Zoological Studies
a SpringerOpen Journal

RESEARCH   **Open Access**

Evolutionary history of Nile perch *Lates* sp. inferred from mitochondrial DNA variation analyses

Matthew Tenywa Mwanja[1*], Vincent Muwanika[2], Charles Masembe[3], Sylvester Nyakaana[3] and Wilson Waiswa Mwanja[4]

Abstract

Background: Evolutionary histories of aquatic species are often characterized by distinct patterns of genetic variation, which in part reflect drainage evolution. In the present study, the consequences of paleo-environmental changes on patterns of genetic variation of the mitochondrial DNA control region in Nile perch *Lates* sp. sampled from seven water bodies across the African continent were investigated.

Results: In a total sample of 124 individual sequences, 37 distinct haplotypes were observed, and 78.4% of these haplotypes were location specific. Haplotypes were found to cluster into two major groups, one composed of individuals sampled from East Africa and another from West Africa, with no haplotypes shared in between.

Conclusion: These lineages may have developed in geographical isolation during the Pleistocene and have remained largely allopatric without gene flow ($Nm = 0.0$) since that time. There was also evidence that both of these genetic lineages have undergone recent population expansions. We interpret these results in light of the recent evolution of Africa's modern drainage network.

Keywords: Africa drainage system; Nile perch; Evolution; mtDNA

Background

Construction of spatio-temporal details of the evolutionary histories of fishes can reveal key events of drainage evolution (Cotterill and de Wit 2011; Goodier et al. 2011). This is because the biogeography of extant freshwater fishes represents responses to landscape evolution. Since the ecology of these fishes incorporates responses to habitat change, these responses leave corresponding signatures in the genetics of these fishes (Balon 1974; Lévêque 1997; Zengeya et al. 2011). Therefore, the evolutionary history of each aquatic species is characterized by a distinct pattern of genetic variation, which includes persisting relics of drainage evolution (Craw et al. 2007).

Goodier et al. (2011) studied the phylogeography of the African tigerfish of the genus *Hydrocynus* and found that the diversity and distribution of *Hydrocynus* reflect a complex history of vicariance and dispersal. Moreover, range expansions in this particular species testify to

changes in drainage basins, and the principal divergence events in *Hydrocynus* have closely interfaced with evolving drainage systems across tropical Africa. Agnèse et al. (1997) using molecular marker polymorphisms found that the Nile tilapia *Oreochromis niloticus* with an almost similar natural distribution on the African continent as the Nile perch, *Lates* sp., can be divided into two groups: one in East Africa and the other comprising the Nile and West African populations. They attributed the subdivisions to tectonic movements and volcanic activity in East Africa that played a major role in the formation of the different basins on the continent, although they also postulated that the alternating dry and humid phases across the continent in the Pleistocene may have had large influences on the expansion and contraction of populations and their speciation.

Nile perch occurs naturally in the Afro-tropical region of the African continent, in the major river basins of the Nile, Congo, Volta, Niger, and Senegal Rivers (Fishbase 2010). It also occurs naturally in Albert Lake in Uganda, Turkana Lake in Kenya (Harrison 1991), Chamo, Abaya, and Gambella Lakes and the Baro River in Ethiopia

Springer

(Golubtsov and Habteselassie 2010), Lake Chad in Central Africa, Marriott Lake in the Nile Delta in Egypt, and in scores of small rivers in tropical West Africa (Moritz and Linsenmair 2005). Several introductions of Nile perch into non-native waters resulted in successful establishment as reported in East Africa in Victoria, Kyoga, and Nabugabo Lakes (Ogutu-Ohwayo 1993, 1990). For effective management and conservation of this fishery, it is prudent to understand its evolutionary history throughout its extant range and also in environs into which it was introduced.

The rate of change in genetic markers allows investigations of diversity on temporal scales resulting from recent history to deep evolutionary time (Nielsen 1998). For conservation efforts, the evolutionary history of the Nile perch will assist in deciphering historical congruencies responsible for the genetic subdivision throughout its extant range, in both endemic and introduced environs. The evolutionary history can also be used to test congruence between distributional histories and paleo-environmental changes that may have led to biological evolutionary diversification (van Tuinen et al. 2004; Dominguez-Dominguez et al. 2008) and address broader ecological and evolutionary processes like changes in populations, demographics, and genetics over time (Ramakrishnan and Hadly 2009). In addition, it can also describe how fish populations expanded or contracted in environments they occupy and provide possible successful establishment mechanisms in environs into which they were introduced (Tzeng et al. 2006). Influences of patterns of post paleo-environmental changes on species distributions and historical demographics of the Nile perch have not yet been studied. For effective management of an exploitable species such as the Nile perch, it is important to define its taxonomic status and historical distributions throughout its extant range and delineate its different evolutionary lineages that may warrant separate management strategies and should be considered different Evolutionarily Significant Units (ESUs) (Soltis and Gitzendanner 1999).

In this study, we analyzed the genetic diversity of mitochondrial (mt)DNA and combined phylogeographic, phylogenetic, and mismatch analyses to infer spatial dynamics of the distributional and demographic history of the Nile perch throughout its native and introduced ranges on the African continent. Based on the results, we provide proposals for effective management to ensure its continued and sustainable exploitation in the future.

Methods
Sample collection and DNA extraction
Nile perch samples were collected from Albert, Turkana, Kyoga, Nabugabo, and Victoria Lakes in East Africa, Chamo Lake in Ethiopia, and the Senegal River of West Africa (Figure 1). In the field, fin clips (about 2×4 cm each) were taken from fish samples, and these were stored in separate 5-ml vials filled with absolute ethanol. The ethanol in the vials was drained off 1 h later, and the vials were refilled with absolute ethanol; in the field, samples were kept at room temperature and in the laboratory at −80°C before DNA extraction.

Total genomic DNA was extracted using the Dneasy™ tissue kit (Qiagen, Germantown, MD, USA) according to the manufacturer's instructions. This involved maceration of part of the fin tissue, cell lysis, digestion with proteinase K at 55°C for $2 \sim 3$ h, followed by binding of the DNA, and then washing and eluting of the DNA. A negative control that contained no tissue was also prepared. DNA extracts were dissolved in 300 μL of elution buffer and stored at −20°C. To test for success of DNA extraction, 4 μL of total genomic DNA was electrophoresed on a 2% agarose gel stained with ethidium bromide and then visualized under ultraviolet light for clear DNA bands.

DNA amplification and sequencing
The mtDNA D-loop region of Nile perch was amplified using primers LN20 (ACCACTAGCACCCAAAGCTA) and HN20 (GTGTTATGCTTTAGTTAAGC) respectively located in the proline and phenylalanine transfer RNA genes (Bernatchez and Danzmann 1993). Polymerase chain reactions (PCRs) consisted of 5 μL genomic DNA, 25 μL AmpliTaq Gold@ Master Mix, 5 μL of a 10 μM solution of each of the two primers, and double-distilled H_2O added to a final volume of 50 μL. Touchdown PCR conditions were used for amplification, which included an initial denaturation step of 10 min at 95°C, followed by 1 cycle each of 94°C for 1 min, 67°C ~ 51°C for 1.5 min, and 72°C for 2 min for each of the annealing temperatures starting at 67 and stepping down by 2°C from each preceding cycle to 51°C. This was followed by 25 cycles of 94°C for 1 min, 61°C for 1.5 min, and 72°C for 2 min with a final extension at 72°C for 7 min. Double-stranded PCR products were cleaned following the manufacturer's protocol (QIAquick, Qiagen).

D-loop sequence analyses
Sequencing reactions were carried out by Macrogen (Seoul, South Korea) using an Applied Biosystems 3730 × l DNA analyzer (Carlsbad, CA, USA). Samples were sequenced in both the forward and reverse directions to guarantee the accuracy of nucleotide identification. Standard chromatographic curves of forward and reverse sequences were imported into the program ChromasPro 1.41 (Technelysium, Tewantin, Queensland, Australia) and manually aligned and edited. Consensus sequences were exported to the program BioEdit version 7.0.9 (Hall 1999), aligned with other sequences using the CLUSTALW algorithm (Thompson et al. 1994), and adjusted by the eye.

Figure 1 Map of Africa showing the seven water bodies from which Nile perch were sampled.
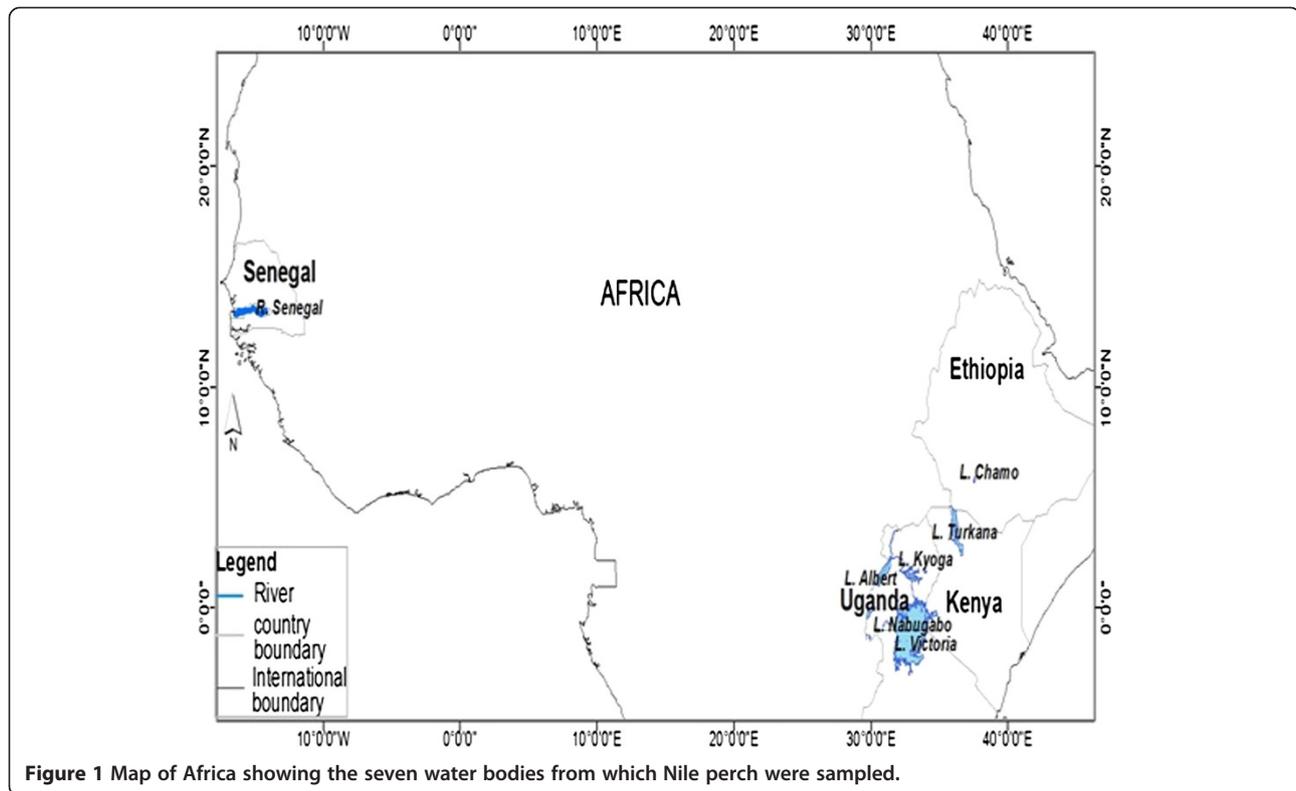
Nucleotide diversity, haplotype diversity, and population differentiation

The number of haplotypes, haplotype diversity (h), number of polymorphic sites based on the methods of Nei (1987), an estimate of nucleotide diversity (π), the average number of nucleotide differences (k; Tajima 1989), and the average number of nucleotide substitutions per site (dxy; Nei 1987) were calculated using DnaSP 5.0 (Librado and Rozas 2009). The MEGA 5.0 software (Tamura et al. 2011) was used to generate a table of variable sites. The number of haplotypes was counted for all polymorphic sites in all populations. Two measures of population differentiation, namely (a) F_{ST} calculated using ARLEQUIN 3.5 (Excoffier and Lischer 2010) and (b) N_{ST} calculated using DnaSP 5.0 (Librado and Rozas 2009), were considered. Population subdivisions were further tested by calculating the nearest neighbor statistic (S_{nn}; Hudson 2000) for mtDNA control region sequences (10^4 permutations and gaps excluded in pairwise comparisons), using DnaSP 5.0 (Librado and Rozas 2009). In addition, to visualize genetic heterogeneity, a minimum spanning haplotype network was estimated using the TCS program (Clement et al. 2000), which implements the statistical parsimony method of Templeton et al. (1992).

Genetic diversity and phylogenetic analyses

Levels of genetic diversity within and among the different geographical populations were compared using haplotype diversity and a maximum-likelihood (ML) estimation of the average number of nucleotide substitutions per site within and among groups (Nei 1987) using MEGA 5.0 (Tamura et al. 2011). A hierarchical analysis of molecular variance (AMOVA; Excoffier et al. 1992) was performed using ARLEQUIN 3.5 (Excoffier and Lischer 2010) to compare the imputable component genetic diversity to the variance among major lineages of the Nile perch with that observed among populations within each lineage. Φ_{ST} analyses were performed using a matrix of Tamura and Nei (1993) distances. The significance of the variance components associated with different levels of genetic structure was tested using 10^4 permutations.

Phylogenetic trees were constructed using neighbor-joining (NJ) and ML methods in MEGA5.0 (Tamura et al. 2011), and the Bayesian inference (BI) method using MRBAYES (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). The tree topology was tested by a bootstrap analysis with 5×10^6 pseudo-replicates to evaluate support for the phylogenetic relationships (Felsenstein 1985). These were further supported by construction of a minimum spanning haplotype network using TCS version 1.6 (Clement et al. 2000) as described in Templeton et al. (1992). MODELTEST 3.6 (Posada and Crandall 1998) was used to determine the optimal substitution model for the mtDNA data. Parameter values estimated by MODELTEST were adopted for further analysis, including phylogenetic relationships of the haplotypes, AMOVA, and estimates of gene flow. The NJ and ML trees used the Nile perch's

putative most closely related sister species, barramundi *Lates calcarifer* [GenBank:DQ012415.1] as an outgroup. The robustness of the NJ and ML tree branches were estimated based on 2,000 bootstraps, using the complete deletion of gaps.

Mismatch analysis and demographic history

The history of demographic changes in the Nile perch among the identified major lineages was investigated by a mismatch distribution analysis and neutrality tests. Since the population structure has a limited effect on the mismatch distribution (Harpending 1994), different populations within the same lineage were pooled together. Tajima's (1989) and Fu's (1997) tests of selective neutrality were used to examine haplotypes for the effect of selection by checking deviations from neutrality on the total number of segregated sites as a means to assess evidence of population expansion. The population demographic history was examined by calculating mismatch distributions (Harpending 1994; Rogers 1995). Both the neutrality tests and mismatch distribution analysis were conducted using ARLEQUIN 3.5 (Excoffier and Lischer 2010). The pairwise frequency distribution of individuals was used to determine species population expansion. Demographic

expansion parameters, tau (τ), theta at time 0 - (θ_0) and theta at time 1 - (θ_1), Harpending's raggedness index, the simulated sum of squared deviation, and the observed sum of squared deviations were calculated using ARLEQUIN 3.5. The raggedness index was included in the mismatch analysis to determine the goodness of fit to a unimodal distribution (Harpending 1994). Following the method of Schneider and Excoffier (1999), the moment estimators of time to expansion τ and the mutational parameters before ($\theta_0 = 2\mu N_0$) and after expansion ($\theta_1 = 2\mu N_1$) were determined and were expressed in units of mutational time, where N_0 and N_1 are respective female effective population sizes before and after an expansion that occurred τ generations ago. An estimate of time since population expansion, t, was calculated as $t = \tau/2u$; τ was determined as in Rogers and Harpending's (1992) equation, $u = 2\mu_0 k$, where k is the number of nucleotides and μ_0 is the mutation rate per site per nucleotide. In this study, since the molecular clock for the control region of fishes seems to vary among different taxa of fishes, we used a sequence divergence rate of 3.6% per 10^6 years (Kim and Raymond 1999) estimated using the geminate species of snook from a sister family, the Centropomidae, to that of the Nile perch, the Latidae, for the Nile perch mtDNA control region analysis.

Figure 2 List of 35 polymorphic sites from 37 haplotypes of the mitochondrial DNA D-loop of Nile perch. Dots indicate nucleotide variants identical to the first sequence. VIC, Lake Victoria; NAU, Nabugabo Lake; KYU, Kyoga Lake; AKU, Albert Lake; TLK, Turkana Lake; TAE Chamo Lake; SEN, Senegal River.

Table 1 Nile perch population statistics

Group	Population	*n*	*N*	*h*	*π*	*D*	*P*	FS	P_{FS}	*τ*	$θ_0$	$θ_1$	N_0	N_1	Hri
East Africa	Albert Lake	9	4	0.583 ± 0.183	0.00963 ± 0.0017	−1.037	0.16	2.217	0.88	20.43	0.250	25,513.639	70.90	8,386,680.09	0.339
	Kyoga Lake	10	3	0.644 ± 0.101	0.00207 ± 0.0007	−0.431	0.32	0.345	0.48	1.048	0.006	89,037.437	1.17	25,557,965.10	0.127
	Nabugabo Lake	6	5	0.933 ± 0.122	0.01570 ± 0.0074	−0.940	0.23	0.240	0.46	2.251	0.080	81,810.182	25.97	24,164,312.56	0.164
	Chamo Lake	4	3	0.833 ± 0.222	0.00259 ± 0.0008	0.592	0.83	−0.658	0.15	2.408	0.255	90,394.584	44.64	26,747,561.57	0.250
	Turkana Lake	8	4	0.786 ± 0.113	0.00302 ± 0.0005	0.717	0.80	−0.621	0.20	1.622	0.064	82,186.554	14.59	23,426,793.30	0.087
	Victoria Lake	72	24	0.865 ± 0.033	0.00934 ± 0.0017	−1.222	0.08	−8.328	0.01	2.279	0.203	50,336.038	69.74	15,089,660.07	0.021
West Africa	Senegal River	15	9	0.845 ± 0.008	0.00461 ± 0.0012	−1.489	0.07	−4.319	0.01	1.606	0.034	79,119.289	11.09	23,117,407.51	0.104

Population age ranges are given with 95% confidence intervals of *τ* for 1,000 mismatch distribution replicates. *n*, sample size; *N*, number of haplotypes; *h*, haplotype diversity; *π*, nucleotide diversity; *D*, Tajima's test; *P*, *p* value of Tajima's test; FS, Fu's FS test; P_{FS}, *p* value of the FS test; *τ*, tau; $θ_0$, theta at time 0; $θ_1$, theta at the time of the recorded measure; N_0, population at time 0; N_1, population at the time of the recorded measure; Hri, Harpending raggedness index.

Results

Nucleotide diversity, haplotype diversity, and population differentiation

Sequencing of the mtDNA control region of 124 individuals of the Nile perch produced a PCR fragment from which 476 nucleotides could be analyzed; among which, 260 sites were variable. Thirty-seven unique haplotypes were found. Observed nucleotide frequencies were A = 0.283, T = 0.336, C = 0.178, and G = 0.203. Overall, haplotype diversity (*h*) was 0.893 ± 0.018, and nucleotide diversity (*π*) was 0.12408. The most common haplotype, i.e., haplotype 5, was shared by 33 individuals and was found in most sampled water bodies except for Albert Lake, Chamo Lake, and the Senegal River (Figure 2). In addition, 78.4% of the haplotypes were restricted to their sampling locality (sampled water bodies) (Figure 2). Haplotype sequences were deposited in GenBank with accession numbers JQ778220 ~ 56.

Haplotype diversities (*h*) for the different collections were relatively high, ranging from 0.583 ± 0.183 for the Albert Lake population to 0.933 ± 0.122 for the Nabugabo Lake population, whereas nucleotide diversities, *π*, ranged

0.00207 ~ 0.01570 (Table 1). The level of genetic diversity varied among the different populations but was comparable among the two major lineages: the East and West African lineages (Table 1).

Population differentiation assessed using the AMOVA showed highly significant heterogeneity ($Φ_{ST}$ = 0.99, N_{ST} = 0.99) when the populations were grouped in two geographic regions: East and West Africa (Tables 2 and 3). The grouping was undertaken following the bifurcation pattern created by the high elevation of sub-basins of the Nile River that determined its northerly flow course. Even when the populations were grouped according to the sample sources, the AMOVA revealed highly significant heterogeneity ($Φ_{ST}$ = 0.94) among the sampled water bodies (Table 2), most probably caused by a lack of shared haplotypes between East and West African Nile perch populations, which may have led to an overestimation of genetic differentiation. To further test the samples' homogeneity, the nearest neighbor statistic (S_{nn}) was calculated for D-loop sequences. When samples were grouped into East and West African populations, the overall test statistic

Table 2 AMOVA of Nile perch from seven water bodies

Source of variation	AMOVA for different water bodies					AMOVA for the two lineages				
	d.f.	Sum of squares	Variance components	Variation (%)	*Φ*-statistic	*d.f.*	Sum of squares	Variance components	Variation (%)	*Φ*-statistic
Among groups	6	6,026.765	76.46248	93.45	0.9345	1	6,465.984	245.03121	99.21	0.9921
Among populations within groups	14	31.644	−0.83513	−1.02	−0.1557	5	16.596	0.12678	0.05	0.0648
Within populations	103	638.323	6.19732	7.57	0.9243	117	214.153	1.83037	0.74	0.9926
Total	123	6,696.733	81.82467			123	6,696.733	123.40335		

Lake Victoria, Kyoga Lake, Nabugabo Lake, Albert Lake, Chamo Lake, Turkana Lake, and the Senegal River populations were used as the top level grouping.

Table 3 mtDNA sequence characteristics of the two lineages

Statistic	East Africa lineage	West Africa lineage	Overall
m	109	15	124
S	35	11	260
N	28	9	37
h	0.864	0.848	0.893
Kt	3.91	2.08	55.83
π	0.0087	0.0046	0.1241
S_{nn}	-	-	1.00
K	-	-	0.9806
Nm	-	-	0.00
Φ_{ST}	-	-	0.99
N_{ST}	-	-	0.99
D	−1.259 ($p = 0.08$)	−1.489 ($p = 0.06$)	−0.545 ($p = 0.36$)
FS	−10.685 ($p = 0.00$)	−4.319 ($p = 0.01$)	−1.589 ($p = 0.32$)

m, number of sequences; S, number of segregated sites; N, number of haplotypes; h, haplotype diversity; K, average proportion of nucleotide differences between lineages; Kt, average number of nucleotide differences; π, nucleotide diversity; S_{nn}, Hudson's statistic (Hudson 2000); Nm, gene flow; D, Tajima's neutrality test; FS, Fu's neutrality test; p, probability values of the tests.

revealed a significant association between D-loop sequence similarity and the geographical location ($S_{nn} = 1.000$, $p = 0.001$) (Table 3).

The minimum spanning haplotype network obtained using TCS 1.21 revealed that both the East and West African Nile perch collections formed a star-shaped network as the general pattern, rooted in the central-most common haplotype which is possibly the ancestral haplotype (Figure 3). This suggests the existence of relationships between populations from either region and recent population expansion in the Nile perch of both regions.

Geographical distribution of Nile perch lineages, hierarchical genetic diversity, and phylogenetic analyses

The optimal substitution model of GTR (Tavere 1986) for Nile perch was obtained with MODELTEST 3.6 (Posada and Crandall 1998). The parameters for this model were as follows: base frequencies were A = 0.290, T = 0.329, C = 0.179, and G = 0.202, and α = 1.46. The NJ tree presented a brush-shaped topology (Figure 4). Nile perch haplotypes clustered into two distinct groups (henceforth referred to as the East and West African lineages). The distribution of the major Nile perch lineages was exclusive to where samples were collected (East or West Africa) with no overlap in distribution. Haplotypes of the Nile perch collected from lakes in East Africa were not found in West African collections and vice versa (Tables 1 and 4). Most of the molecular variance observed in Nile perch was imputable to differences among lineages, $\Phi_{ST} = 0.9345$ (Table 2). The phylogenetic trees showing the 37 haplotypes observed in African Nile perch are shown in Figures 4 and 5. The two clusters were separated by an observed 98.06%

Figure 3 Minimum spanning network analysis of haplotypes of the African Nile perch. The sizes of the circles and rectangles are representative of the haplotype frequency, and the rectangles represent the ancestral haplotypes in both clades. Haplotypes are marked by names that correspond to those in Figure 2. The dots on the lines represent putative mutational steps between haplotypes obtained using the TCS program (Clement et al. 2000).

Figure 4 Neighbor-joining tree describing the phylogenetic relationships among 37 mtDNA CR haplotypes of Nile perch. Branch lengths are expressed as observed percentage nucleotide substitutions. Bootstrapped values based on 5,000 replications are reported at the internodes. The tree was rooted using the mtDNA CR sequence of *L. calcarifer* (barramundi).

nucleotide divergence (Table 3) with 99% bootstrap value support for the NJ tree (Figure 4) and a bootstrap value of 1.00 for the 50% majority rule consensus tree based on the BI analysis (Figure 5). Haplotypes of East and West Africa defined separate monophyletic clades (Figures 4 and 5). The genetic distance by the simple P method, $P = Zd/Zt$ (Avise 1994), where Zd is the total number of mutations, deletions, and substitutions and Zt is the total number of bases compared, produced a genetic distance of $P = 58.2\%$ between the two lineages.

The relationship among haplotypes and the frequency of each haplotype of the African Nile perch were further visualized by constructing a minimum spanning network (Figure 3). In the network, two lineages of East and West Africa were recognized. There was substantial structuring of groups among lineages with an overall Φ_{ST} of 0.0648 (Table 2).

Mismatch analysis and demographic history

The mismatch distribution was clearly bimodal with one mode corresponding to the number of differences between lineages (approximately 240) and the other to differences among individuals within lineages (approximately 2) (Figure 6). The goodness of fit test showed that no mismatch distributions within the two lineages significantly deviated ($X^2 = 0.224$, $p > 0.05$) from predicted values under the sudden expansion model (Table 2). Considering the 95% confidence interval generated by simulations, values of the observed numbers of polymorphic sites were consistent with simulated sites, meaning that the parameters estimated by the model were sufficiently accurate to account for the observed polymorphism within the two Nile perch lineages. Observed values for the age expansion parameter, τ (Table 4), did not substantially differ among populations of the different sampled water bodies, and respective values for the two lineages were 1.691 and 1.563 for the East and West African lineages. The respective estimated times since expansion before the present for the two Nile perch lineages were 49,603 and 45,576 years. The expansion model was further supported by the fact that its findings for all collections agreed with the expectations of a historically expanding population model (Rogers 1995), where P [simulated sum of squared deviations (SSD$_{sim}$)] $\geq P$ [observed SSD (SSD$_{obs}$)] (Table 4), the low Harpending's raggedness indices (Hri = 0.156, $P = 0.570$), and negative values of Tajima's D and
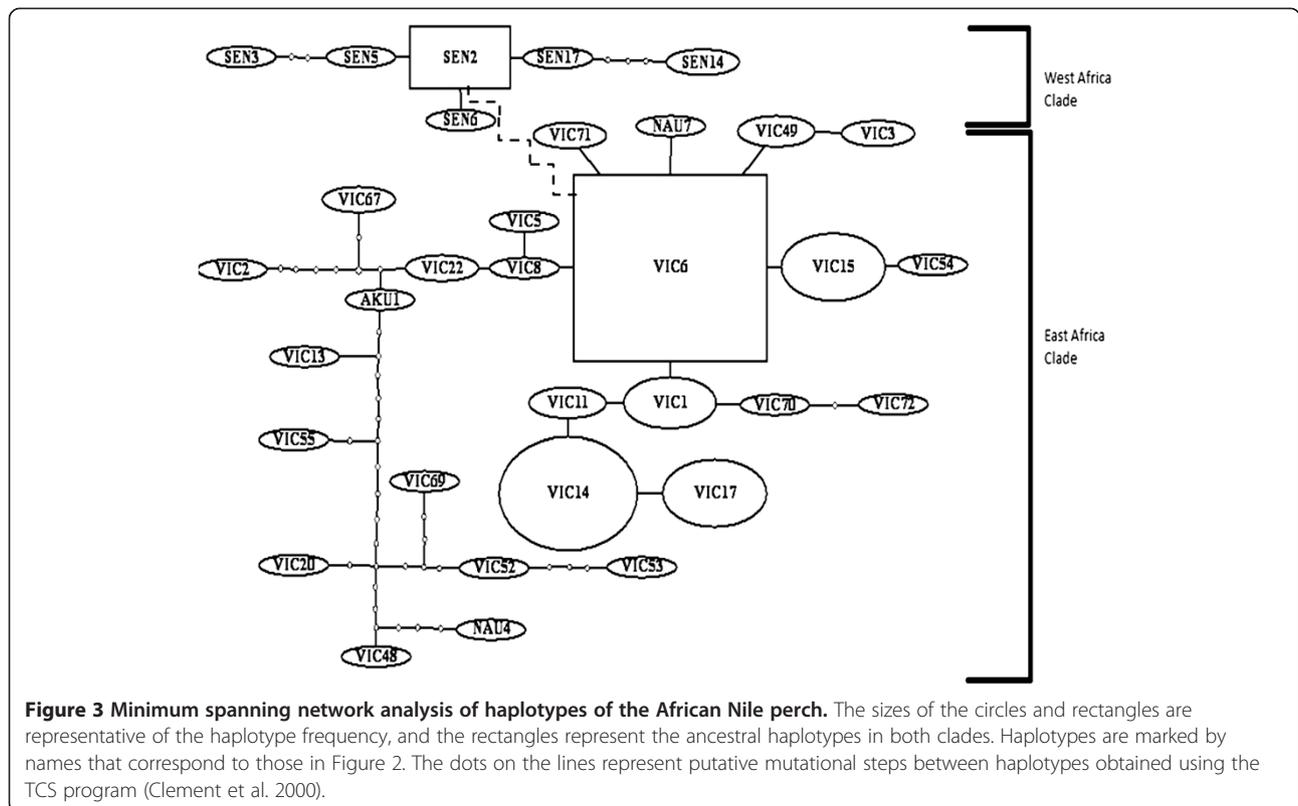
Table 4 Observed 95% CIs of simulated *S* and demographic parameters among Nile perch lineages

Group	Population	S	95% CI	τ	θ_0	θ_1	P (SSD$_{sim}$)	P (SSD$_{obs}$)	Hri	P (Hri)
East Africa	Albert Lake	15	1 ~ 14	13.81 (0.04 ~ 68.81)	0.00 (0.00 ~ 3.57)	2.11 (0.16 ~ 99,999.00)	0.240	0.126	0.339	0.20
	Kyoga Lake	3	1 ~ 10	0.97 (0.00 ~ 2.45)	0.00 (0.00 ~ 0.05)	99,999.00 (1.35 ~ 99,999.00)	0.580	0.012	0.127	0.60
	Nabugabo Lake	19	2 ~ 12	2.20 (0.04 ~ 4.18)	0.00 (0.00 ~ 1.40)	99,999.00 (2.41 ~ 99,999.00)	0.130	0.085	0.164	0.48
	Chamo Lake	2	0 ~ 7	1.51 (0.00 ~ 4.30)	0.00 (0.00 ~ 2.53)	99,999.00 (16.84 ~ 99,999.00)	0.650	0.037	0.250	0.74
	Turkana Lake	3	3 ~ 12	1.60 (0.00 ~ 3.23)	0.00 (0.00 ~ 0.97)	99,999.00 (4.79 ~ 99,999.00)	0.680	0.011	0.087	0.76
	Victoria Lake	33	19 ~ 41	1.25 (0.24 ~ 5.65)	1.25 (0.00 ~ 2.16)	26.25 (4.05 ~ 99,999.00)	0.300	0.008	0.021	0.85
West Africa	Senegal River	11	6 ~ 18	1.56 (0.48 ~ 2.75)	0.00 (0.00 ~ 0.53)	99,999.00 (2.92 ~ 99,999.00)	0.170	0.028	0.104	0.34

CIs, confidence intervals; *S*, number of polymorphic sites. τ, θ_0, and θ_1 are the age of expansion and population size before and following expansion, expressed in units of mutational time, respectively. SSD is sum of square deviations between the observed and the expected mismatch distribution for a demographic history of the lineage defined by estimates of the τ, θ_0, and θ_1 parameters. P (Hri) is the probability of observing by chance a higher value of the raggedness index than the observed one, under a hypothesis of population expansion. Population age ranges are given with 95% CIs of τ for 1,000 mismatch distribution replicates.

Fu's *F* (*D* = −0.550, FS = −1.589) were consistent with population expansion.

Discussion

Results of the phylogeographic, phylogenetic, and mismatch analyses revealed that Nile perch on the African continent could be divided into two lineages that most likely identify ancestral populations of Nile perch that evolved separately as a result of allopatric separation. This discussion focuses on the species' evolutionary history as may have been dictated by various paleo-environmental changes on the African continent in the species' extant range that is most compatible with differential patterns of genetic diversity observed within and among the major evolutionary lineages beginning with the Pliocene-Pleistocene epoch to modern times.

Nucleotide diversity, haplotype diversity, and population differentiation

The overall haplotype diversity was relatively high compared to those of other freshwater fishes (Mccusker and Bentzen 2010). This high haplotype diversity can be explained by the presence of divergent lineages, the East and West African Nile perch lineages. In addition, the genetic signature of high haplotype diversity and low nucleotide diversity found in the African Nile perch can be attributed to rapid population expansion after a reduction in the effective population size (Ludt et al. 2012). The rapid population size increases immediately after the bottlenecks and a possibility that the populations may have had high pre-bottleneck genetic diversities, factors known to augment genetic variation (Avise 1994), may have countered any negative bottleneck effects. The high genetic heterogeneity as exhibited by the high nucleotide

divergences and the lack of shared haplotypes between the two major Nile perch lineages can be explained by the long separation of the two lineages (McCracken and Sorenson 2005) caused by the formation of the novel Nile River basin through rifting, tilting, and climate changes on the African continent in the Late Pleistocene (Talbot et al. 2000).

Phylogenetic relationships among Nile perch populations

Nucleotide variability of control region sequences showed geographical structuring in Nile perch. Two major haplotype groups were found on the African continent, and these were well supported. The two groups showed high bootstrap values, clearly resolving the Nile perch into East and West African groupings that were monophyletic in the MP, NJ, and ML analyses. The high sequence divergence between the two groups and high number of sequences analyzed explain the high bootstrap value on the node, separating the East and West African haplotypes. The genetic distance between the two lineages was high, and there were no haplotypes shared between the two lineages, suggesting historical interruption of gene flow for a number of generations, a period long enough for the species to reach a stage of reciprocal monophyly (Avise et al. 1987; Kazirian and Donnelly 2004).

The parsimonious network of mtDNA haplotypes of the Nile perch D-loop for the two lineages, where the most common haplotypes were rooted in the center and shared with or connected to all of the other haplotypes, is an indicator of a recent population expansion in both lineages. Theoretical and empirical expectations suggest that mtDNA which shares a single recent ancestor shows a strong tendency for the ancestral haplotype (often most frequent) to be most interior followed in declining
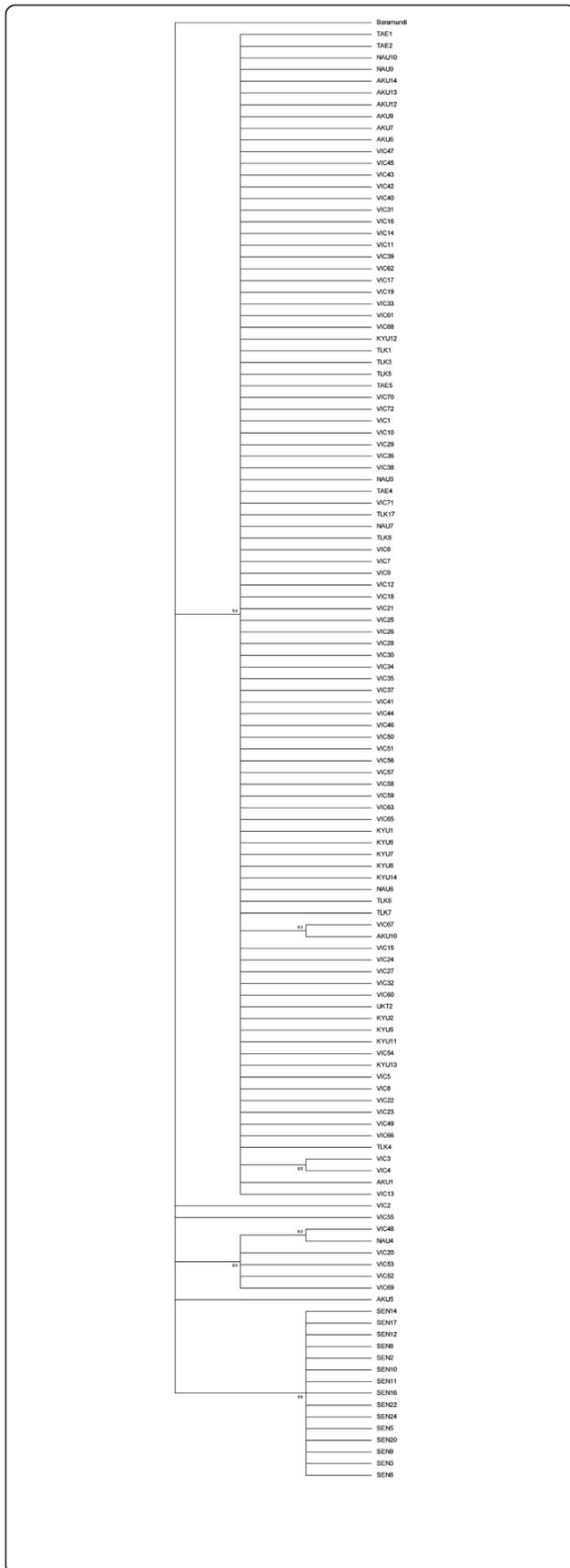
Figure 5 Maximum likelihood tree describing the phylogenetic relationships among 37 mtDNA CR haplotypes of Nile perch. Branch lengths are expressed as observed percentage nucleotide substitutions. Bootstrapped values (>50) based on 100 replications are reported at the internodes. The tree was rooted using the mtDNA CR sequence of *L. calcarifer* (barramundi).

frequency by more distantly related haplotypes (Avise et al. 1988).

Mismatch analysis and demographic history

In the current study, both Nile perch lineages, the East and West African lineages, fit the sudden expansion model, and when comparing the number of generations as revealed by τ values, the expansions may have occurred around the same time. The population expansion model was further supported by the fact that both lineages had relatively high haplotype diversity (0.864 and 0.848, respectively) in contrast to having low nucleotide diversity (0.0087 and 0.0046, respectively), a scenario which suggests that the two lineages experienced population bottlenecks followed by population expansions (Avise 2000). The mismatch distribution topologies were very similar, and expansion times were close, suggesting that the expansion events may have occurred in the same period, most probably during the late Pleistocene, when the African continent went through alternating dry and wet climatic phases that may have led to bottlenecks and expansions in fish populations and speciation (Agnèse et al. 1997; Beuning et al. 1997). The negative values of Fu's index and Tajima's *D* further support the hypothesis of recent population expansion. The formation of the novel Nile basin from previous sub-basins through rifting, tilting, and climatic changes in the late Pleistocene may have cut off all outflows from the Nile to West African lakes and river basins, thereby stemming any possible genetic exchange between East and West African fish. This bifurcation was associated with the rifting and tilting of the African continent that led to the rise of the left escarpment of the river Nile basin separating the Nilo-Sudan ichthyofaunal province into West and East Africa (Adamson 1982; Adamson et al. 1993; Woodward et al. 2007). Our findings are similar to those of previous studies that worked on Nile tilapia phylogenetics, a species with a similar distribution to that of the Nile perch on the African continent (Agnèse et al. 1997; Rognon and Guyomard 1997). Those studies also indicated that populations of the Nile tilapia can be divided into East and West African lineages.

Taxonomic status of the Nile perch

The amount of genetic variation as measured by F_{ST}, genetic distance, and phylogenetic nodal support (bootstrap scores) found among East African Nile perch populations was insufficient to support the earlier classification of Nile
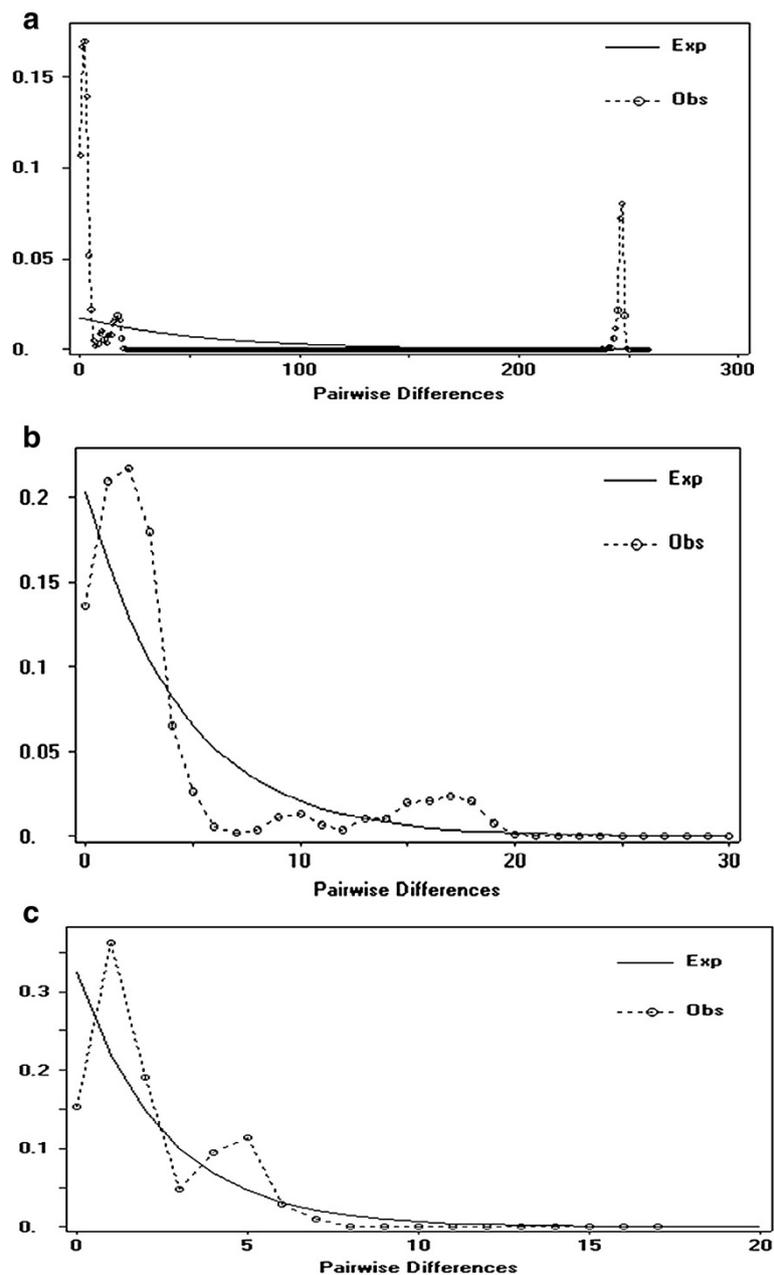
Figure 6 Mismatch distribution. Frequency distributions of pairwise numbers of mutational differences between Nile perch observed in both lineages combined **(a)** and separately in the two Nile perch lineages defined by a mitochondrial DNA analysis: East Africa **(b)** and West Africa **(c)**. Dots represent observed data, and the bold curve is the model fitted to the data.

perch by ecological biologists into four different species, *Lates niloticus*, *Lates macropthalamus*, *Lates longispinis*, and *Lates albertianus*, in the East African region (Harrison 1991). Our findings do suggest that there are two genetic groupings of Nile perch on the African continent, including one East African group with two genetically divergent lineages and one West African group. Speciation into separate East and West African groups may have been caused by long geographical separation with no genetic

exchange. Apparently, phylogenetic relationships among East African Nile perch populations are still problematic and require further investigation.

Conclusions

MtDNA control region data indicate that there are two genetically divergent lineages of Nile perch on the African continent, an eastern lineage defined by populations from East African and Ethiopian waters and a western lineage

consisting of populations from the Senegal River. Genetic differences between these lineages are likely indicative of two different species and two Evolutionarily Significant Units (ESUs).

Since depletion of one species cannot easily be replenished by natural migration by the other, the ESUs should be conserved as two separate management units, and genetic exchange due to anthropologic activities between the two should be avoided. The status of fisheries of the Nile perch on the African continent, although listed as of 'least concern' because of its wide distribution and there being no major threats, is vulnerable to overexploitation (Azeroual et al. 2010). The Nile perch fishery on the African continent has relatively high genetic diversity compared to other freshwater fish species, meaning the fishery is healthy and if effectively managed can ably support both the species' sustainable exploitation and conservation efforts.

Author details
[1]ARDC-Kajjansi, PO Box 530, Kampala, Uganda. [2]Department of Environmental Management, Makerere University, PO Box 7098, Kampala, Uganda. [3]Biological Sciences Department, Makerere University, PO Box 7062, Kampala, Uganda. [4]Department of Fisheries Resources, PO Box 4, Entebbe, Uganda.

References
Adamson DA (1982) The integrated Nile. In: Williams MAJ, Adamson DA (ed) A land between two Niles: quaternary geology and biology in Central Sudan. AA Balkema, Rotterdam, pp 221–234
Adamson D, McEvedy R, Williams MAJ (1993) Tectonic inheritance in the Nile basin and adjacent areas. Isr J Earth Sci 41:75–85
Agnèse FJ, Adépo-Gourène B, Abban KE, Fermon Y (1997) Genetic differentiation among natural populations of the Nile tilapia Oreochromis niloticus (Teleostei, Cichlidae). Heredity 79:88–96
Avise JC (1994) Molecular markers, natural history and evolution. Chapman and Hall, New York, p 511
Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge, p 447

Avise JC, Arnold J, Ball RM, Jr, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematic. Annu Rev Ecol Systemat 18:489–522
Avise JC, Ball MR, Arnold J (1988) Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory of neutral mutations. Mol Biol Evol 5(4):331–344
Azeroual A, Entsua-Mensa M, Getahun A, Lalèyè P, Moelants T, Ntakimazi G (2010) Lates niloticus. In: IUCN 2013. IUCN red list of threatened species. Version 2013.1. www.iucnredlist.org. Accessed 24 July 2013
Balon EK (1974) Fish production of a tropical ecosystem. In: Balon EK, Coche AG (ed) Lake Kariba: a man-made tropical ecosystem in Central Africa, mono-graphiae biologicae 24. Dr W Junk Publishers, The Hague, pp 248–748
Bernatchez L, Danzmann GR (1993) Congruence in control-region sequence and restriction-site variation in mitochondrial DNA of Brook charr (Salvelinus fontinalis Mitchill). Mol Biol Evol 10:1002–1014
Beuning KRM, Talbot MR, Kelts K (1997) A revised 30000 year paleoclimatic and paleohydrologic history of Lake Albert, East Africa. Paleogeogr Paleoclimatol Palaeoecol 136(3–4):259–279
Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9:1657–1660
Cotterill FPD, de Wit MJ (2011) Geoecodynamics and the Kalahari epeirogeny: linking its genomic record, tree of life and palimpsest into a unified narrative of landscape evolution. S Afr J Geol 114:493–518
Craw D, Burridge C, Anderson L, Waters JM (2007) Late Quaternary river drainage and fish evolution, Southland, New Zealand. Geomorphology 84:98–110
Dominguez-Dominguez O, Alda F, Perez-Ponce G, de Leon JL, Garcia-Garitagoitia ID (2008) Evolutionary history of the endangered fish Zoogoneticus quitzeoensis (Bean, 1898) (Cyprinodontiformes: Goodeidae) using a sequential approach to phylogeography based on mitochondrial and nuclear DNA data. BMC Evol Biol 8:161
Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resourc 10:564–567
Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491
Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791
FishBase (2010) Lates niloticus, Nile perch. FishBase. http://fishbase.org.cn/Country/CountryList.php?ID=347&GenusName=Lates&SpeciesName=niloticus. Accessed 6 Aug. 2011
Fu XY (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925
Golubtsov SA, Habteselassie R (2010) Fish faunas of the Chamo-Abaya and Chew Bahir basins in southern portion of the Ethiopian Rift Valley: origin and prospects for survival. Aquat Ecosyst Health Manag 13(1):47–55
Goodier SA, Cotterill FP, O'Ryan C, Skelton PH, de Wit MJ (2011) Cryptic diversity of African tigerfish (genus Hydrocynus) reveals palaeogeographic signatures of linked Neogene geotectonic events. PLoS One 6(12):e28775. doi:10.1371/journal.pone.0028775
Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acid Symp Ser 41:95–98
Harpending RC (1994) Signature of ancient population growth in a low resolution mitochondrial DNA mismatch distribution. Hum Biol 66:591–600
Harrison K (1991) The taxonomy of East African Nile perch, Lates spp. (Perciformes, Centropomidae). J Fish Biol 38:175–186
Hudson RR (2000) A new statistic for detecting genetic differentiation. Genetics 155:2011–2014
Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754–755
Kazirian D, Donnelly AM (2004) The criterion of reciprocal monophyly and classification of nested diversity at the species level. Mol Phylogen Evol 32:1072–1076
Kim AD, Raymond RW (1999) Amphi-panaimic geminates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in mtDNA control region of fishes. Mol Phylogen Evol 13(1):208–213
Lévêque C (1997) Biodiversity dynamics and conservation: the freshwater fish of tropical Africa. Cambridge University Press, Cambridge

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

Ludt BW, Bernal AM, Bowen WB, Rocha AL (2012) Living in the past: phylogeography and population histories of Indo-Pacific wrasses (genus *Halichoeres*) in shallow lagoons versus outer reef slopes. PLoS One 7(6):e38042. doi:10.1371/journal.pone.0038042

McCracken K, Sorenson M (2005) Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear gene trees in the stiff-tailed ducks (Nomonyx-Oxyura)? Syst Biol 54(1):35–55

McCusker RM, Bentzen P (2010) Positive relationship between genetic diversity and abundance in fishes. Mol Ecol 19:4852–4862

Moritz T, Linsenmair K (2005) West African fish diversity–distribution patterns and possible conclusions for conservation strategies. In: Huber AB, Sinclair JB, Karl-Heinz L (ed) African biodiversity. Springer, New York, pp 187–196

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nielsen LJ (1998) Population genetics and the conservation and management of Atlantic salmon (*Salmo salar*). Can J Fish Aquat Sci 55(S1):145–152

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

Ogutu-Ohwayo R (1993) The effects of predation by Nile perch, *Lates niloticus*, L. on fish of Lake Nabugabo, with suggestions for conservation of endangered cichlids. Conserv Biol 7(3):701–711

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818

Ramakrishnan U, Hadly EA (2009) Using phylochronology to reveal cryptic population histories: review and synthesis of four ancient DNA studies. Mol Ecol 18:1310–1330

Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. Evolution 49:608–615

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic difference. Mol Biol Evol 9:552–569

Rognon X, Guyomard R (1997) Mitochondrial DNA differentiation among East and West African Nile tilapia populations. J Fish Biol 51:204–207

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574

Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics 152:1079–1089

Soltis PS, Gitzendanner AM (1999) Molecular systematic and conservation of rare species. Conserv Biol 13:471–483

Tajima F (1989) Statistical method for testing neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595

Talbot RM, Williams JAM, Adamson AD (2000) Strontium isotope evidence for late Pleistocene reestablishment of an integrated Nile drainage network. Geology 28(4):343–346

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Tavere S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures Math Life Sci (Am Math Soc) 17:57–86

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotype associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132:619–633

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22:4673–4680

Tzeng CS, Lin YS, Lin SM, Wang TY, Wang FY (2006) The phylogeography and the population demographics of selected freshwater fishes in Taiwan. Zool Stud 45:285–297

van Tuinen M, Ramakrishnan U, Hadly EA (2004) Studying the effect of environmental change on biotic evolution: past genetic contributions, current work and future directions. Phil Trans A Math Phys Engin Sci 362:2795–2820. doi:10.1098/rsta.2004.1465

Woodward JC, Macklin GM, Krom DM, Williams JAM (2007) The Nile: evolution, quaternary river environments and material fluxes. In: Maddy D, Macklin MG, Woodward JC (ed) River basin sediment systems: archives of environmental change. AA Balkema, Rotterdam, pp 261–292

Zengeya TA, Decru E, Vreven E (2011) Revalidation of *Hepsetus cuvieri* (Castelnau, 1861) (Characiformes: Hepsetidae) from the Quanza, Zambezi and southern part of the Congo ichthyofaunal provinces. J Nat Hist 45(27):1723–1744