Genetic Diversity and Structure of *Opsaridium microlepis* along Lake Nyasa, Tanzania

***Nehemia, A.[1] and Mwakalesi, A.J.[2]**

[1]Department of Biosciences, Sokoine University of Agriculture,
P.O. Box 3038, CHUO KIKUU, Morogoro-Tanzania
[2]Department of Chemistry and Physics, Sokoine University of Agriculture,
P.O. Box 3038, CHUO KIKUU, Morogoro-Tanzania

*****Corresponding author e-mail: *nehemiah@sua.ac.tz*

Abstract

Opsaridium microlepis is a fish species that serves as a source of revenue and protein to most people living near Lake Nyasa. However, the population of this species has witnessed a worrisome decline, leading the International Union for Conservation of Nature (IUCN) to classify it as a threatened fish species. The current work used partial mitochondrial cytochrome oxidase subunit I (COI) sequences to investigate the genetic diversity, effective population size, and structure of O. microlepis along the Lake Nyasa areas of Tanzania. The findings indicated that Kafyofyo had the highest nucleotide diversity (π = 0.20%) and haplotype diversity (h = 0.78), while Katumba had the lowest nucleotide diversity (π = 0.08%). The lowest haplotype diversity (h = 0.50) was recorded at Katumba and Mbambabay. The molecular analysis indicated significant differences across the subpopulations investigated (Overall PhiST (ΦST) = 0.093, P < 0.05). The mean Sum of Square deviation (SSD), Harpending's raggedness index (HRI), Tajima's D (D), and Fu's Fs (Fs) were 0.014, 0.157, 0.108, and -0.88680 respectively and neither the mismatch distribution nor the neutrality test findings were significantly different from zero. The mismatch distribution supports the idea of sudden population expansion. Consequently, the effective population size estimates are large for sampling sites with higher genetic diversity. Thus, the current study's findings can serve as a foundation for long-term strategic plans to conserve and manage populations of O. microlepis in areas with low genetic diversity and effective population size.

Keywords: Lake Nyasa, Tanzania, genetic diversity, effective population size, *Opsaridium microlepis*

Introduction

Lake Nyasa has been a substantial source of support for the local communities residing along its shores, primarily through its abundant fishery resources. However, the economically valuable Lake Salomon fish, *Opsaridium microlepis* has been dramatically declining. As a result, the International Union for Conservation of Nature (IUCN) has red-listed it as a threatened species. The species' abundance was predicted to decrease by 30% in a decade because of increased overfishing, siltation, and habitat degradation in rivers and nearshore areas (Tweddle 2018). Thus, understanding the status of genetic diversity and effective population size for proper managing the *O. microlepis* species and other aquatic species in the area is essential. Genetic diversity plays an important role in the adaptation ability of a population in the face of fluctuating environmental conditions (Markert *et al.* 2010). Declining population size may reduce genetic diversity (Petit-Marty *et al.* 2022). It has been shown that habitat type and life cycle variation influence fish genetic diversity (Martinez *et al.* 2018). The previous work on, "Migratory behavior shapes spatial genetic structure of cyprinid fishes within the Lake Malawi catchment," evaluated the genetic diversity and structure of *O. microlepis*, *O. tweddleorum,* and *O. microcephalum* along the Lake Nyasa catchment (Sungani *et al.* 2016). However, the sampling approach dominated a

portion of Malawi. Only one sampling location (Songwe) in Tanzania was used for the *O. microlepis* samples, making it difficult to justify the species' current genetic status, particularly in Tanzania. The current genetic diversity, structure, and effective population size of *O. microlepis* in Tanzania's Lake Nyasa area was investigated in this study. Additionally, the present study assesses the demographic history of *O. microlepis* in Lake Nyasa, Tanzania.

Materials and methods

Sampling

Lake Nyasa, with a surface size of 31,000 km^2, is Africa's third largest Lake after the Lake of Victoria and Tanganyika and the world's third deepest freshwater lake. The Lake is the southernmost of the Great African Rift Valley lakes, lying between Malawi, Mozambique, and Tanzania. The Rivers that drain their water into Lake Nyasa include River Songwe, River Kiwira, River Mbaka, River Lufiryo and River Ruhuhu. Lake Nyasa is situated between the latitudes of 9°30'-14°40'S and the longitudes of 33°50'-33°36'E (Nindi 2007). *Opsaridium microlepis* samples were obtained at six locations: 1. Kafyofyo, 2. Buloma, 3. Katumba, 4. Lupingu, 5. Ruhuhu, and 6. Mbambabay (Fig. 1). Twenty samples of *O. microlepis* were taken from each site for molecular analysis. *Opsaridium microlepis* fin tissue samples were collected and immediately stored in 99.9% Ethanol. While sampling, other common fish found in the area included *O. tweddleorum, O. microcephalum*, and *O. karongae*.

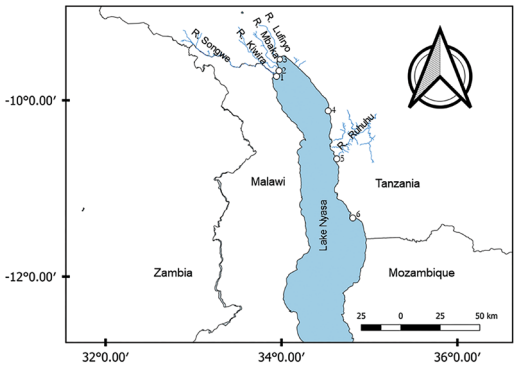

Figure 1: Sampling sites of the *Opsaridium microlepis* along the Lake Nyasa,Tanzania: white circle with black line colour represents the sampling sites

DNA extraction

About 30 mg of preserved tissue was used for DNA extraction following the protocol of Quick-DNA™ Miniprep Plus Kit (ZYMO Research). The DNA extraction success was checked using 2% TBE agarose gels.

Polymerase chain reactions and DNA sequencing

A segment of the COI gene containing 629 base pairs was amplified in a T100TM Thermal cycler device (Bio-Lab Inc, GA, USA) using the forward primer FishF1 and the reverse primer FishR2 (Ward *et al.,* 2005). The polymerase chain reactions were performed in a total volume of 35 µL, which included 2 µL of DNA template, Multiplex PCR 5X Master Mix, 0.4 mg of BSA, 0.3 µM of each primer, and 11.7 µL of RNAse free water. The temperature profile included an initial temperature of 94°C for 5 minutes, 35 cycles of 94°C for 40 seconds, 54°C for 45 seconds, 72°C for 1 minute, and a final extension of 72°C for 15 minutes. The primers FishF1 and FishR2 were used to Sanger sequence the PCR products at the Macrogen Europe Laboratory using an automated sequencer (AB 3730XL; Applied Biosystems, Foster City, USA).

Mitochondrial DNA analysis

Some PCR product sequences were found to be of poor quality and were eliminated from the dataset. The remaining sequences were first edited by removing the ends with MEGA 11 software, and the species identity was confirmed with the web tool BLAST. Using MEGA 11 software, the sequences were translated into amino acids CLUSTAL W (Tamura *et al.* 2013) was used for multiple alignments of the sequences as implemented in the software MEGA 11. The Haplotype network of partial mitochondrial cytochrome oxidase subunit I (COI) sequences from *Opsaridium microlepis* was constructed using the software Pop Art v.1.7 (Leigh and Bryant 2015) following the collapsing of the sequences using the online FaBox 1.41 (Villesen 2007). The genetic diversity was assessed using Arlequin v. 3.5.2.2

(Excoffier and Lischer 2010). Analysis of molecular variance was used to determine the within and among population differentiation. Arlequin v. 3.5.2.2 (Excoffier and Lischer 2010) examined the historical demography and neutrality parameters. Migrates-n v. 3.11 was used to calculate mutation-scaled effective population size. The first runs were done in three replicates, each with one long chain of 5,000 recorded steps and 50 increments. With an exponential prio distribution, the sampling parameter value and burn-in were 7,50,000 and 10,000, respectively. The final runs had 50,000 recorded steps, 50 increments, four replicates, 2,500,000 sampled parameter values, and a burn-in values 100,000. The heating approach used multiple Markov chains and four static temperatures with exponential prio distribution that began at 1, 1.5, 3, and 1,000,000.

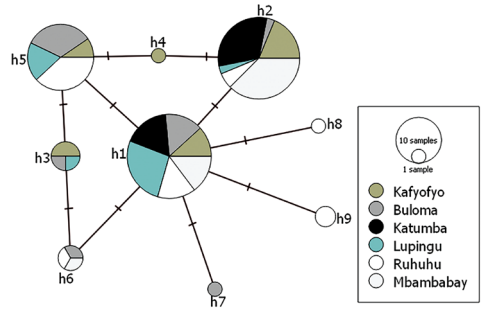

Figure 2: *Opsaridium microlepis's* haplotype network of partial mitochondrial cytochrome oxidase subunit I (COI) sequences from Lake Nyasa in Tanzania. The middle circle in the haplotype network represents 34 individuals. Other circles' sizes correlate to the number of individuals, as seen on the right side of the haplotype network. The haplotype network shows the proportion of haplotypes from each sampling site.

Results

Haplotypes and nucleotide diversities

The overall nucleotide diversity and haplotype diversity for the 629 bp sequences acquired in this investigation were 0.17± 0.12 and 0.74± 0.02, respectively. The sequences have been submitted to GenBank with the accession numbers OQ318303 to OQ318401. The haplotype diversity ranged between 0.50 and 0.78, while the nucleotide diversity ranged between 0.08 and 0.2%. Nine haplotypes were obtained, with Buloma having the most singleton haplotypes, followed by Lituhi and Ruhuhu (Table 1). The minimum spanning haplotype network resembles a star, with rare haplotypes having only one mutational step from the core haplotype (Fig. 2).

Population genetic structure

Analysis of molecular variances of the COI sequences for Opsaridium microlepis revealed significant genetic differences among the subpopulations studied (Overall ΦST= 0.093, P < 0.05). After performing sequential Bonferroni correction, some of the pairwise FST-values were significantly different (Table 2).

Demographic history

Tajima's D values were negative but

Table 1: *Opsaridium microlepis's* cytochrome oxidase subunit I (COI) nucleotide diversity, haplotype diversity and haplotype distribution along Lake Nyasa, Tanzania.

Sampling sites	Codes	π (%)	h	N	Haplotypes distribution									
					1	2	3	4	5	6	7	8	9	Nh
Kafyofyo	KF	0.20±0.15	0.78±0.07	15	4	6	2	1	2	0	0	0	0	5
Buloma	BL	0.16±0.13	0.74±0.08	16	5	1	1	0	7	1	1	0	0	6
Katumba	KT	0.08±0.08	0.50±0.07	16	6	10	0	0	0	0	0	0	0	2
Lupingu	LP	0.12±0.10	0.60± 0.11	15	9	1	1	0	4	0	0	0	0	4
Ruhuhu	RH	0.18±0.14	0.77±0.07	19	5	2	0	0	8	1	0	1	2	6
Mbambabay	MB	0.10±0.09	0.50 ± 0.10	18	5	12	0	0	0	1	0	0	0	3

π = nucleotide diversity, h= haplotype diversity and Nh = the total number of haplotypes

Table 2: Pairwise fixation index (FST) values for *Opsaridium microlepis* cytochrome oxidase subunit I (COI) sequencing data from Lake Nyasa, Tanzania

CODE	KF	BL	KT	LP	RH	MB
KF						
BL	0.10	-				
KT	0.10	0.42*	-			
LP	0.09	-0.03	0.39*	-		
RH	0.07	-0.03	0.33*	-0.04	-	
MB	0.10	0.42*	-0.05	0.40*	0.35*	-

** indicates pairwise FST-value that is significantly different after sequential Bonferroni correction*

insignificant for Buloma, Lupingu, and Ruhuhu samples. Buloma samples have significant Fu's Fs test results. The Rogers test and mismatch distribution analysis support the idea of sudden population expansion (Table 3). The observed (bars) and expected (line) frequency of pairwise differences is unimodal (Fig. 3).

Effective population size

Buloma had the highest mutation-scaled effective population size, followed by Kafyofyo at the mean and 95% confidence interval. Mbambabay had the smallest mutation-scaled effective population size (Fig. 4).

Table 3: Demographic and neutrality variables derived from *Opsaridium microlepis* sequences from Lake Nyasa, Tanzania

Codes	SSD	HRI	D	Fs
KF	0.01	0.10	1.08	-0.95
BL	0.02	0.16	-0.50	-2.63*
KT	0.02	0.25	1.31	1.25
LP	0.01	0.14	-0.58	-0.99
RH	0.02	0.16	-0.67	-2.00
MB	0.01	0.13	0.00	-0.01

*SSD= Sum of the squared deviations, HRI= Harpending's raggedness index D= Tajima's D and Fs = Fu's Fs. * represents a value with $p < 0.05$.*

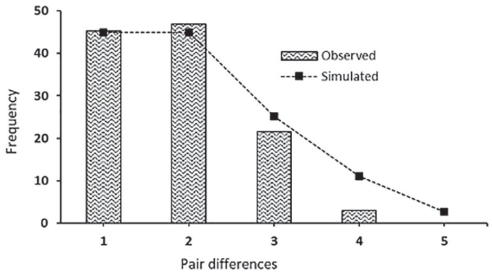

Figure 3: The observed (bars) and expected (line) frequency of pairwise differences in the cytochrome oxidase subunit I sequencing from *Opsaridium microlepis*, along Lake Nyasa, Tanzania.

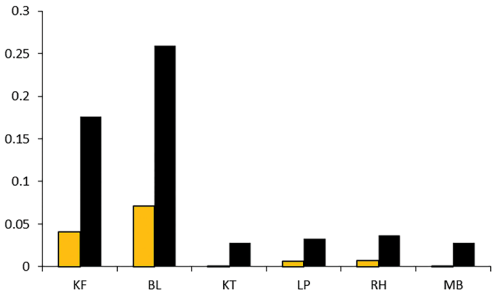

Figure 4: *Opsaridium microlepis's* effective population size estimated using mutation-scaled data from Lake Nyasa, Tanzania. The yellow bars reflect the mean confidence interval, whereas the black bars represent the 95% confidence interval.

Discussion

Population resilience depends on genetic diversity and effective population size (Nikolic *et al.* 2010, Faulks *et al.* 2011). We found variable levels of genetic diversity among the subpopulations investigated. Some subpopulations exhibited lower genetic diversity. Population reduction results in a genetic diversity loss (Petit-Marty *et al.* 2022). However, the overall genetic diversity obtained in the current study is more than that reported for *Opsaridium microlepis* in the same Lake in Malawi (Sungani *et al.* 2016). However, the genetic diversity in the current study is more significant than that of *O. tweddleorum* found in Malawi. These findings may suggest that cyprinid fishes in Lake Nyasa are subjected to various stressors that influence their genetic diversity. Environmental stress during recruitment and unfavourable environmental conditions are linked to lower genetic diversity (Jennifer *et al.* 2011). Indeed, the International Union for Conservation of Nature (IUCN) Red List species classified *O. microlepis* as a vulnerable species due to rising overexploitation, including the use of smaller mesh sizes that catch juveniles, as well as ongoing habitat decline due to deforestation and siltation from soil erosion and water abstraction for irrigation purposes, which hinders the downstream movement of young fish (Tweddle 2018). These variables may result in poor dispersal potential among the *O. microlepis* and limited connectedness due to diminished capacity for recolonization and gene interchanges among the populations (Garcia-Cisneros *et al.* 2016).

The analysis of molecular variances in this study revealed more genetic differentiation among subpopulations of *O. microlepis* compared to that obtained for the species *O. microcephalum,* but lower compared that reported for *O. tweddleorum*, both of which are endemic to Lake Nyasa. However, a research conducted in Malawi along Lake Nyasa found no genetic differentiation among *O. microlepis* species (Sungani *et al.* 2016).

Tajima's D values were not found to be significant across all populations. One group, however, shows significant negative Fu's Fs-values. Tajima's significant D values are a signal of the presence of evolutionary factors, such as a recent increase in population size. Fu's Fs-values are more effective than Tajima's D in detecting the recent effect of rapid population expansion, and significant low Fu's Fs-values indicate recent population growth. Most subpopulations have non-significant negative Fu's Fs values, demonstrating the absence of evolutionary processes that would otherwise create deviations from the genetic marker's neutrality. However, the non-significant sum of square deviations (SSD) (Rogers 1995) supports the idea of recent rapid population growth. For all populations, the raggedness index (Rogers and Harpending 1992) and mismatch distribution results corroborate the recent population expansion model hypothesis. A star-like shape in the haplotype network and an unimodal mismatch distribution for all populations also support the occurrence of recent population increase following a bottleneck impact.

Conclusion

The findings show that some sampling sites have limited nucleotide and haplotype diversity. The Migrates-n v. 3.11 program analysis found that some sampling sites had smaller effective population sizes. This necessitates urgently implementing appropriate management measures to prevent the population from becoming extinct. The IUCN has already designated *O. microlepis* as a threatened species. Future studies will need to use sensitive markers with high resolution to confirm the patterns of genetic diversity and effective population size found in this study.

Acknowledgements

We express our sincere gratitude to the Sokoine University of Agriculture Research and Innovation Support (SUARIS) for the financial support that made this work possible. Without their invaluable support, this project (Project Name: SUARIS1, Funding Number: DCOP06) could not have been realized. Their contribution has been instrumental in advancing our understanding and contributing to the field of conservation molecular genetics. We distinctively acknowledge SUARIS for their noteworthy support.

References

Faulks, L.K., Gilligan, D.M., and Beheregaray, L.B (2011). The role of anthropogenic vs. natural in-stream structures in determining connectivity and genetic diversity in an endangered freshwater fish, Macquarie perch (Macquaria australasica). Evolutionary Applications, 4 (4), 589–601.

Garcia-Cisneros, A., Palacín, C., Ben Khadra, Y., and Pérez-Portela, R (2016). Low genetic diversity and recent demographic expansion in the red starfish Echinaster sepositus (Retzius 1816). Scientific Reports, 6 (August), 33269.

Jennifer, P., Planes S., Rochel E., David, L., and Cécile, F. (2011). Genetic diversity loss associated to high mortality and environmental stress during the recruitment stage of a coral reef fish. Coral Reefs, 30 (2) 399-404.

Markert, J.A., Denise M. Champlin, Ruth Gutjahr-Gobell, Jason S. Grear, Anne Kuhn, Thomas J. McGreevy, Annette Roth, Mark J. Bagley, and Diane E. Nacci (2010). Population genetic diversity and fitness in multiple environments. BMC Evolutionary Biology, 10 (205), 1–13.

Martinez, A.S., Willoughby, J.R., and Christie, M.R (2018). Genetic diversity in fishes is influenced by habitat type and life-history variation. Ecology and Evolution, 8(23), 12022–12031.

Nikolic, N., Butler, A.J., Bagliniere, L.-J., Laughton, R., McMyn, G.A.I., and Chevalet, C (2010). An examination of genetic diversity and effective population size in Atlantic salmon populations. Genetic Research Cambridge, (2009).

Nindi, S.J. (2007). Changing livelihoods and the environment along Lake Nyasa, Tanzania. African Study Monographs, Suppl, 36, 71–93.

Petit-Marty, N., Liu, M., Tan, I.Z., Chung, A., Terrasa, B., Guijarro, B., Ordines, F., Ramírez-Amaro, S., Massutí, E., and Schunter, C (2022). Declining population sizes and loss of genetic diversity in commercial fishes: A simple method for a first diagnostic. Frontiers in Marine Science, 9 (May), 1–11.

Pini, J., Planes, S., Rochel, E., Lecchini, D., and Fauvelot, C (2011). Genetic diversity loss associated to high mortality and environmental stress during the recruitment stage of a coral reef fish. Coral reefs, 30, 399–404.

Rogers, A.R. (1995). Genetic evidence for a pleistocene population explosion. Evolution, 49 (4), 608–615.

Rogers, A.R. and Harpending, H (1992). Population growth makes waves in the distribution of pairwise genetic differences. Molecular Biology and Evolution, 9 (3), 552–569.

Sungani, H., Ngatunga, B.P., and Genner, M.J (2016). Migratory behaviour shapes spatial genetic structure of cyprinid fishes within the Lake Malawi catchment. Freshwater Biology, 61, 1062–1074.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution, 30 (12), 2725–2729.

Tweddle, D (2018). Opsaridium microlepis, Lake Salmon. The IUCN Red List of Threatened Species 2018:, 8235.

Villesen, P (2007). FaBox: An online toolbox for FASTA sequences. Molecular Ecology Notes, 7 (6), 965–968.