



Population Genetic Structure and Demographic History of *Opsaridium microcephalum* along Lake Nyasa

Alex Nehemia¹ and Alinanuswe J. Mwakalesi²

¹Department of Biological Sciences, College of Natural and Applied Sciences, Sokoine University of Agriculture, P.O. Box 3038, Morogoro, Tanzania

²Department of Chemistry and Physics, College of Natural and Applied Sciences, Sokoine University of Agriculture, P.O. Box 3038, Morogoro, Tanzania

*Corresponding author: nehemiah@sua.ac.tz

Received Sept 2023 Revised Dec 2023 Accepted Jan 2024 Published Jan 2024

<https://dx.doi.org/10.4314/tjs.v49i5.8>

Abstract

The Sanjika, *Opsaridium microcephalum* is among the fish species that serve as a major source of proteins and income to people along Lake Nyasa. However, the information on its genetic diversity and structure particularly in the Tanzania part of the Lake is limited. Therefore, the assessment of the genetic diversity and structure of *O. microcephalum* along the Lake Nyasa part of Tanzania was conducted in the current study using a fragment of the mitochondrial COI gene from 92 individuals of *O. microcephalum*. The findings showed that an average haplotype diversity and average nucleotide diversity were 0.8065 ± 0.0314 and 0.002380 ± 0.001640 , respectively. The molecular Variance (AMOVA) indicated significant genetic variations among the subpopulations studied (Overall $\Phi_{st} = 0.31560$, $p < 0.001$). The lowest haplotype diversity was recorded at Manda and the highest haplotype diversity was recorded at Buloma. The lowest nucleotide diversity was recorded at Manda and the highest was recorded at Lupingu. The differences in genetic diversity can be a possible indicator of different localised evolutionary forces that require attention to conservationists for the sustainable management of *O. microcephalum*.

Keywords: Lake Nyasa, genetic diversity, demographic history, population genetics structure and effective population size

Introduction

The *Opsaridium microcephalum* is one of the three endemic species of *Opsaridium* (Cyprinidae) which is riverine, potamodromous and lake spawning populations occurring in Lake Nyasa and its tributaries. The other members of Cyprinidae which are endemic to Lake Nyasa catchments are the kabyabya (*Opsaridium tweddleorum*) which is a small species restricted to inflowing rivers and streams and the mpasa (*Opsaridium microlepis*) which is a large lake-dwelling species that ascends to the lower and reaches of inflowing rivers during the rains to spawn (Tweddle and Turner 2014). While *Opsaridium microcephalum*

and *Opsaridium microlepis* are large-bodied species which is endemic to the catchment and migrate from the lake to rivers to breed, the *Opsaridium tweddleorum* is small-bodied that is only found in streams and rivers (Sungani *et al.* 2017).

The *Opsaridium microcephalum* species has been observed to grow to more than 30cm in total length (TL) and is commercially important around Lake Nyasa (Morioka and Matsumoto 2003). Despite its importance, the population trend has been reported by The IUCN Red List of Threatened Species to decrease (Morioka and Matsumoto 2003, Tweddle 2019). The previous study conducted in Lake Nyasa in

Malawi reported the existence of some genetic structure in this species (Sungani et al. 2016). However, no genetic study has been conducted on the *Opsaridium microcephalum* in Lake Nyasa in Tanzania. This study sought to assess the current genetic diversity and structure of *Opsaridium microcephalum* along Lake Nyasa in Tanzania using partial mitochondrial cytochrome oxidase subunit I (COI) sequences. The study also aimed at assessing the demographic history of and estimation of the effective population size of this species along Lake Nyasa in Tanzania. Assessment of the genetic variation of fish species is important because it helps the researchers to understand the abilities of the fish species to adjust to changing environment which is important for their survival (Mahboob et al. 2019). Estimation of genetic diversity in COI is useful in predicting declines in fish species by fishing pressures because the decline in populations might result in a decrease in genetic diversity important for the adaptive potential to face current and future environmental changes (Petit-Marty et al. 2022). Determining the effective population size (N_e) is an important parameter in the conservation of genetic diversity (Nikolic et al. 2010). It is essential to ascertain a given species' status or population size to conserve

and manage them effectively. However, examining past population size changes in particular provides an evolutionary perspective on the dynamics of the current population (Faulks et al. 2022). It is also crucial to evaluate fish species' population structures because ignoring them in fisheries management could result in unexpected dangers of overexploitation (Ying et al. 2011). The COI barcode region can give important insights into a species' evolutionary processes, demographic history, structure connectivity, and population genetic diversity (Bucklin et al. 2021)

Materials and methods

Sampling

Samples of *Opsaridium microcephalum* were collected from fishermen in six landing sites located along Lake Nyasa (Figure 1), between February and November 2020. Twenty samples of *O. microcephalum* were collected from each sampling sites, stored in absolute ethanol (99.9%) and transported to the Sokoine University of Agriculture for molecular analysis. Other fish species that were also found in the sampling sites were *Oreochromis Karongae* and *Opsaridium microlepis*.

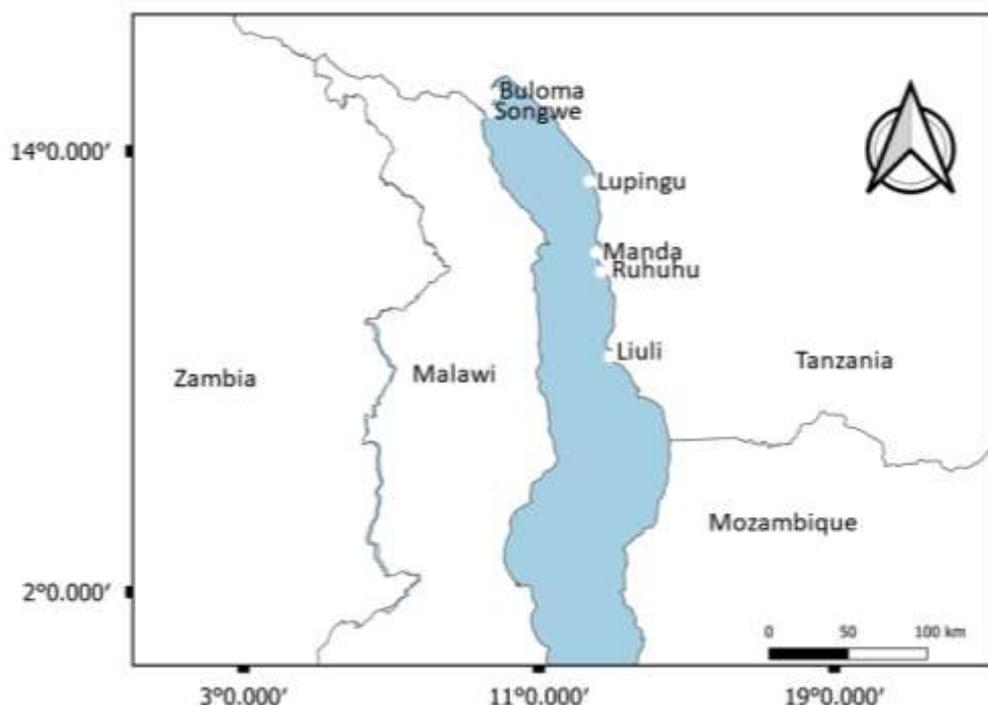


Figure 1. A map showing sampling sites of the *Opsaridium microcephulum* along Lake Nyasa: the white circles indicates sampling sites

DNA extraction

About 30 mg of sample tissue of *Opsaridium microcephulum* preserved in 99.9% ethanol was used for the extraction of DNA following instructions provided by the protocol of Quick-DNA™ Miniprep Plus Kit (ZYMO Research) and 2% TBE agarose gels were used to visualize the extracted DNA.

Mitochondrial DNA (mtDNA) analysis

The primers FishF1: 5'-TCAACCAACCACAAAGACAT TGGCAC-3' and FishR2: 5'-ACTT CAGGGTGACCGAAGAATCAGAA-3' (Ward et al. 2005), were used to amplify a segment of the COI gene containing 578 base pairs using a T100™ Thermal cycler device (Bio-Lab Inc, GA, USA). Polymerase chain reactions (PCR) were performed in a total volume of 35 µL, which included 2 µL of DNA template, 1x Multiplex PCR Master Mix, 0.4mg of BSA, 0.3 M of each primer, and 11.7 µL of RNase-free water. The temperature profile that was applied to the

reaction mixture involved an initial temperature of 94°C for 5min, followed by 35 cycles of 94°C for 40sec, 54°C for 45sec, and 72°C for 1min, with a final extension of 72°C for 15min. The success of sample amplification was visualized using a 1% agarose gel. The PCR products were sequenced at the Macrogen Europe company using an automated sequencer (AB 3730XL; Applied Biosystems, Foster City, USA) using the FishF1 and FishR2 primers. Sequences from *Opsaridium microcephulum* were edited using the software MEGA 11. The editing involved removing sequences with bad quality and trimming the ends. The final sequences consisted of the final length of 578bp and are deposited in GenBank with the accession numbers OQ991382-OQ991473. The software MEGA 11 software was also used to inspect the presence of stop codons which might be indicating sequencing errors or false genes. The online program BLAST was then used for species identity (Wheeler and Bhagwat 2007). The program CLUSTAL

W (Tamura et al. 2013) implemented in the software MEGA 11 was used for multiple alignments of the sequences and the online FaBox 1.41 (Villesen 2007) was used to collapse haplotypes. The software Pop Art v.1.7 (Leigh and Bryant 2015) was used for the haplotype network analysis. The software Arlequin v. 3.5.2.2 (Excoffier and Lischer 2010) was used for assessing genetic diversity, historical demography, and neutrality parameters. Analysis of molecular variance (AMOVA) implemented in Arlequin v. 3.5.2.2 was used for analysing genetic differentiations among the populations. Hierarchical AMOVA was carried out to test if there is a population genetic structure for subpopulations collected along the Lake in the part of Tanzania and Malawi.

Results

Haplotypes and nucleotide diversities

The overall haplotype and nucleotide diversity were 0.8065 ± 0.0314 and 0.002380 ± 0.001640 , respectively. The highest haplotype diversity was observed at Buloma and the lowest was recorded at Manda. The highest nucleotide diversity was recorded at Lupingu and Buloma while the lowest at Manda (Table 1). The haplotype numbers are 38 with private haplotypes accounting for 73.68%. The haplotype network consists of three star-like structures with most of the rare haplotypes in the minimum spanning network having only one mutational step from the central haplotype (Figure 2). The haplotype number varied from 5 to 10, with Buloma recording the highest number (Table 2).

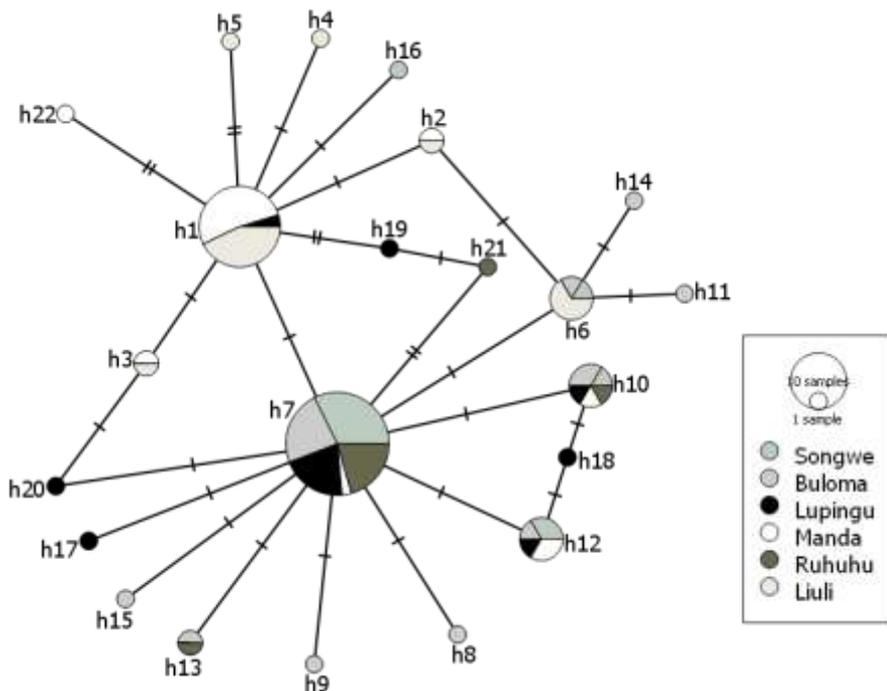


Figure 2: Haplotype network of partial mitochondrial cytochrome oxidase subunit I (COI) sequences from the *Opsaridium microcephalum* in Tanzania, Lake Nyasa.

Tajima's D values were negative for samples collected from all sites and significant for samples from Songwe, Buloma and Manda. Negative significant Fu's Fs test values were also observed for samples

from all sites. The Rogers test and mismatch distribution analysis supported the theory of recent sudden population expansion (Table 1). The mismatch distribution frequency is a unimodal graph (Figure 3).

Table 1: Nucleotide and haplotype diversity, demographic and neutrality parameters based on cytochrome oxidase I (COI) sequences from the *Opsaridium microcephulum* from Tanzania, Lake Nyasa. π : nucleotide diversity, h : haplotype diversity, SSD: sum of squared deviations, HRI: Harpending's raggedness index, D: Tajima's D and Fs: Fu's Fs

Sampling sites	Codes	π (%)	h	SSD	HRI	D	Fs
Songwe	SN	0.1269	0.5333	0.006	0.100	-1.69167*	-2.23705*
Buloma	BL	0.2216	0.8105	0.022	0.160	-1.80873*	-
Lupingu	LP	0.2358	0.7692	0.022	0.250	-1.40568	-
Manda	MD	0.1153	0.4762	0.013	0.135	-	-
Ruhuhu	RH	0.1749	0.7033	0.017	0.161	1.91084**	2.59631**
Liuli	LU	0.1516	0.6286	0.008	0.131	-1.23596	-1.60596
						-1.45121	-1.86416*

* Indicates $p < 0.05$ and ** $p < 0.001$

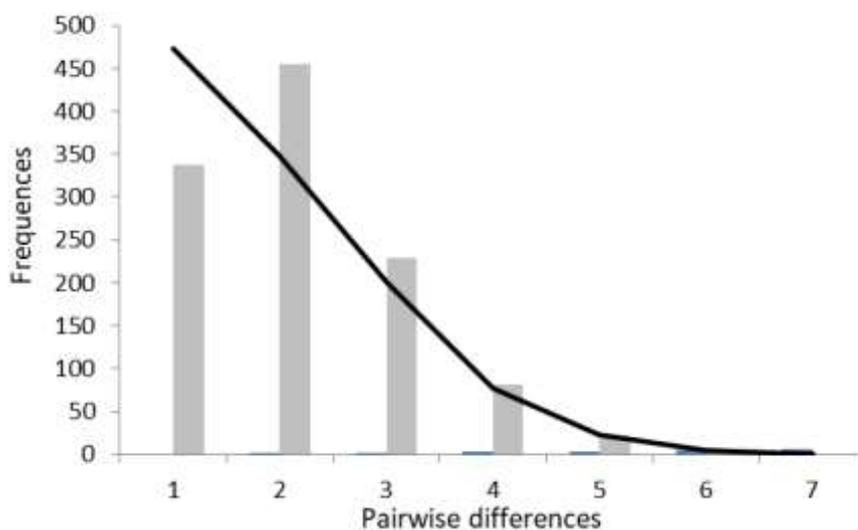


Figure 3: The observed (bars) and expected (line) frequency of pairwise differences of the cytochrome oxidase subunit I sequence from the *Opsaridium microcephulum* in Tanzania, Lake Nyasa.

Table 2: The haplotype distribution of *Opsaridium microcephulum* cytochrome oxidase subunit I (COI) along Lake Nyasa in Tanzania; N is the number of samples, and Nh is the total number of haplotypes at each site. For site codes see Table 1.

Codes	N	Haplotype																				Nh		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		21	22
SN	16						1	11			1		2				1							5
BL	18						1	8	1	1	2	1	1	1	1									10
LP	14	1						7			1		1					1	1	1		1		8
MD	15	11	1	1				1																5
RU	14						4	7			1			1								1		5
LU	15	9	1	3	1	1																		5
Total	92	21	2	4	1	1	6	34	1	1	5	1	4	2	1	1	1	1	1	1	1	1	1	38

The highest mean mutational-scaled effective population sizes was recorded at Buloma (0.037) and lowest at Songwe (0.0037). The mean mutational-scaled effective population sizes were 0.029 at Lupingu, 0.0056 at Manda, 0.0039 at Liuli and 0.0043 at Ruhuhu.

Genetic population structure and demographic history

Table 3. Pairwise Φ_{st} -values for cytochrome oxidase subunit I (COI) sequence data in the *Opsaridium microcephulum* from Tanzania, Lake Nyasa. For site codes see Table 1.

CODE	SN	BL	LP	MD	RU	LU
SN	-					
BL	-0.01596	-				
LP	-0.02727	0.00915	-			
MD	0.51678*	0.46351	0.38426*	-		
RU	0.03461	-0.02572	0.04666	0.51820*	-	
LU	0.52977*	0.48402*	0.40264*	-0.00914	0.52894*	-

* indicate significance value after bonferroni correction

Discussion

The overall haplotype diversity of *Opsaridium microcephulum* recorded in this study was higher than the one recorded in the same lake in Malawi but the nucleotide diversity is lower than the one recorded in Malawi. However, the genetic diversity obtained in this study is higher than previously reported genetic diversity of *O. microlepis* and *O. tweddleorum* (Sungani et al. 2016). The difference in genetic diversity observed by the author of the study conducted in Malawi was suggested to be due to divergent life-history evolution (Sungani et al. 2016).

Effective population size (Ne) is an important parameter in the conservation of genetic diversity (Nikolic et al. 2010). Genetic diversity is one of three levels of biological diversity requiring conservation. Genetic theory predicts that levels of genetic variation should increase with effective

The analysis of molecular variances (AMOVA) of the COI sequences revealed population genetic differentiation among the population (Overall $\Phi_{st} = 0.31560, p < 0.001$). However, the pairwise Φ_{st} values revealed the genetic differences between all the samples from Manda with samples from other sites. The same trend is also observed for all samples from Liuli (Table 3).

population size (Frankham 1996). However, the theory can be violated when dealing with an open population (Palstra and Ruzzante 2008). The lack of a positive relationship between effective population size and genetic diversity in the present study may be partly explained by the same reason. Since the *O. microcephalum* is assigned as Least Concern by the IUCN RedList (Tweddle 2019) the populations are not threatened and therefore are open populations with extensive gene flow in the region. However, the difference between the global test observed in this species with other species found in the same area may be due to contrasting migration life histories among these species (Sungani et al. 2016).

The analysis of molecular variances (AMOVA) demonstrated higher population genetic differentiation among the population with pairwise F_{ST} values indicating genetic differences between Manda samples and with

most other samples. The same pattern was observed for Liuli samples, which also showed pairwise F_{ST} values that suggested genetic differences with most other samples. A global test with F_{ST} results greater than 0.15 suggests greater genetic differentiation (Wright 1978). The results of this study demonstrated greater genetic differentiation than previously reported in Malawi. However, the global test achieved in this study is higher than that reported for this species and *O. microlepis*, but lower than that recorded for *O. tweddleorum* in the same study area (Sungani et al. 2016). Connectivity is vital in maintaining and restoring natural ecological processes, and genetic diversity contributes to the adaptation and persistence of any species in the face of global perturbations (Sahyoun et al. 2016). The higher F_{ST} value recorded in this species reveals the presence of factors that may be limiting gene flow among the populations. Environmental conditions and habitat quality can mediate gene flow by providing different levels of resistance to inter-population dispersal (Chiu et al. 2023). The study previously conducted in the study area demonstrated the presence of human activities that were likely to degrade some catchments that acts as a spawning area for fish species (Nindi 2007). Differences in habitat quality among the sampling sites may be one of the factors that have contributed to the observed higher F_{ST} value in this species.

The Tajima's D values were negatively significant for samples from Songwe, Buloma and Manda, indicating the presence of evolutionary forces, such as a recent increase in population size following the bottleneck event, that produce deviations from the genetic marker's neutrality. However, the negative non-significant Tajima's D values obtained in samples from other populations may indicate an excess of rare nucleotide site variants compared to what would be expected under a neutral evolution model (Joshi et al. 2013). Negatively significant F_u 's F_s results, which are based on the distribution of haplotypes recorded for samples from all sites, also

suggest the presence of recent population expansion in this species.

The non-significant sum of square deviations (SSD) discovered was in favour of the rapid population growth scenario (Rogers 1995). The raggedness index supports the theory of the recent population expansion model for all populations (Rogers and Harpending 1992). We did, however, discover a star-like pattern in the haplotype network, which could imply recent population growth following a bottleneck.

Conclusion

The results obtained in this study demonstrated that the genetic diversity of *Opsaridium microcephulum* from Tanzania differs from that of Malawi in the same Lake. Genetic diversity and effective population size were found to differ among the sampling locations, with certain sites having the lowest, indicating the presence of factors that may be driving genetic erosion at some sampling sites. The genetic diversity was shown to have a negative association with effective population size, indicating that the area had open populations with considerable gene flow. The increased population genetic differentiation indicated by analysis of molecular variances (AMOVA) may suggest the presence of processes that cause genetic erosion and, as a result, contribute to the difference in genetic variations observed among the sampling sites. This could be a call to conservationists to improve fishing management and promote the conservation of fish habitats in the area. However, more research with a large number of samples and sensitive molecular markers like microsatellites is needed to confirm the patterns of genetic diversity and effective population sizes 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 SUARIS 1 could not have been

realized. Their contribution has been instrumental in advancing our understanding and contributing to the field of Molecular genetics. We distinctively acknowledge SUARIS for their noteworthy support and Mr. Nguliyati Maulidi for assistance with molecular analysis.

References

- Bucklin A, Peijnenburg KTCA, Kosobokova KN, O'Brien TD, Blanco-Bercial L, Cornils A, Falkenhaus T, Hopcroft RR, Hosiá A, Laakmann S, Li C, Martell L, Questel JM, Wall-Palmer D, Wang M, Wiebe PH and Weydmann-Zwolicka A 2021 Toward a global reference database of COI barcodes for marine zooplankton. *Mar. Biol.* 168:1–26.
- Chiu MC, Nukazawa K, Resh VH and Watanabe K 2023 Environmental effects, gene flow and genetic drift: Unequal influences on genetic structure across landscapes. *J. Biogeogr.* 50: 352–364.
- Excoffier L and Lischer HEL 2010 Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10: 564–567.
- Faulks L, Kaushik P, Taniguchi S, Sekino M, Nakamichi R, Yamamoto Y, Fujimori H, Okamoto C, Kodama S, Daryani A, Manwong AF, Galang I, Mochioka N, Araki K, Suzuki M, Kaji Y, Ichiki T, Matsunaga T and Hakoyama H 2022 Inferring the demographic history of Japanese eel (*Anguilla japonica*) from genomic data: Insights for conservation and fisheries management. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 32:1092–1098.
- Frankham R 1996 Relationship of genetic variation to population size in wildlife. *Conserv. Biol.* 10:1500–1508.
- Joshi J, Salar RK, Banerjee P, Upasna S Tania M S and Vijn RK 2013 Genetic variation and phylogenetic relationships of Indian buffaloes of Uttar Pradesh. *Asian-Australasian J. Anim. Sci.* 26: 1229–1236.
- Leigh JW and Bryant D 2015 POPART : full-feature software for haplotype network construction. *Methods. Ecol. Evol.* 6: 1110–1116.
- Mahboob S, Al-ghanim KA, Al-misned F, Al-balawi HFA, Ashraf A and Al-mulhim NMA 2019 Saudi Journal of biological sciences genetic diversity in tilapia populations in a freshwater reservoir assayed by randomly amplified polymorphic DNA markers. *Saudi. J. Biol. Sci.* 26: 363–367.
- Morioka S and Matsumoto S 2003 Otolith features and utility of lapillus for daily increment analysis in *Opsaridium microcephalum* (Cyprinidae) juveniles collected from Lake Malawi. *Ichthyol. Res.* 50: 82–85.
- Nikolic N, Butler AJ, Bagliniere LJ, Laughton R, McMyn GAI and Chevalet C 2010 An examination of genetic diversity and effective population size in Atlantic salmon populations. *Genet. Res. Cambridge.* (2009).
- Nindi SJ 2007 Changing livelihoods and the environment along lake Nyasa, Tanzania. *African. study. Monogr. Suppl.* 36:71–93.
- Palstra FP and Ruzzante DE 2008 Genetic estimates of contemporary effective population size: What can they tell us about the importance of genetic stochasticity for wild population persistence? *Mol. Ecol.* 17: 3428–3447.
- Petit-Marty, N, Liu M, Tan IZ, 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. *Front. Mar. Sci.* 9: 1–11.
- Rogers AR 1995. Genetic evidence for a pleistocene population explosion. *Evolution.* 49: 608–615.
- Rogers AR and Harpending H 1992 Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9: 552–569.
- Sahyoun R, Guidetti P, Di Franco A and Planes S 2016 Patterns of fish connectivity between a marine protected area and surrounding fished areas. *PLoS One.* 11: 1–15.
- Sungani H, Ngatunga BP and Genner MJ

2016. Migratory behaviour shapes spatial genetic structure of cyprinid fishes within the Lake Malawi catchment. *Freshw. Biol.* 61: 1062–1074.
- Sungani H, Ngatunga BP, Koblmüller S, Mäkinen T, Skelton PH and Genner MJ 2017 Multiple colonisations of the Lake Malawi catchment by the genus *Opsaridium* (Teleostei: Cyprinidae). *Mol. Phylogenet. Evol.* 107: 256–265.
- Tamura K, Stecher G, Peterson D, Filipski A and Kumar S 2013 MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30: 2725–2729.
- Tweddle D 2019 *Opsaridium microcephalum* (amended version of 2018 assessment). *The IUCN Red List of Threatened Species 2019: e.T60845A155041563*.
- Tweddle D and Turner GF 2014 Evidence for lacustrine breeding by sanjika *Opsaridium microcephalum* (Teleostei: Cyprinidae) in Lake Malawi. *African. J.Aquat. Sci.* 39: 479–480.
- Villesen P 2007 FaBox: An online toolbox for FASTA sequences. *Mol. Ecol. Notes.* 7: 965–968.
- Ward RD, Zemlak TS, Innes BH, Last PR and Hebert PDN 2005 DNA barcoding Australia’s fish species. *Philos. Trans. R. Soc. B. Biol. Sci.* 360: 1847–1857.
- Wheeler D and Bhagwat M 2007 Comparative Genomics Volumes 1 and 2. In: BLAST QuickStart Example-Driven Web-Based BLAST Tutorial. *Methods. Mol. Biol.* 395:149-176.
- Wright S 1978 Evolution and the genetics of populations, variability Within and among natural populations. Chicago: The University of Chicago Press.
- Ying Y, Chen,Y, Lin L and Gao T 2011 Risks of ignoring fish population spatial structure in fisheries management. *Can. J. Fish. Aquat. Sci. Aquat. Sci.* 68: 2101–2120.