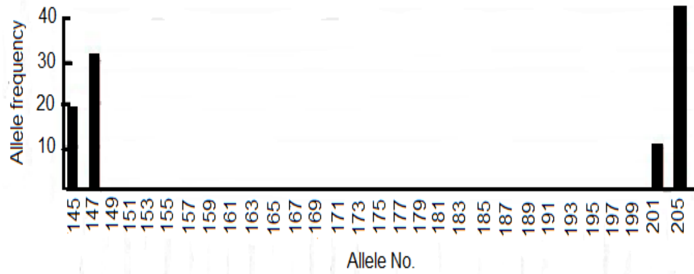

Figure 3. Frequencies of alleles observed at OSU20d locus in the Karungu predefined population.

This population has only four alleles but with a very wide allelic range, i.e., 145 to 205. The alleles occur at 145, 147 and again at 201 and 205, but with no alleles at all between 149 and 199. Apart from low allelic richness, this characteristic is diagnostic of bottleneck events (reduction in the number of alleles relative to the total allele size range; Garza and Williamson, 2001).

No site-specific alleles were identified at any of the loci studied. Genetic differentiation between predefined populations was considerably low, ranging from F_{ST} = 0.0024 (Sio and Karungu) to F_{ST} = 0.054 (Mbita and Dunga) (Table 2).

Inspection of the delta-*K* plot for models with a range of *K* values from one to 10 for the entire dataset revealed the highest peaks at $K = 2$ and another at $K = 4$. The peak at $K = 4$ was stronger, indicating that the highest hierarchy of structure had four populations when all samples were considered. Despite the stronger statistical support for a four clusters model, this outcome may have been artefactual because in all tested models ($K = 1$ to 10), using the default settings or model parameters, individuals showed no structure; instead, all individuals were symmetrically admixed (Figure 4), a phenomenon interpreted to code for an absence of genetic structure (Pritchard and Wen 2004).

Discussion

This study quantified the level of genetic variation and structure within and between populations of Nile perch in Lake Victoria, compared with some samples from lakes Albert and Turkana. There was no variation in fish from the Kenyan Gulf at a usually variable and diagnostic portion of CO 1, and a single character variation in fish from Lake Albert. Interestingly, samples from Lake Turkana were not significantly differentiated from those of lakes Albert and Victoria (essentially the same stock) in spite of the fact that Lake Turkana has been isolated from the Nile for about 7,500 years (Beadle, 1981).

A comparison of the Nile perch with Indo-Pacific *Lates* species (*L. calcarifer* and *L. uwisara*) indicated a considerable between species variation (94 nucleotides). Analysis of neutral genetic variation in the same samples showed a considerably low level of genetic variation and an absence of genetic structure, where all individuals were symmetrically admixed in a Bayesian framework. An analysis of molecular variance (AMOVA) indicated that in Lake Victoria nearly all the variation (95.9%) resided within individuals, with only 2.3% being partitioned within populations, and a paltry 0.18% between geographically predefined populations. Three evolutionary scenarios could explain the absence of variation at CO 1 and the unexpectedly low level of genetic differentiation at the otherwise variable SSR loci, (1) a recent selective sweep affecting large portions of the Nile perch genome, (2) long range gene flow in a highly mobile species, or (3) a recent bottleneck, consistent with an eroded genetic variation and absence of genetic structure.

Table 2. Genetic differentiation (F_{ST}) between predefined populations. A distance of 0.00 would mean a very close genetic similarity, while 1.00 would mean complete genetic separation between groups, where no single alleles is shared. All FST values are below 0.05, suggesting a very low genetic differentiation (absence of genetic structure).

	Sio	Mfangano	Dunga	Mbita	Turkana	Albert
Karungu	0.0024	0.0024	0.0048	0.0024	0.0024	0.0024
Sio Port		0.0024	0.0024	0.0024	0.0024	0.0024
Mfangano			0.0286	0.0024	0.0333	0.0024
Dunga				0.0545	0.0414	0.0024
Mbita					0.0381	0.0024
Turkana						0.0024

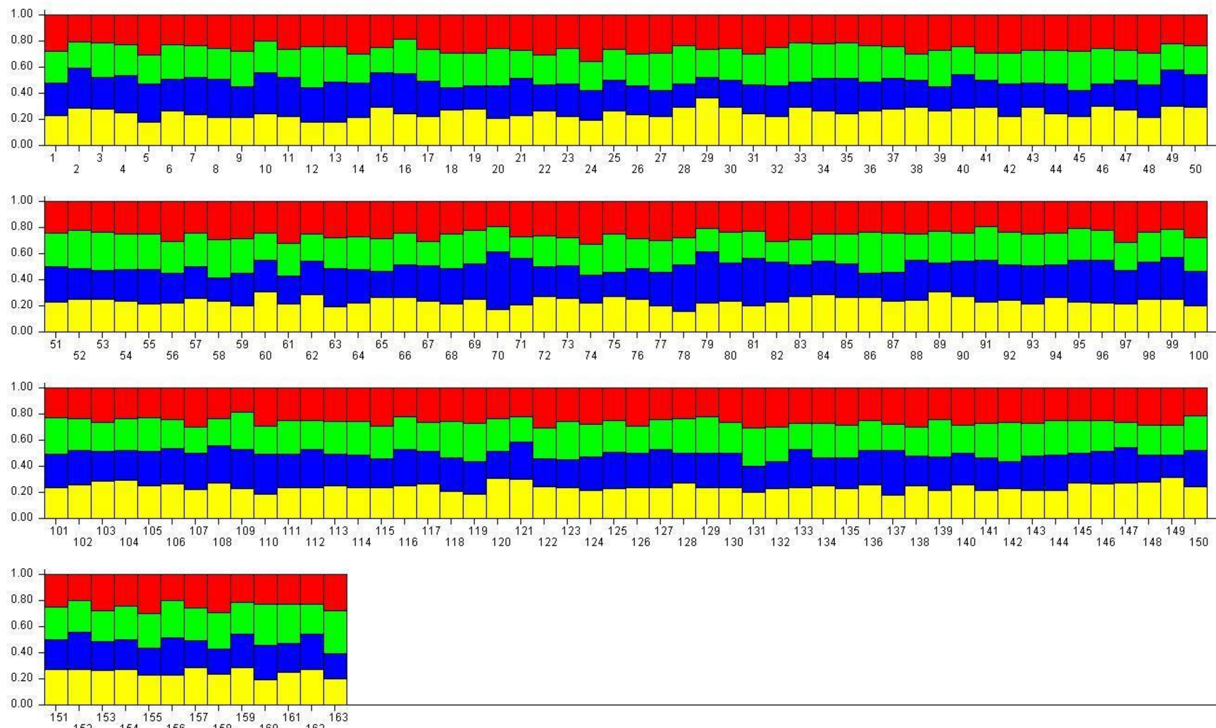

Figure 4. Ancestry assignments for 163 individuals from Lake Victoria determined by STRUCTURE using a K = 4 model. The *x*-axis is the separate individuals, while the *y*-axis shows the proportion of ancestry for each group as represented by the four different colours. Each individual appears admixed having mixed colours; the presence of structure would mean that some individuals (vertical bars) would have only one colour, with others having alternative colours.

The resolution power of the CO 1 locus has low species resolution power in fish because strong purifying selection expunges the most deleterious non-synonymous mutations within populations over time and before divergence of species (Ward and Holmes, 2007). The term “selective sweep” describes the erosion of genetic variation around a genomic region that has undergone a recent and strong positive selection, such as the rise of a favoured mutation. In this situation, genetic variation at loci around such a locus would be expected to show considerably reduced genetic variation, as observed in this study. Under a purifying selection regime, most observed variations will be recent, apomorphic, and will often be found in a small number of individuals implying they are deleterious variants not yet selected against.

This reasoning would seem to suggest that the single character variation observed among a few

samples from Lake Albert is either another *Lates* species or a recent deleterious variant waiting to be expunged by selection. The total lack of variation suggests that there is only a single species in Lake Victoria. However, considering that CO 1 has failed to accurately resolve fish species differences in other studies due to strong purifying selection (Ward and Holmes, 2007), these results have to be interpreted cautiously. In this study, only two characters differed between barramundi from Australia and those from China, while barramundi from Myanmar (now known to be a different species, *L. uwisara*) differed from the other barramundis by 53 characters.

Under the Infinite Allele Model (IAM) a population that has undergone a recent reduction in its effective size is expected to have a reduced number of alleles and an excess of heterozygosity (Maruyana and Fuerst, 1985; Cornuet and Luikart, 1996), indicative of a bottleneck effect. Another common signature is a reduction in the number of alleles relative to the total allele size range (Garza and Williamson, 2001).

In this study, all populations showed an excess of heterozygotes ($H_O > H_E$), consistent with a case of hybridization or a recent bottleneck. Many of the loci in this study, as exemplified by the locus OSU20d (Figure 3), showed a tendency to have few alleles within the large allele ranges observed. For example, locus OSU20d (Figure 3) had only four alleles within a 60 base-pair range (149-205), suggesting that many of the alleles expected within this range were eroded following a bottleneck.

It would be expected that this apparent bottleneck effect would be seen only in Lake Victoria populations and not in the source lakes, Albert and Turkana. This lack of genetic variation could be explained by the prolific reproduction of Nile perch in a new unexplored environment in the 54 years since introduction resulting in a full post-bottleneck recovery. Inbreeding in the Nile perch may also account for the reduction of genetic variation.

This work was initiated by recent reports that Nile perch has declined because of overfishing. The effective management of natural populations requires a clear definition of biological differences among and within species. Results obtained in this study confirm that the Nile perch of Lake Victoria belong to only one species *Lates niloticus*. Like its parent population in Lake Albert, and those in Lake Turkana, it has extremely low genetic variation, evidenced by both a mitochondrial gene and nuclear multilocus genotypes. Theoretically, such a low genetic variation points to low evolutionary potential for its populations so, hypothetically, Nile perch is genetically threatened. It could be argued that they were introduced into Lake Victoria only a few decades ago, which is not long enough to recover from the bottleneck-like effects of introducing few individuals. This argument can be countered by the fact that its parent population in Lake Albert is genetically identical and there is no evidence that they suffered a bottleneck effect. Similarly, fish in Lake Turkana have not differentiated even after a long period of isolation, which also counters the bottleneck argument.

Acknowledgements

This work was supported by Lake Victoria Fisheries Organization (LVFO) through the Implementation of Fisheries Management Program I (IFMP) to D. Lusweti, and the National Council for Science, Technology and Innovation Fellowship to J.W Ochieng. We thank the government of Kenya through the State Department of Fisheries for logistical support during sampling for this study.

References

Batargias, C., Dermitzakis, E., Magoulas, A. and E. Zouros (1999). Characterization of six polymorphic markers in gilt-head seabream, *Sparus aurata* (Linnaeus 1758). *Molecular Ecology* **8**:897-898.

Beadle, L.C. (1981). *The Inland Waters of Tropical Africa,* 2nd edition. Longman, London.

Cornuet, J.M. and Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**:2001-2014.

Evanno, G., Regnant, S. and J. Goudet (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**:2611-2620.

Frankham, R., Lees, K., Montgomery, M., England, P.R., Lowe, E.H. and Briscoe, D.A. (1999). Do population size bottlenecks reduce evolutionary potential? *Animal Conservation* **2**:255-260.

Garza, J.C. and Williamson, E.G. (2001). Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* **10**:305-318.

Goudet, J. (2001). A computer program to calculate F-statistics. *Heredity* **86**:485-456.

Hauser, L., Carvalho, G. R., Pitcher, T. J. and R. Ogutu-Ohwayo (1998). Genetic affinities of an introduced

predator: Nile perch in Lake Victoria, East Africa. *Molecular Ecology* **7**:849-857.

Hebert, P.D.N., Cywinska, A., Ball, S.L. and de Waard J.R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London*, B **270**:313-322.

IUCN (2008). Red list of threatened species. International Union for the Conservation of Nature, Cambridge, UK.

Lipscomb, D., Platnick, N. and Wheeler, Q. (2003). The intellectual content of taxonomy: a comment on DNA taxonomy. *Trends in Ecology and Evolution* **18**:65-66.

Kenya Fisheries Department (2003). Annual Report

Kenya Gazette (2003). Fisheries Act CAP 389 Revised of 1991: Legal Notice No. 214, The Fisheries (Prohibition) Regulations, 2003. Section 2(a). Government Printer, Nairobi: p. 712.

Lewontin, R.C. and Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**:175–195.

Maruyana, T. and Fuerst, P.A. (1995). Population bottlenecks and non-equilibrium models in population genetics. II. Number of allele in a small population that was formed by a recent bottleneck. *Genetics* **111**:675-689.

Moritz, C. and Cicero, C. (2004). DNA barcoding: promise and pitfalls. *PLoS Biology* **2**(10): e354.

Mwanja, M.T. and Mwanja W.W. (2008). Preliminary results of the analysis of the population structure of the Nile perch (*Lates niloticus* Linnaeus, 1756) of Lake Victoria using microsatellite markers. *African Journal of Biotechnology* **7**:2885-2892.

Nei, M. (1987). *Molecular Evolutionary Genetics.* Columbia University Press, New York.

Pethiyagoda, R. and Gill A.C. (2012). Description of two new species of sea bass (Teleostei: Latidae: *Lates*) from Myanmar and Sri Lanka. *Zootaxa* 3314: 1-16.

Pringle, R.M. (2005). The origins of the Nile perch in Lake Victoria. *BioScience* **55**:780-787.

Pritchard, J.K., Stephens, M. and P. Donnelly (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**:945-959.

Pritchard, J.K. and Wen, W. (2004). Documentation for STRUCTURE Software: Version 2. Available at http://pritch.bsd.uchicago.edu.

Raymond, M. and Rousset, F. (1995). An exact test for population differentiation. *Evolution* **49**:1280-1283.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* **43**:223-225.

Sambrook, J., Fritsch, E.F. and T.M. (eds.) (1989). *Molecular Cloning: A Laboratory Manual,* 2nd edition. Cold Spring Harbor Laboratory Press, New York.

Slatkin, M. (1995). Hitchhiking and associative overdominance at a microsatellite locus. *Molecular Biology and Evolution* **12**:473-480.

Taabu-Munyaho, A., Marshall, B.E., Tomasson, T. and Marteindottir, G. (2016). Nile perch and the transformation of Lake Victoria. *African Journal of Aquatic Science* **41**:127-142.

Ward, R.D., Zemlak, T.S., Ines, B.H., Last, P.R. and P.D.N. Hebert (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society of London* B **360**:1847-1857.

Ward, R.D., and Holmes B. (2007). Analysis of nucleotide and amino acid variability in the barcode of cytochrome C oxidase in fishes. *Molecular Ecology Notes* **7**:899-907.

Weir, B.S. and Cockerham C.C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**:1358-1370.

Wu, L., Booton, G.C., Kaufman, L., Chandler, M. and Fuerst, P.A. (1996). Use of DNA microsatellite loci to identify populations and species of Lake Victoria haplochromine cichlids. *In*: E.M. Donaldson and D.D. MacKinlay (eds.), *Aquaculture Biotechnology*. Proceedings of the International Congress on the Biology of Fishes, American Fisheries Society: p. 105-114.