academicJournals

Vol. 12(48), pp. 6709-6717, 27 November, 2013
DOI: 10.5897/AJB2013.12459
ISSN 1684-5315 ©2013 Academic Journals
http://www.academicjournals.org/AJB

African Journal of Biotechnology

*Full Length Research Paper*

# Genetic diversity and structure of potamodromous *Opsaridium microlepis* (Günther) populations in the in-let rivers of Lake Malawi

**Wisdom Changadeya[1]*, Hamad Stima[2], Daud Kassam[2], Emmanuel Kaunda[2] and Wilson Jere[2]**

[1]Molecular Biology and Ecology Research Unit (MBERU) DNA Laboratory, Department of Biological Sciences, Chancellor College, University of Malawi, P.O BOX 280, Zomba, Malawi.
[2]Aquaculture and Fisheries Science Department, Bunda College of Agriculture, Lilongwe University of Agriculture and Natural Resources, P.O. Box 219, Lilongwe, Malawi.

Accepted 21 November, 2013

**Studies were carried out to determine the genetic diversity and structure of endangered *Opsaridium microlepis* (Mpasa) populations in the affluent rivers of Lake Malawi; Linthipe, Bua, Dwangwa and North Rukuru. A total of 200 DNA samples of *O. microlepis* from four river populations were analyzed at 20 microsatellite loci. The primers' discriminating power was high (mean PIC, 0.76) yielding a total of 295 alleles with a range of 10-22 and an average of 15 alleles per locus. All the populations were not in Hardy Weinberg Equilibrium probably due to outbreeding that leads to heterozygosity excess. This observation was further supported by heterozygosity excess exhibited by 100% of the population-locus combinations (mean $F_{IS}$, -0.30) and lack of evidence for genetic bottleneck. The populations exhibited high genetic diversity as evidenced by high mean Shannon information Index (I=1.64) and high observed heterozygosity ($H_o$ = 0.98). Genetic relationships among the populations appear to be less influenced by geographical distance (Mantel's statistics Z, 0.18; *p* = 0.6369) implying that the populations do not fit into the isolation by distance model. Nevertheless, the populations are highly differentiated ($F_{ST}$ = 0.17; AMOVA among populations = 16%). This is supported by inter-deme migration of less than one individual per generation (*Nm*=0.91) as determined by Slatkin' private allele method. Therefore, these populations are probably still large and distinct requiring separate monitoring and management due to inferred restricted gene flow and considerable population differentiation.**

**Key words:** Mpasa, Lake Malawi, population structure, genetic diversity, microsatellites, threatened species, conservation.

## INTRODUCTION

The genetic diversity and population structuring of native species populations is an important feature to be considered in conservation and management programs. The use of molecular techniques to assess patterns of genetic variation has supported conservation programs, by indicating which species require greater conservation efforts and also by selecting areas where natural populations are viable (Johnson et al., 2001). Recently, molecular markers have been commonly used in population studies. Simple sequence repeat (SSR) markers are preferable because they are co dominant, highly polymorphic and are considered the most powerful in terms of their resolving power (Creste et al., 2004; Buhariwalla et al., 2005). In addition, microsatellites have a wide distribution in the

*Corresponding author. E-mail: wchanga@cc.ac.mw or wisdomchangadeya@gmail.com. Tel: +265 888 877 858.

genome and can be easily identified, making them ideal markers for genetic variability studies of populations (Collevatti et al., 2001). The polymorphism obtained with microsatellite markers has provided useful information necessary for management of fish stocks (Alam and Islam, 2005), population analysis and biodiversity conservation (Romana-Eguia et al., 2004; Changadeya et al., 2012).

The fisheries production from the wild is generally declining and with an increase in human population, the pressure is even greater on the resource (Government of Malawi, 1998). Most of the fish species in Lake Malawi including cichlids and cyprinids are endemic. Studies by Ambali et al. (2002) have also shown that there are species flocks that are endemic to specific areas in the lake. Pollution, lakeshore development and other human activities have degraded critical habitats for the spawning and early life history stages of many fishes. In addition, recruitment, overfishing and reduction of spawning stock below a critical threshold have prevented many populations from rebuilding to previous levels of abundance. Riverine potamodromous cyprinids species are most vulnerable because they are targeted by fishermen during the spawning period when they swim up stream. This development has led to the fishery of some in-let rivers of Lake Malawi such as North Rukuru, Bua, Dwangwa and Linthipe to be threatened with extinction. Though there are several commercially important riverine cyprinids such as Mpasa (*Opsaridium microlepis*) that are under threat, they have been studied to a limited extent resulting in inadequate knowledge of their biogeography, origin and genetic diversity (Snoeks, 2004).

Mpasa, belong to the Cyprinidae family and is endemic to Lake Malawi. This species was described by Günther in 1864, and is mostly caught while on spawning runs in large rivers (Lowe, 1975; Tweddle, 1981). In recent years, *O. microlepis* catches in Lake Malawi and its in-let rivers have declined (Anon, 2008). The inclusion of Mpasa in the 2006 IUCN Red List of Threatened species (IUCN, 2006) was because it had been heavily fished during its spawning runs up rivers and was considered to have consequently suffered approximately 50% decline in the past ten years, due to overfishing, habitat degradation and pollution (Kazembe et al., 2005). This is a cause for concern in the conservation of Mpasa and development of fisheries in Malawi in general. Therefore, the objective of this study was to investigate genetic diversity and population structuring among populations of *O. microlepis* in the Lake Malawi' river in-lets of Bua, Dwangwa, Linthipe and North Rukuru using microsatellite markers.

## MATERIALS AND METHODS

### Sample collection

Muscles tissues of 200 *O. microlepis* specimens (50 each from North Rukuru, Bua, Dwangwa and Linthipe river) (Figure 1) were collected in February and April 2010. Tissue of about 5-10 mm$^2$ was extracted from individual fish and preserved in vials with 95% ethanol. The samples were brought to Molecular Biology and Ecology Research Unit DNA Laboratory, Department of Biological Sciences, Chancellor College, University of Malawi, where DNA analysis was conducted using 20 microsatellite markers.

### DNA extraction and amplification

Total genomic DNA was extracted from *O. microlepis* samples following a standard SDS- proteinase K/phenol-chloroform protocol as described by Hillis et al. (1996). The study used 20 polymorphic microsatellite loci (Table 1). Amplification of DNA was done in 13.11 µL final reaction volume with 5.7 µl PCR grade water, 1 µl of 10 mM dNTP mix, 1.25 µl of 10 x PCR buffer, 1.6 µl of 25 mM magnesium chloride ($MgCl_2$), 0.75 µl of 15 pmol of both forward and reverse primers, 0.06 µl of 5u/µl *Taq* DNA Polymerase stored in buffer A (Promega, USA) and 2 µl of 25ng/ µl template DNA. The amplifications were carried out in a Master-cycler gradient 5331 Eppendorf Version 2.30.31-09 with the following PCR conditions: initial denaturation at 94°C for 2 min, then 10 amplification cycles of denaturing at 94°C for 30 s, annealing at an optimal temperature for a specific primer pair for 15 s and elongation at 72°C for 30 s. This was followed by 30 amplification cycles of denaturing at 89°C for 30 s, annealing at an optimal temperature for a specific primer pair for 15 s and elongation at 72°C for 30 s. The final extension was at 72°C for 20 min followed by a soaking temperature of 4°C.

### Electrophoresis of PCR products on polyacrylamide gel using silver staining technique

The amplified PCR products were resolved using 6% polyacrylamide gel electrophoresis. The 6% polyacrylamide gel was poured in BIORAD Sequi-Gen® GT Nucleic Acid Electrophoresis Cell. A total of 6 µL of STR 3X Loading Solution (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol FF) was added to the PCR products in 0.2 mL PCR tubes and denatured in Mastercycler gradient at 95°C for 5 min. Then 6 µL of denatured PCR products was loaded on the 6% polyacrylamide gel and ran at 50 W. The gel plates were fixed, stained and developed following procedures in the Promega Silver Sequence™ DNA Sequencing System Technical Manual. The microsatellites bands were scored over a light box using pGem DNA marker (Promega, USA) and $\phi$ X174 DNA/*Hinf* 1 (Promega, (USA) as band size standard markers.

### Data analysis

Polymorphism information content (PIC), a variability measure of each locus, was calculated as described by Zhao et al. (2007):

$$PIC = 1 - \left( \sum_{i=1}^{n} p_i^2 \right) - \sum_{j-1}^{k-1} \sum_{j=i+1}^{n} 2P_i^2 P_j^2$$

Where, $p_i$ is the frequency of the *i*th allele out of the total number of alleles at a microsatellite locus, and *n* is the total number of different alleles for that locus.

Genepop on the web (Raymond and Rousset, 1995) was used to conduct the following analyses: test for conformity to Hardy-Weinberg Equilibrium (Haldane, 1954; Weir, 1990; Guo and Thompson, 1992), test for genotyping linkage equilibrium using Fisher' method, tests of genic and genotypic differentiation, estimation of effective number of migrants using Slatikin's private allele method (Slatkin, 1985) and computation of Wrights statistics (Wright, 1969).

Bottleneck computer software version 1.2.02 (Cornuet and

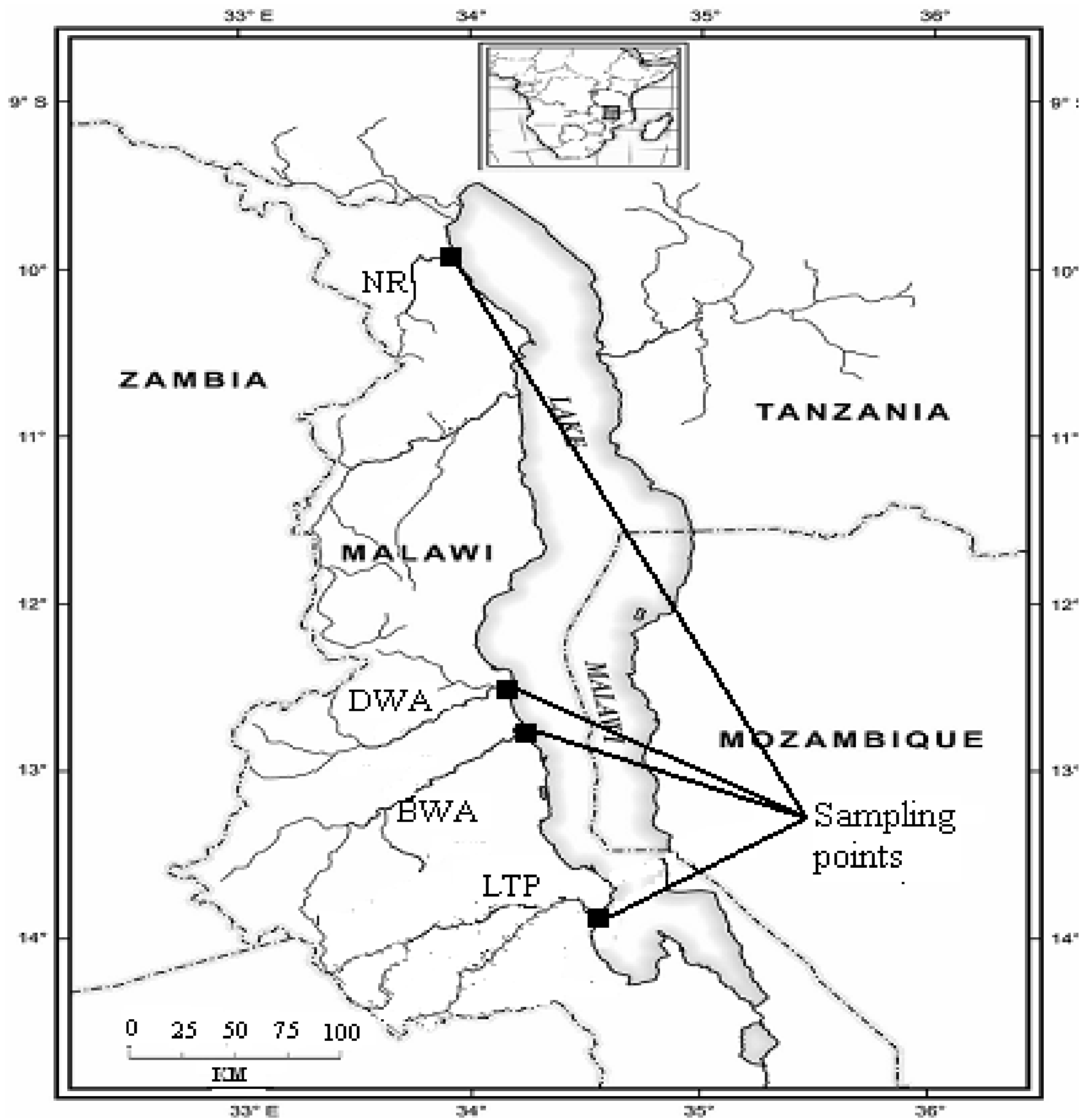


**Figure 1.** Rivers from which *O. microlepis* populations were sampled. NR, North Rukuru River; BWA, Bua River; DWA, Dwangwa River; LTP; Linthipe River (Source: www.map of the world.com).

**Table 1.** Total number of alleles (A) and allele size range (SR) in base pairs and microsatellite primer polymorphic information content (PIC).

| Locus | A | Gene Bank accession | $T_A$ | Allele size range (SR) | Repeat motif | Primer source (reference) | PIC |
|---|---|---|---|---|---|---|---|
| CypG49 | 12 | AY349167 | 53.6 | 170-192 | $(TA)_{11\text{-}21}$ | Baerwald and May (2004) | 0.74 |
| CypG3 | 11 | AY349122 | 56.4 | 150-202 | $(CAGA)_2$ | Baerwald and May (2004) | 0.76 |
| CypG13 | 13 | AY349132 | 53.4 | 140-172 | $(TAGA)_{10}$ | Baerwald and May (2004) | 0.73 |
| Ca3 | 10 | AF277575 | 54.5 | 162-182 | | Dimsoski et al., (2000) cited by Hamilton and Tyler (2008) | 0.78 |
| CypG5 | 12 | AY349124 | 54.3 | 114-172 | $(TAGA)_{12}$ | Baerwald and May (2004) | 0.76 |
| CypG4 | 17 | AY349123 | 54.3 | 134-202 | $(TAGA)_{12}$ | Baerwald and May (2004) | 0.79 |
| CypG30 | 22 | AY349148 | 54.3 | 118-182 | $(TAGA)_7$ | Baerwald and May (2004) | 0.77 |

**Table 1.** Contd.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Lid1 | 17 | | 55.3 | 140-194 | (TTTC)7 | Barinova et al., (2004) cited by Hamilton and Tyler (2008) | 0.75 |
| CypG48 | 16 | AY349166 | 55.3 | 126-158 | | Baerwald and May (2004) | 0.80 |
| MFW | 11 | EF144124 | 55.3 | 106-126 | | Mohindra et al., (2005) | 0.76 |
| CypG22 | 12 | AY349140 | 55.7 | 202-224 | | Baerwald and May (2004) | 0.70 |
| CypG6 | 12 | AY349125 | 52.5 | 192-218 | $(TAGA)_7$ | Baerwald and May (2004) | 0.71 |
| CypG8 | 18 | AY349127 | 52.5 | 132-166 | $(CAGA)_6$ | Baerwald and May (2004) | 0.73 |
| CypG21 | 14 | AY349139 | 52.7 | 158-184 | $(CAGA)_6$ $(TAGA)_7$ | Baerwald and May (2004) | 0.75 |
| CypG27 | 14 | AY349145 | 52.7 | 104-188 | $(TAGA)_8$ | Baerwald and May (2004) | 0.76 |
| Lid11 | 16 | | 53.7 | 200-228 | (TTTG)8 | Barinova et al., (2004) cited by Hamilton and Tyler (2008) | 0.77 |
| AP1 | 13 | AJ428582 | 53.4 | 158-182 | $(TA)_{11-21}$ | Hsu et al.,(2004) | 0.74 |
| AP2 | 18 | AJ428583 | 55.0 | 110-188 | $(AC)_{18-20}$ | Hsu et al.,(2004) | 0.78 |
| Ru2 | 17 | | 53.6 | 142-174 | | Barinova et al., (2004) cited by Hamilton and Tyler (2008) | 0.75 |
| CypG15 | 20 | AY349134 | 53.8 | 116-158 | | Baerwald and May (2004) | 0.82 |
| Mean | 14.75 | | | | | | 0.76 |

$T_A$, annealing temperature.

Luikart, 1996) was employed to determine if the populations in the study had undergone recent effective population size reductions (Genetic bottlenecks). The bottleneck tests estimations were based on 1000 replications using three mutation models; Infinite Allele Model (IAM), Stepwise Mutation Model (SMM) and Two-Phased Model (TPM). Three statistical tests were employed namely; Sign test, Standardised Differences Test and Wilcoxon Sign Rank Test.

POPGENE Version 1.31 freeware (Yeh et al., 1999) was used to compute several measures of genetic variability within and between sample populations. The following variables were computed: Shannon's information index (I) (Lewontin, 1974), genetic distance and other population variance measurements such as observed heterozygosity (Ho) and expected heterozygosity (He).

Shannon' Information Index means for the four populations were compared at 95% level of significance using unpaired *t* test with Welch's correction performed by Graph Pad Prism version 3.00 (1999), for Windows, GraphPad Software, San Diego, California, USA, www.graphpad.com.

NTSYSpc version 2.11c (Rholf, 2001) used genetic distances matrices for the four populations generated by POPGENE Version 1.31 to construct a genetic relationships dendrogram from the Sequential Agglomerative Hierarchical and Nested (SAHN) clustering method using the Unweighted Pair-Group Method with Arithmetical averages (UPGMA) algorithm (Sneath and Sokal, 1973).

Mantel's test was undertaken to determine correlations between genetic and geographical distance matrices among the populations. The MXCOMP programme of NTSYS was used to compute a product-moment correlation coefficient (normalized mantel's statistics Z) for the two matrices (Rholf, 2001). In order to determine if the correlation was significant, actual coefficient was compared with the values produced by randomly permuting the matrix pair 2000 times.

The significance of the spatial variation in gene diversity among populations was estimated by performing a hierarchical analysis of genetic diversity among populations using the analysis of molecular variance model (AMOVA), as described in Michalakis and Excoffier (1996) using ARLEQUIN version 3.1 (Excoffier et al., 2006).

## RESULTS

### Genetic diversity

All 20 loci were polymorphic in all populations with polymorphic information content (PIC) values averaging 0.76 with a range of 0.70 at locus CypG22 to 0.82 at locus CypG15 (Table 1). This implies that all SSR loci used in this study had high discriminating power since PIC ≥ 0.5 is indicative of high polymorphism locus (Botstein et al., 1980). A composite genotypic linkage disequilibrium analysis for each locus pair across all four populations at 20 loci, showed that 19 (10%) locus pairs out of 190 possible locus pairs were in significant linkage disequilibrium (data not shown). This shows that genotypes at 90% of all loci were not linked. A total of 295 alleles with a mean of 15 per locus and a range of 10 to 22 alleles were generated reflecting a rich allelic diversity in the populations. A mean Shannon Information Index (I) of 1.64 was generated and showed that the populations exhibited high genetic diversity.

All populations deviated significantly from Hard-Weinberg Equilibrium (HWE) at all loci ($p \leq 0.05$) (Table 2). The mean observed heterozygosity ($H_o$) was higher than expected heterozygosity ($H_e$) averaging 0.98 and 0.87 and ranging from 0.92 to 1.00 and from 0.78 to 0.92, respectively (Table 2). The Wright's Fixation index ($F_{IS}$),

**Table 2.** Summary of estimates of observed and expected homozygosity and heterozygosity, Wright' F-Statistics and results of exact test of Hardy Weinberg Equilibrium (HWE) at 20 loci for all populations.

| Locus | HWE | $H_o$ | $H_e$ | Obs. Hom. | Exp. Hom. | $F_{IS}$ | $F_{IT}$ | $F_{ST}$ |
|---|---|---|---|---|---|---|---|---|
| CypG49 | 0.0000 | 0.98 | 0.86 | 0.02 | 0.14 | -0.32 | -0.09 | 0.18 |
| CypG3 | 0.0000 | 0.92 | 0.78 | 0.08 | 0.22 | -0.27 | -0.15 | 0.10 |
| CypG13 | 0.0000 | 1.00 | 0.81 | 0.01 | 0.19 | -0.35 | -0.18 | 0.12 |
| Ca3 | 0.0000 | 0.93 | 0.83 | 0.07 | 0.17 | -0.18 | -0.10 | 0.06 |
| CypG5 | 0.0000 | 0.97 | 0.79 | 0.03 | 0.21 | -0.28 | -0.22 | 0.05 |
| CypG4 | 0.0000 | 0.96 | 0.85 | 0.04 | 0.15 | -0.27 | -0.10 | 0.14 |
| CypG30 | 0.0004 | 0.96 | 0.90 | 0.04 | 0.10 | -0.32 | -0.01 | 0.23 |
| Lid1 | 0.0000 | 0.96 | 0.90 | 0.04 | 0.10 | -0.27 | -0.02 | 0.19 |
| CypG48 | 0.0000 | 0.97 | 0.91 | 0.03 | 0.09 | -0.23 | -0.03 | 0.16 |
| MFW | 0.0000 | 1.00 | 0.87 | 0.00 | 0.14 | -0.31 | -0.12 | 0.15 |
| CypG22 | 0.0000 | 1.00 | 0.86 | 0.00 | 0.14 | -0.44 | -0.09 | 0.24 |
| CypG6 | 0.0000 | 1.00 | 0.86 | 0.00 | 0.14 | -0.40 | -0.10 | 0.22 |
| CypG8 | 0.0000 | 1.00 | 0.90 | 0.00 | 0.10 | -0.37 | -0.05 | 0.23 |
| CypG21 | 0.0000 | 1.00 | 0.90 | 0.00 | 0.14 | -0.30 | -0.13 | 0.13 |
| CypG27 | 0.0000 | 1.00 | 0.90 | 0.00 | 0.14 | -0.31 | -0.12 | 0.14 |
| Lid11 | 0.0000 | 1.00 | 0.91 | 0.00 | 0.09 | -0.31 | -0.04 | 0.21 |
| AP1 | 0.0000 | 1.00 | 0.86 | 0.00 | 0.14 | -0.34 | -0.11 | 0.17 |
| AP2 | 0.0000 | 1.00 | 0.90 | 0.02 | 0.11 | -0.26 | -0.06 | 0.16 |
| Ru2 | 0.0000 | 1.00 | 0.90 | 0.01 | 0.10 | -0.32 | -0.05 | 0.20 |
| CypG15 | 0.0001 | 0.97 | 0.92 | 0.03 | 0.08 | -0.25 | -0.01 | 0.19 |
| Mean | 0.0000 | 0.98 | 0.87 | 0.02 | 0.13 | -0.30 | -0.09 | 0.17 |

$H_o$, Observed heterozygosity; $H_e$ expected heterozygosity; obs. Hom. observed homozygosity; Exp. Hom. expected homozygosity; Wright' F-Statistics ($F_{IS}$, $F_{IT}$ & $F_{ST}$)

values were negative across all loci in all populations indicating an excess of heterozygotes (heterozygotes excess), in contrast to the positive values of $F_{IS}$ which indicate excess of homozygotes (heterozygotes deficiency) (Table 2).

Bottleneck tests revealed that all the populations had a normal L-shaped allele frequency distribution as expected under mutation drift equilibrium, suggesting that the populations had not experienced a recent genetic bottleneck (data not shown).

## Genetic structure and differentiation among populations

Tests for a genic and genotype differentiation showed that all population pairs were significantly differentiated ($p \leq 0.05$) (Table 3). Similar results were reflected by mean Shannon Information Index which showed that all the population pairs were significantly different except LTP-DWA pair ($p \leq 0.05$) (Table 3). The estimate of population differentiation ($F_{ST}$) among the population pairs ranged from 0.14 to 0.19 with BWA-LTP pair being the least differentiated and LTP-DWA and NR-DWA pairs the most differentiated (Table 3). The overall differentiation among the populations was 17% ($F_{ST}$=0.17) (Table 2). The 17% level of differentiation was indicative of high genetic differentiation and was validated by Analysis of Molecular Variation (AMOVA) which showed that among population variation was at 16.4% (Table 4). Population differentiations of more than 15% are considered high rather than moderate (Wright, 1978; Hart and Clark 1997) and are associated with low gene flow among the populations. The study registered an overall low gene flow of less than one migrant per generation among the four populations (Nm=0.91). The highest number of migrants per generation (N*m* = 0.70) was observed between LTP and BWA populations and the lowest gene flow was between NR and DWA populations (Nm=0.38) (Table 3).

## Cluster analysis and genetic relationships among populations

Cluster analysis dendrogram indicated that BWA and LTP populations were the most genetically close and DWA population was genetically isolated from the rest of the populations (Figure 2). Although BWA and LTP populations were genetically close, but geographically, Bua River is close to Dwangwa River (Table 3). Mantel's test revealed weak insignificant positive correlation between genetic and geographical distances among the populations (r = 0.18; *p*=0.6369). Thus, the structuring in *O. microlepis* populations is not necessarily due to isolation

**Table 3.** Summary of contingency tests of differentiation (Fisher' exact Probability test), Nei's (1972) genetic distance, gene flow (Nm), Population Fixation Index ($F_{ST}$), geographical distance and Shannon Information Index (I) between population pairs.

| Population pair | Parameter | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Genic and genotypic Differentiation | Genetic distance | Nm | $F_{ST}$ | Geographical distance (km) | Shannon Information Index (I) | | |
| | *P -values* | | | | | Mean±SD | Mean±SD | *P-values* |
| NRand BWA | 0.0000 | 1.05 | 0.67 | 0.16 | 335.9 | 1.66±0.14 | 1.73±0.16 | 0.008*** |
| NR and LTP | 0.0000 | 0.82 | 0.49 | 0.16 | 451.9 | 1.66±0.14 | 1.58±0.15 | 0.006*** |
| BWAand LTP | 0.0000 | 0.74 | 0.70 | 0.14 | 122.11 | 1.73±0.16 | 1.58±0.15 | 0.0001*** |
| NR and DWA | 0.0000 | 1.24 | 0.38 | 0.19 | 294.24 | 1.66±0.14 | 1.58±0.13 | 0.0001*** |
| BWAand DWA | 0.0000 | 0.90 | 0.56 | 0.16 | 43.98 | 1.73±0.16 | 1.58±0.13 | 0.0001*** |
| LTP and DWA | 0.0000 | 1.06 | 0.45 | 0.19 | 158.85 | 1.58±0.15 | 1.58±0.13 | 0.8201ns |

NR, North Rukuru River; BWA, Bua River; DWA, Dwangwa River; LTP; Linthipe River.

**Table 4.** Analysis of molecular variation (AMOVA) among and within population of *O. microlepis* (based on 1023 permutation and over 20000 bootstraps).

| Source of variation | Sum of squares | Variance components | Percentage Variation |
|---|---|---|---|
| Among Population | 458.83 | 1.48 | 16.40 |
| Within Population | 2931.2 | 7.54 | 83.60 |
| Total | 3390.02 | 9.02 | - |

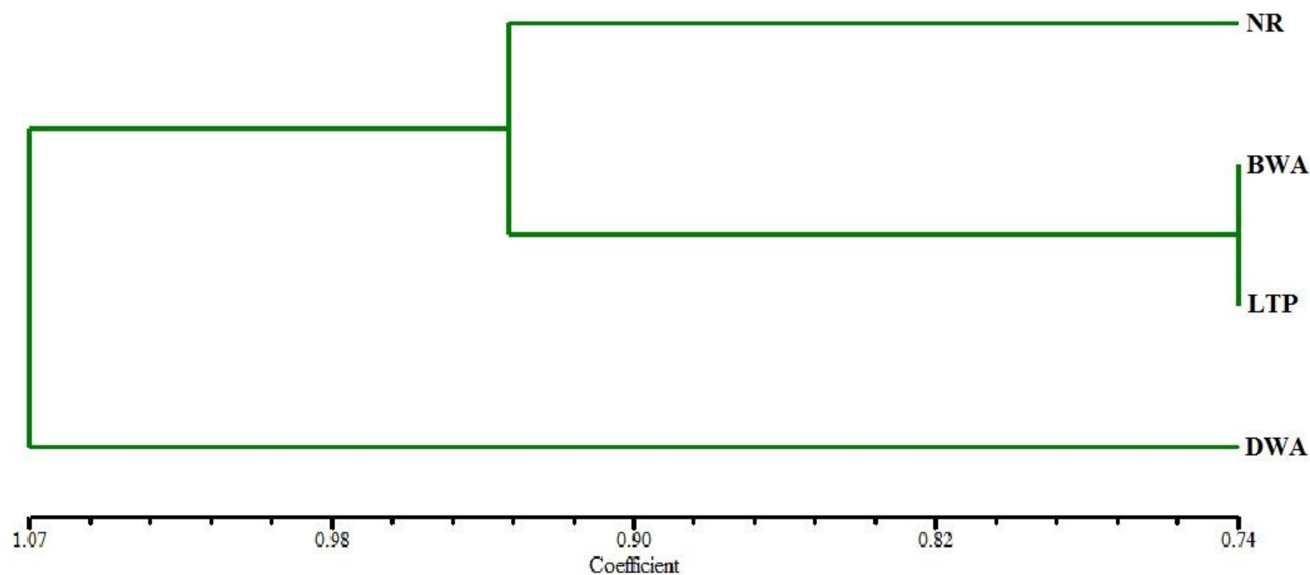


**Figure 2.** Genetic relationships among the populations based on Nei' genetic distances.

by geographical distances.

## DISCUSSION

### Genetic diversity

Maintenance of genetic diversity is a major component of many species conservation programs, since loss or critical reduction of genetic diversity is one indicator of diminution in evolutionary potential. In this study, Mpasa, an endangered species due to declining numbers caused by overfishing, demonstrated high degree of genetic diversity, based on total allele count of 295 resulting in a high average number of alleles per locus of 15. This observation was further supported by high mean observed heterozygosity (0.98) and high mean Shannon Information Index (1.64). The mean values of number of alleles per locus and observed heterozygosity reported in this

study are higher than those found in other cyprinid species that have experienced a reduction in historic population size but still retain high genetic diversity, such as the Cape Fear shiner (A = 8.2, $H_e$ = 0.70) reported by Saillant et al. (2004); *Anaecypris hispanica* (A = 10.3, $H_e$ = 0.68) reported by Salgueiro et al. (2003) and the critically endangered cyprinid, *Squalius aradensis* (A=4.9, $H_e$=0.45) reported by Mesquita et al. (2005). The observed phenomenon (high genetic diversity after presumed reduction in population size) can generally occur when a species has recently experienced a modest decline in population size which has not been sustained over many generations. Continued reductions of even modest proportions, can lead to serious decrease in genetic diversity (Frankham et al., 1995) so continuous monitoring is recommended. However, bottleneck tests results in this study cast some doubts on the presumed reduction in population size among the four Mpasa populations, since they reveal that the populations have not experienced recent reduction in effective population size. Therefore, the observed high genetic diversity could be due to out breeding within the populations.

## Conformity to Hardy-Weinberg Equilibrium and test for linkage disequilibrium

Determination of whether a fish stock is a mixture of fish from more than one population is commonly done using a test of Hardy-Weinberg Equilibrium and a test for linkage disequilibrium. Both tests are based on the genetics principle that a mixture of gene pools will exhibit a Wahlund effect, that is, homozygote excess relative to binomial distribution (Kamonrat, 1996). All the populations of *O. microlepis* were not in Hardy-Weinberg Equilibrium at all loci and had negative values for $F_{IS}$ (Table 2) at all loci suggesting that an excess of heterozygotes was responsible for the departure from HWE (Rosewich et al., 1999). The heterozygosity excess could be due to out breeding of the populations since contrary to popular expectations; bottleneck tests have shown that the *O. microlepis* populations in the study have not experienced a recent genetic bottleneck despite known over exploitation over the years. However, Peery et al. (2012) argued through a review of published literature that typically applied, microsatellite-based bottleneck tests often do not detect bottlenecks in populations of vertebrates known to have experienced declines. Their simulations revealed that bottleneck tests displayed limited statistical power to detect bottlenecks due to limited samples sizes and loci used in reviewed published studies (median=8-9 loci and 31-38 individuals). If bottleneck tests limitations are applied to this study, then excess of heterozygotes, which imply occurrence of recent genetic bottlenecks, probably would be due to small number of breeders producing the next generation leading to allelic frequencies in males and females parents differing due to binomial sampling error (Pudovkin et al., 1996; Stoeckel et al., 2006). This difference in allele frequencies in males and females results in an excess as regards HWE of heterozygotes in progeny hence populations deviating from HWE (Pudovkin et al., 1996). Further, Balloux (2004) argued that in small sexual or self-incompatible populations, the fact that individuals cannot reproduce with themselves decrease the probability of creating homozygote offspring.

The study registered 10% of loci pair wise linkage disequilibrium at 5% level indicating a possible admixture though at a lower level.

## Genetic structure and differentiation within and among population

AMOVA results revealed that 84% of genetic variation resides within populations indicating possibly, the existence of large enough populations that allow reasonable random mating. The 16% among population genetic variance implies great genetic differentiation among the populations as validated by $F_{ST}$ value of 0.17. Wright (1978) considered any $F_{ST}$ value above 0.15 as showing high genetic differentiation. Level of genetic differentiation demonstrated by Mpasa in this study (Table 2) is higher as compared to values seen in Pacific herring ($F_{ST}$ = 0.023), Atlantic herring ($F_{ST}$ = 0.035) and widespread anadromous fish like Atlantic salmon ($F_{ST}$ = 0.054) (McConnell et al., 1995). Low gene flow among populations (Nm=0.91) and absence of recent genetic bottlenecks which implies that despite heavy exploitation, the populations are outbreed due to possible presence of large numbers of fish, are the plausible reasons for the observed high genetic differentiation among the populations. The high $F_{ST}$ obtained in this study hence signify that the populations are distinct requiring independent conservation management for each river system. Mills and Allendorf (1996) concluded that rate of migration of $Nm \geq 1$ leads to considerable homogeneity among populations but population divergence and structuring occurs when $Nm \leq 1$. The populations in the present study have overall migration rate of $Nm \leq 1$ rendering them to structuring and divergence.

## Genetic relationships and cluster analysis of the populations

Although comparison of Shannon Information Index population means found that Linthipe and Dwangwa population pair was insignificantly different ($p$=0.8201), Wright F statistics ($F_{ST}$ = 0.19) revealed that the pair was among the most differentiated with second lowest gene flow (Table 3). Bua and Linthipe population pair though not the closest geographically is the most genetically close (Figure 2) possibly due to common founding population which is reflected in form of more shared alleles (highest gene flow) among the population pairs resulting in the least genetic differentiation (Table 3).

The low relationship between genetic and geographical distance as revealed by Mantel's test denotes that the populations do not fit into the isolation by distance model. This model states that gene flow is the highest between close populations and it is expected that close populations should show similar genetic composition but it is not the case with these populations. Findings of this study concur with other studies of *Lenthrinops* species flock (Duponchelle et al., 1999; Changadeya et al., 2001) which reported fish flocks not fitting the isolation by distance model though in those studies the fish populations experienced high migration rates.

## Conclusions

The results showed that populations in the study are probably still large, have high genetic variation and are highly differentiated as to be considered as distinct populations. Therefore, the four inlet river populations (North Rukuru, Bua, Linthipe and Dwangwa) require separate monitoring and management strategies due to inferred restricted gene flow and population differentiation.

## ACKNOWLEDGEMENTS

The research was funded and supported by the Regional Fish Node NEPAD/SANBio at Bunda College of Agriculture, of the Lilongwe University of Agriculture and Natural Resources. We would like to acknowledge Y. Kazembe and E. Acquaron, technicians in the MBERU DNA laboratory of Department of Biological Sciences, Chancellor College, for all technical assistance provided during sample collection in the field and data generation in the laboratory.

### REFERENCES

Alam MS, Islam MS (2005). Population genetic structure of Catla catla (Hamilton) revealed by microsatellite DNA markers. Aquaculture 246: 151-160.

Ambali AJD, Malekano LB, Mkoko BJ, Mwale GP, Changadeya W (2002). Assessment of impacts and potential mitigative interventions of fishing and other anthropogenic activities on the 100-metre aquatic zone of Lake Malawi. Report submitted to the Department of National Parks and Wildlife, Lilongwe, Malawi.

Anon (2008). Annual Report. Division of Fisheries, Ministry of Livestock Development and Fisheries, Dar es Salaam, Tanzania.

Baerwald MR, May B (2004). Characterisation of microsatellite loci for five members of the minnow family Cyprinidae found in the Sacramento- San Joaquin Delta and Its tributaries. Mol. Ecol. Notes 4: 385-390.

Balloux F (2004). Heterozygosity excess in small populations and the heterozygosity-excess effective population size. Evolution 58: 1891-1900.

Botstein D, White RL, Skolnick M, Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32: 314-331.

Buhariwalla HK, Jarret RL, Jayashree B, Crouch JH, Ortiz R (2005). Isolation and characterization of microsatellite markers from *Musa balbisiana*. Mol. Ecol. Notes 5:327–330.

Changadeya W, Ambali, AJD, Malekano L (2001). Genetic variation of *Taeniolenthrinops praeorbitalis* (Chisawasawa) in the central and southern Lake Malawi. UNISWA J. Agric. 10: 30-39.

Changadeya W, Kaunda E, Ambali A (2012). Molecular characterization of *Musa* L. cultivars cultivated in Malawi using microsatellite markers. Afr. J. Biotechnol. 11 (18): 4140-4157.

Collevatti RG, Grattapaglia D, Hay JD (2001). Population genetic structure of the endangered tropical tree species Caryocar brasiliense, based on variability at microsatellite loci. Mol. Ecol. 10:349-356.

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

Creste S, Neto AT, Vencovsky R, Silva SO, Figueira A (2004). Genetic diversity of *Musa* diploid and triploid accessions from the Brazilian banana breeding program estimated by microsatellite markers. Genet. Resour. Crop Evol. 51: 723-733, Kluwer Academic Publishers.

Duponchelle F, Snoeks J, Agnese JF, Hanssens H, Mafuca J, Mandere D, Msukwa A (1999). Comparison of taxonomic, genetic and life history characteristics of two species (*Mylochromis anaphyrmus* and *Lenthrinops gossei*) in various parts of lake Malawi/Nyasa. In Ribbink, A.J. and Ribbink, A.C. (eds). Abstracts of the SADC/GEF Lake Malawi/Nyasa Biodiversity Conservation Project Final Workshop,4-3/03/1999, Senga bay, Salima, Malawi: 34-35.

Excoffier L, Laval G, Schneider S (2006). ARLEQUIN ver. 3.1 An Integrated Software package for Population Genetics Data Analysis, University of Berne, Switzerland.

Frankham R (1995). Effective population size/adult population size ratios in wildlife: review. Genet. Res. 66: 95-107.

Government of Malawi (1998). National Environmental Action Plan. 1.0 Department of Research and Environmental Affairs, Lilongwe, Malawi.

GraphPad Prism Version 3.00 (1999). GraphPad Software Incorporated, San Diego, California, USA, www.graphpad.com

Gunther (1864). Lake salmon genus Opsaridium, *Opsaridium microlepis*, Available on The free encyclopedia wikipedia, [Accessed on 29/10/2011].

Guo SW, Thompson EA (1992). Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. Biometrics 48:361-372.

Haldane JBS (1954). An exact test for randomness of mating. J. Genet. 52: 631-635.

Hamilton PB, Tyler CR (2008). Identification of microsatellite loci for parentage analysis in roach *Rutilus rutilus* and eight other cyprinid fish by cross-species amplification, and a novel test for detectimng hybrids between roach and other cyprinids. Mol. Ecol. Resour. 8: 462-465.

Hart DL, Clark AG (1997). Principles of Population Genetics, 3rd Edition. Sinauer Associates, Inc, Sunderlands, Massachusetts, p. 542.

Hillis DM, Mable BK, Larson A, Davis SK, Zimmer EA (1996). Nucleic Acids IV: Sequencing and Cloning. In; Hillis, D.M., Moritz, C., & Mable, B.K. Molecular Systematics, 2nd ed. Massachusetts Sinauer Associates Inc.

Hsu KC, Wang JP, Chen XL, Chiang TY (2004). Isolation and characterization of microsatellite loci in *Acrossocheilus paradoxus* (Cyprinidae) using PCR-based isolation of microsatellite arrays (PIMA). Genetics 5:113-115.

IUCN (2006). The Threatened IUCN Red List of Threatened species. www.iucnredlist.org. Accessed on 7 December 2006.

Johnson MS (2001). Measuring and interpreting genetic structure to minimise the genetic risks of translocations. Aquac. Res. 31: 133-143.

Kamonrat W (1996). Spartial genetic structure of Thai Silver Barb *Puntius gonionotus* (Bleeker) populations in Thailand. PhD. Thesis, Dalhousie University, Canada.

Kazembe J, Magombo Z, Khawa D, Kaunda E (2005). *Opsaridium microlepis*. In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.1. www.iucnredlist.org. html.[accessed on 26/06/2012].

Lewontin RC (1974). The genetic basis of evolutionary change. Columbia University Press, New York.

Lowe RH (1975). Fish communities in tropical freshwaters. London,

Longman. p. 337.

McConnell SK, O'Reilly P, Hamilton L, Wright JM, Bentzen P (1995). Polymorphic microsatellite markers loci from Atlantic salmon (*Salmo salar*): genetic differentiation of North American and European populations. Can. J. Fish Aquat. Sci. 52: 1862-1873.

Mesquita N, Hanfling B, Carvalho GR, Coelho MM (2005). Phylogeography of the cyprinid Squalius aradensis and implications for conservation of the endemic freshwater fauna of southern Portugal. Mol. Ecol. 14:1939–1954

Michalakis Y, Excoffier L (1996). A generic estimation of population sub division using distances between alleles with special reference for microsatellite loci. Genetics 142: 1061–1064.

Mills LS, Allendorf FW (1996). The one-migrant-per-generation rule in conservation and management. Conserv. Biol. 10 (6): 1509 -1518.

Mohindra V, Anshumala PP, Narain L, Kapoor D, Lal KK (2005). Microsatellite loci to determine population structure of *Labeo dero* (Cyprinidae). Aquat. Living Resour. 18: 83-85.

Nei M (1972). Genetic distance between populations. Am. Nat. 106: 283-292.

Peery MZ, Kirby R, Reid BN, Stoelting R, Doucet-Bëer E, Robinson S, Vasquez-arrillo C, Pauli JN, Palsbøll PJ (2012). Reliability of genetic bottleneck tests for detecting recent population declines. Mol. Ecol. 21: 3403-3418.

Pudovkin AI, Zaykin DV, Hedgecock D (1996). On the potential for estimating the effective number of breeders from heterozygotes excess in progeny. Genetics 144:383-387.

Raymond M, Rousset F (1995).GENEPOP: a population genetics software for exact test sand ecumenicism. J. Hered. 86: 248–249.

Rholf JR (2001). NTSYSpc Version 2.11c Numerical Taxonomy and Multivariate Analysis System, Exeter Software, New York.

Romana-Eguia MR, Ikeda M, Basiao Z, Taniguchi N (2004). Genetic diversity in farmed Asian Nile and red hybrid tilapia stocks evaluated from microsatellite and mitochondria DNA analysis. Aquaculture 236: 131-150.

Rosewich UL, Pettway RE, Mcdonald BA, Kistler HC (1999). High levels of geneflow and heterozygosity excess characterize *Rhizoctonia solani* AG-1IA (*Thanatephorus cucumeris*) from Texas. Fungal Genet. Biol. 28: 148-159.

Saillant E, Patton JC, Ross KE, Gold JR (2004). Conservation genetics and demographic history of the endangered *Notropis mekistocholas*. Mol. Ecol. 13: 2947-2958.

Salgueiro P, Carvalho GR, Collares-Pereira MJ, Coelho MM (2003). Microsatellite analysis of genetic population structure of the endangered cyprinid *Anaecypris hispanica* in Portugal: Implications for conservation. Biol. Conserv. 109:47-56.

Slatkin M (1985). Rare alleles as indicators of gene flow. Evolution 39 (1): 53-65.

Sneath PHA, Sokal RR (1973). Numerical taxonomy. Freeman, San Francisco. p. 573.

Snoeks J (ed) (2004). The cichlids diversity of Lake Malawi/Nyasa/ Niassa: identification, distribution and taxonomy. Cichlid Press, El Paso, USA, p. 360.

Stoeckel S, Grande J, Fernandez-Manjarres JF, Bilger I, Frascaria-Lacoste N, Mariette S (2006). Heterozygosity excess in a self-incompatible and partially clonal forest trees species-*Pinus avium* L. Mol. Ecol. 15: 2109-2118.

Tweddle D (1981). The importance of long-term data collection on river fisheries, with particular reference to the cyprinid *Opsaridium microlepis* (Gunther 1864) fisheries of the affluent rivers of Lake Malawi. In: Seminar on river basin management and development. CIFA/T, 8: 145-163.

Weir BS (1990). Genetic data Analysis. Sinauer Associates, Sunderland, MA.

Wright S (1969). Evolution and genetics of populations. University of Chicago Press, Chicago, Illinois.

Wright S (1978). Evolution and the Genetics of Populations, Variability Within and Among Natural Populations. The University of Chicago Press, Chicago.

Yeh FC, Yang R, Boyle T (1999). POPGENE VERSION 1.31, Microsoft Window-based Freeware for Population Genetics Analysis. Quick User Guide. Univeristy of Alberta and Centre for International Forestry Research.

Zhao W, Wang Y, Chen T, Jia G, Wang X, Qi J, Pang Y, Wang S, Li Z, Huang Y, Pan Y, Yang, YH (2007). Genetic structure of mulberry from different ecotypes revealed by ISSRs in China: An implication for conservation of local mulberry varieties. Sci. Hortic. 115: 47-55.