

Genetic Diversity and Population Structure of *Heterobranchus bidorsalis* Inferred from Cytochrome b gene Sequences

*ABUBAKAR, M.Y.¹, IPINJOLU, J.K.¹, ESA, Y.B.³, MAGAWATA, I.¹ AND HASSAN, W.A.²
Department of Fisheries and Aquaculture, Faculty of Agriculture, Danfodiyo University Sokoto.
Department of Animal Science, Faculty of Agriculture, Danfodiyo University Sokoto.
³Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia.
*Corresponding Author
E-mail: yahaya.abubakar@udusok.edu.ng

Abstract

Mitochondria DNA (mtDNA) diversity *Heterobranchus bidorsalis* from Rivers Niger and Benue were studied. The mtDNA Cytochrome-b gene was inferred using 1118bp partial sequence of 27 and 8 samples from Rivers Niger and Benue Respectively. Low Guanine (13%) and higher cytosine (31%) composition were revealed. Ten haplotypes(MG334177 to MG334186)were distinguished. River Niger samples had seven unique haplotypes while two haplotypes were shared among the two populations. Low nucleotide (0.0008 – 0.001) and high (0.67 – 0.68) haplotype diversity were recorded. F_{ST} 's (-0.02966 to 0.000) was non-significant ($p > 0.05$) among the populations. AMOVA(- 2.97%)shows no genetic structure between the populations. *Tajima's D* and *Fu's FS* neutrality tests suggested recent population growth for the populations. Phylogenetic reconstruction revealed monophyletic relationship among the populations. Analysis of data demonstrated suitability of partial Cytochrome b sequence in determining the genetic diversity in *H. bidorsalis* and provides new information for the species' conservation and exploitation.

Keywords: *Heterobranchus bidorsalis*, mtDNA cytochrome b, Genetic diversity, population structure.

Introduction

Aquaculture of catfish of the clariidae family remains the largest contributor and well developed facet of aquaculture in Nigeria. According to Adewumi and Olaleye (2011), the story of aquaculture in Nigeria is essentially that of catfish culture and the favoured species of the clariidae family are those of the genus *Heterobranchus* and *Clarias*. One of these species is *Heterobranchus*

bidorsalis. To ensure sustainability in this species production, continuous controlled reproduction is an indispensable prerequisite for domestication (Bilio, 2007). Accordingly, domestication allows production increases far beyond the maximum sustainable yields of the capture fisheries. However, such increases can seriously threaten con-specific wild populations through escapes from aquaculture installations. However, the genetic intermingling of farmed and wild fish pose great problems to species

diversity and integrity that must be overcome, or at least substantially reduced (Bilio, 2007). To be successful in acquiring high quality seeds, proper identification of the candidate species and the particular strain becomes very important. Recent observations on the field showed problems of improper identification and indiscriminate hybridization of the clariid species, which would affect the genetic make-up of the species (Aluko, 1998; Yisa and Olufeagba, 2005). Despite the large distribution of *H. bidorsalis* in Northern Nigerian water bodies and its aquaculture significance, little is known about the genetic diversity.

Advancement in molecular techniques has increased the availability of different DNA-based markers, which have become efficient tool in conservation genetic studies and had been successfully applied in fisheries sector to identify closely related fish species, determination of genetic diversity of fish population and species identification (Ponzoni *et al.*, 2010; Avise, 2004). Among of which is the mitochondrial DNA (Abdul Azeez *et al.*, 2015). The use of mitochondrial DNA (mtDNA) markers offer several advantages such as maternal mode of inheritance and relatively limited recombination that make it effective in conducting studies regarding phylogenetics analysis (Hurst *et al.*, 1999). According to Habib *et al.* (2011), the fast rate of mtDNA evolution coupled with maternal inheritance have made mtDNA an extremely useful genetic system for studying gene flow, hybrid zones, population structure and other population related questions. Even conservative protein coding genes like Cytochrome b tend to show intraspecific variation mainly in 3rd position of codon which can be used to identify stocks. The molecular characterization of the wild population of this *H. bidorsalis* then becomes very important so as to provide background for effective conservation of the original attributes and the level

of genetic diversity which serve as a base for sustainable management and aquaculture improvement, before the wild population's pedigree are seriously and negatively altered.

Materials and Methods

Fish samples and Sampling sites

A total of 35 samples of *H. bidorsalis* were collected from commercial catches of River Niger (27) and Benue (8), from July to September 2015 (Table 1). Samples were identified using taxonomic keys provided in Reed *et al.* (1967). 2cm³ caudal fin clip of each sample was collected and preserved in 2ml Eppendorf tube filled with Aldrich ethyl absolute alcohol, pure 200 proof, for molecular biology (Dingley *et al.*, 2005) and stored in -20°C and then transported in an air tight ice filled (GINT[®] TITAN 220) vacuum flask via airplane for 48 hours to Fish Genetics Laboratory, Department of Aquaculture, Universiti Putra, Malaysia, where the samples were stored in -20°C for further analysis.

DNA extraction, Polymerase Chain reaction (PCR) Amplification, Agarose Gel Electrophoresis and sequencing

Total genomic DNA extraction was performed using the Wizard[®] SV Genomic DNA Purification System (Promega, Madison, WI, USA), according to manufacturers' protocol by using approximately 20mg of caudal fin clip properly homogenized. The DNA quality and approximate yield were determined by electrophoresis in a 1% agarose gel containing GelRed[™] Nucleic Acid Gel Stain (Biotium) solution at 90V for 30min. the isolated genomic DNA was used for genetic analysis. The presence of band viewed in a gel imager indicates the presence of DNA (Plate 1). The DNA templates were then stored in -20°C for further genetic analysis.

Table 1: Collection sites for fish samples used for genetic characterization

Species of fish	No. of Samples	Collection Sites of Samples	Geographical Coordinates
<i>Heterobranchus bidorsalis</i>	8	Lower River Benue	Lower River Benue at Longitude 7° 47' 0" E and Latitude 8° 0' N.
<i>H. bidorsalis</i>	27	Kainji Lake on River Niger	River Niger at Longitude. 4° 20' 0" E and Latitude 10° 5' 0" N

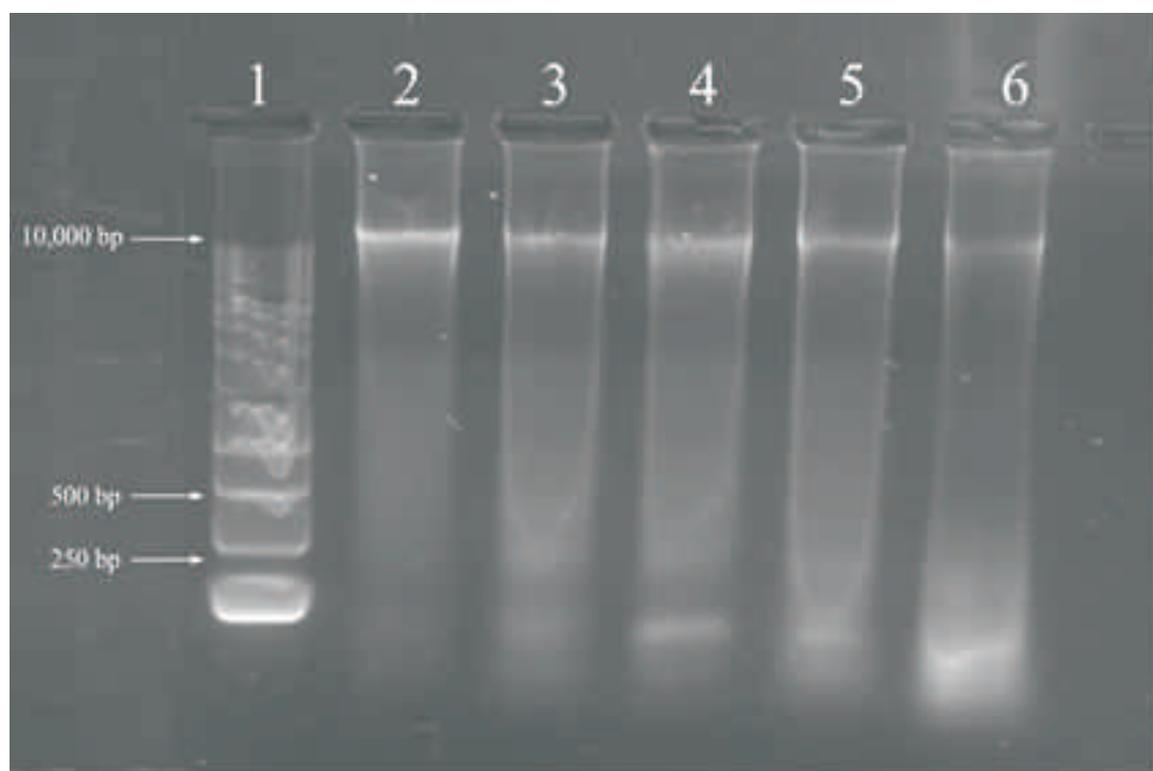


Plate 1: DNA Bands visualized under UV light

To investigate genetic diversity among the population, oligonucleotide primer was used to amplify Cytochrome b (Cyt-b) region of the mitochondria DNA template of each sample extract. 1158bp partial sequence of mitochondria Cyt-b gene were amplified with primers 5'-GACTTGAAAAACCACCGTTG-3' and 5'-CTCCGATCTCCGGATTACAAGAC-3'. (Xiao *et al.*, 2001) using polymerase chain reactions (PCR). The PCR optimization procedures were conducted in an overall volume of 25 μ l which contained 0.5 μ l of each deoxynucleotide triphosphate (dNTPs), 0.2 μ l Taq polymerase, 2.5 μ l MgCl₂, 0.5 μ l of each primers, and 2 μ l of concentrated genomic DNA, 5 μ l of Taq buffer and 13.8 μ l of distilled H₂O (Xiao *et al.*, 2001). The PCR thermal cycling was carried out using an Eppendorf Master cycler based on the following thermal regime: 5 min of 95°C initial denaturation

step; 40 cycles of 94°C denaturation step for 30s, a 50°C annealing temperature for 30 s and a 72°C extension period of 2 min; followed by 72°C final extension step for 10 min and a routine 4°C final hold (Xiao *et al.*, 2001) with slight modification. In order to confirm that the PCR reaction generated sufficient amplicon proportions, PCR amplification products were visualized using a 2.0% laboratory grade agarose gel containing 5 ml GelRed stain a digested lambda DNA ladder (GeneRuler™ 100-bp DNA ladder) was used as marker. The gels were run in an electrophoresis medium for 45min. at 75V and viewed on gel imager (Plate 2). Purified DNA samples were then sent to private sector institution (1st Base 129 laboratories Sdn Bhd) for sequencing to generate associated trace files and continuous read 130 lengths intended for genetic and statistical analysis of mitochondrial DNA.

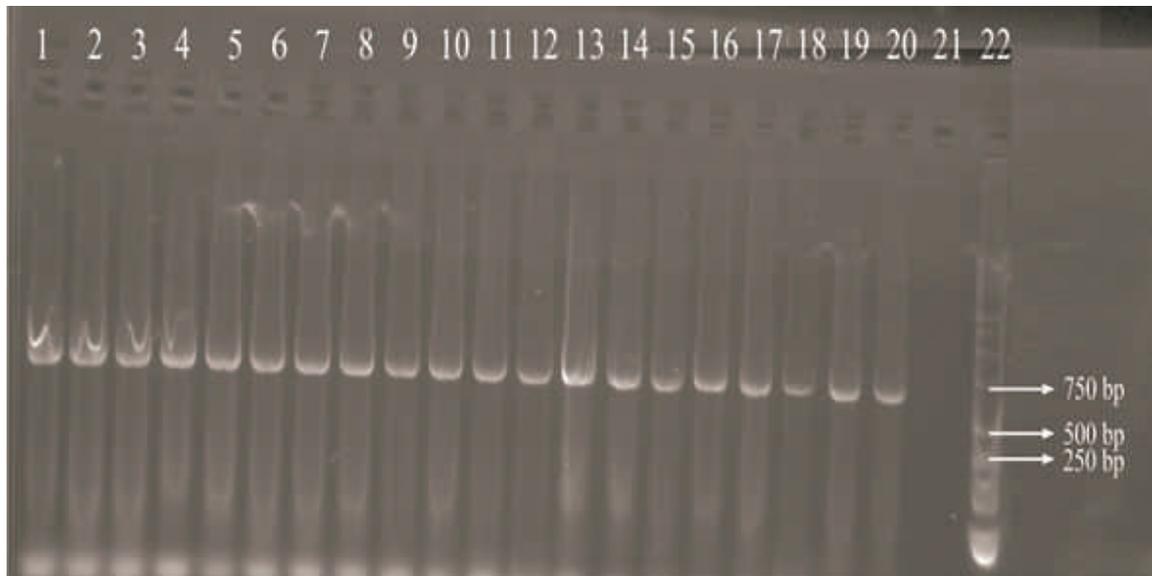


Plate 2: PCR amplified DNA products bands visualized under UV light. Lane 1 -20 = PCR product band of 20 fish amplified DNA templates, Lane 21 = Negative control, Lane 22 = 100 bp ladder

Sequence Analysis

Sequence data were subjected to nucleotide mega blast (high similarity) search (<http://www.ncbi.nlm.nih.gov/BLAST/>) in order to identify the best taxon match for the resultant sequence for each specimen to ensure the sequence were from *Heterobranchus bidorsalis*.

Trace files were manually end-trimmed using BioEdit software 7.2.5 (Hall, 1999) regarding to their homologous section. Afterwards, ClustalX 2.1 (Thompson *et al.*, 1997) was applied to progressively manipulate, align and analyze the DNA sequences. Finally, haplotypes were detected with DnaSP software 5.10.01 (Librado & Rozas, 2009) and deposited into NCBI GenBank.

Sequence diversity and phylogenetic analysis

FaBox programme version 1.41 was used online to group sequence into haplotypes. MEGA 7 (Kumar *et al.*, 2015) was used to perform amino acid translation for the sequences, detect if exist stop codon in the aligned sequences and also used to carry out phylogenetic and molecular evolutionary analysis. Calculation of the pairwise distance was obtained through 1,000 bootstrap variance estimation and Jukes and contour model (Jukes and Contour, 1969). Phylogenetic trees among haplotypes were done using Neighbour-joining (NJ) tree based on the

Tamura 3-parameter (Tamura and Nei 1993), based on the best fit Hasegawa-Kishino-Yano (HKY) model (Nei & Kumar, 2000) and The robustness of statistical support for branches on the NJ trees was determined by 1000 bootstrap replicates (Felsenstein, 1985). *Clarias microcephalus* sequence with GenBank accession number: AJ548462 was used as outgroup. Haplotypes (h), haplotype diversity and nucleotide diversity (π) (Nei, 1987) were calculated for each population using the software Arlequin programme V3.5 (Excoffier and Lischer, 2011). Tajima's D test (Tajima, 1989) and Fu's F_s (Fu, 1997) neutrality test was done. Also Calculations of pairwise F_{st} value (Slatkin 1995), genetic distance analyses were performed. Partitioning of genetic variation within and among populations was calculated using analysis of molecular variance (AMOVA), by computation of conventional F-statistics from haplotypes with 1000 permutations using the software Arlequin programme V3.5 (Excoffier and Lischer, 2011).

Results

The *H. bidorsalis* sequences which were queried against the GenBank database showed low percentage match below 90% with *Heterobranchus* spp. lodged earlier by different authors. The sequence composition analyzed to check for

genetic variation revealed no evidence for pseudo-gene amplification (no ambiguous alignment, low sequence quality, double peaks or stop codons) was found resulting in 1158bp consensus sequence after alignment for sample of the two population. The mean frequency of nucleotide composition revealed low Guanine (G) =13.01 composition, followed by thiamine (T)=27and adenine (A)= 8.98 while cytosine (C) =31.01 had the highest sequence composition. The variable/polymorphic sites observed were ten (0.86%) including five (5) parsimoniously informative sites, while 1148 sites (99.14%) were conserved. Ten haplotypes were distinguished and submitted on NCBI Genbank with assigned accession numbers (MG334177 -MG334186). Samples from River Niger had seven (7) unique haplotypes while River Benue had (1).

Meanwhile two (2) haplotypes were shared among the two populations among of which haplotype 3 (MG334179) had the highest relative frequency (0.556 and 0.5) in both population. One of the haplotypes (Hap 8) was rare and from River Benue samples (Table 2). However, number of polymorphic sites found among the haplotypes were nine (9), including seven (7) transitions and two (2) trans-version sites. The genetic diversity indices presented in Table 2, revealed Low nucleotide and high haplotype diversity. Highest nucleotide diversity (0.001363 ±0.000974) was in River Niger population while the haplotype diversity were similar in both River Niger (0.6809±0.0935) and River Benue (0.6786±0.1220) population. The number of pairwise difference was higher in River Niger population (1.327635±0.851891).

Table 2: Summary of variable sites distribution and relative frequency of haplotype among *H. bidorsalis* populations, based on sequences of the mitochondrial DNA cytochrome b gene

Haplotypes	Variable Sites										GenBank Accession Numbers	Population/Number of sequence		
	4	8	8	9	9	9	9	9	9	9		River (n=27):(D)	Niger	RiverBenue (n=8):(D)
Hap 1	A	T	T	C	T	C	A	G	G	G	MG334177	0.074 (2)		0.000 (0)
Hap 2	G		A			T		A		A	MG334178	0.037 (1)		0.000 (0)
Hap 3								A		A	MG334179	0.556 (15)		0.5 (4)
Hap 4								G	A		MG334180	0.148 (4)		0.375 (0)
Hap 5	G								A		MG334181	0.037 (1)		0.000 (0)
Hap 6				T	C	T			A	A	MG334182	0.037 (1)		0.000 (0)
Hap 7									A		MG334183	0.037 (1)		0.000 (0)
Hap 8							G	A	A	A	MG334184	0.000 (0)		0.125 (1)
Hap 9		A							A		MG334185	0.037 (1)		0.000 (0)
Hap 10								A	A		MG334186	0.037 (1)		0.000 (0)
Total												1.00		1.00
Nucleotide diversity (Pi _{Jc})												0.001±0.0009		0.0008 ± 0.0007
No. of haplotypes												9		3
Haplotype Diversity												0.68±0.0935		0.67±0.1220
No. of Pairwise Difference												1.328±0.852		0.785 ±0.6333
No of Polymorphic sites												9		2
No of Observed trans versions												2		0
No of Observed transitions												7		2

n = sample size, (D) = Distribution of samples in each haplotype. C = cytosine, T = thiamine, A = adenine, G = guanine

The genetic distance, molecular diversity indices and neutrality tests revealed that the location based Population Pairwise F_{ST} 's shown in Table 3, indicates non-significance ($p > 0.05$) genetic variation among the two population with a genetic distance of (-0.02966) and divergence between the two populations is (-0.02966 to 0.000). The analysis of molecular variance showed no variation (-2.96628) among the two population (Table 4) but with high variation (102.96628) within each population. Tajima's D and Fu's F_s neutrality of the sequence data among the two population were negative and

non-significant ($p > 0.05$) with mean values (-0.65226±1.02052 and -2.22636±2.83232) respectively as depicted in Table 5.

The evolutionary relationship among the ten haplotypes obtained from the two population constructed with neighbor joining (NJ) tree where *C. microcephalus* was used as outgroup is presented in Fig. 1. It shows very low and shallow genealogical divergence with clusters among the haplotypes of the two population. The NJ tree depict that haplotype 8 represents the ancestral allelic group where other haplotypes evolved.

Table 3: Location-based population Pairwise F_{ST} s of *H. bidorsalis* populations based on cytochrome b gene

	River Niger	River Benue
River Niger	0.000	
River Benue	-0.02966	0.000

$p > 0.05$

Table 4: Analysis of molecular variation (AMOVA) in *Cyt-b* gene testing the genetic structure among the two *H. bidorsalis* populations

Source of variation	Sum of squares	Variance components	Percentage variation
Among population	0.391	-0.01747	-2.96628
Within populations	20.009	0.60634	102.96628
Total	20.400	0.58887	

Table 5: Neutrality tests for *H. bidorsalis* populations based on cytochrome b (*Cyt-b*) gene

Statistics	River Niger	River Benue	Mean	S.D
Tajima's D test				
Sample size	27	8	17.50000	13.43503
S	9	2	5.50000	4.94975
Pi	1.32764	0.78571	1.05667	0.38320
Tajima's D	-1.37388	-0.06935	-0.65226	1.02052
Tajima's D p-value	0.08600	0.64100	0.36350	0.39244
Fu's FS test				
No. of Alleles unchecked	9	3	6.00000	4.24264
Theta pi	1.32764	0.78571	1.05667	0.38320
Exp. No. of alleles	4.59943	2.44629	3.52286	1.52250
FS	-4.22911	-0.22360	-2.22636	2.83232
FS p-value	0.00300	0.22600	0.11450	0.15768

Neighbour joining

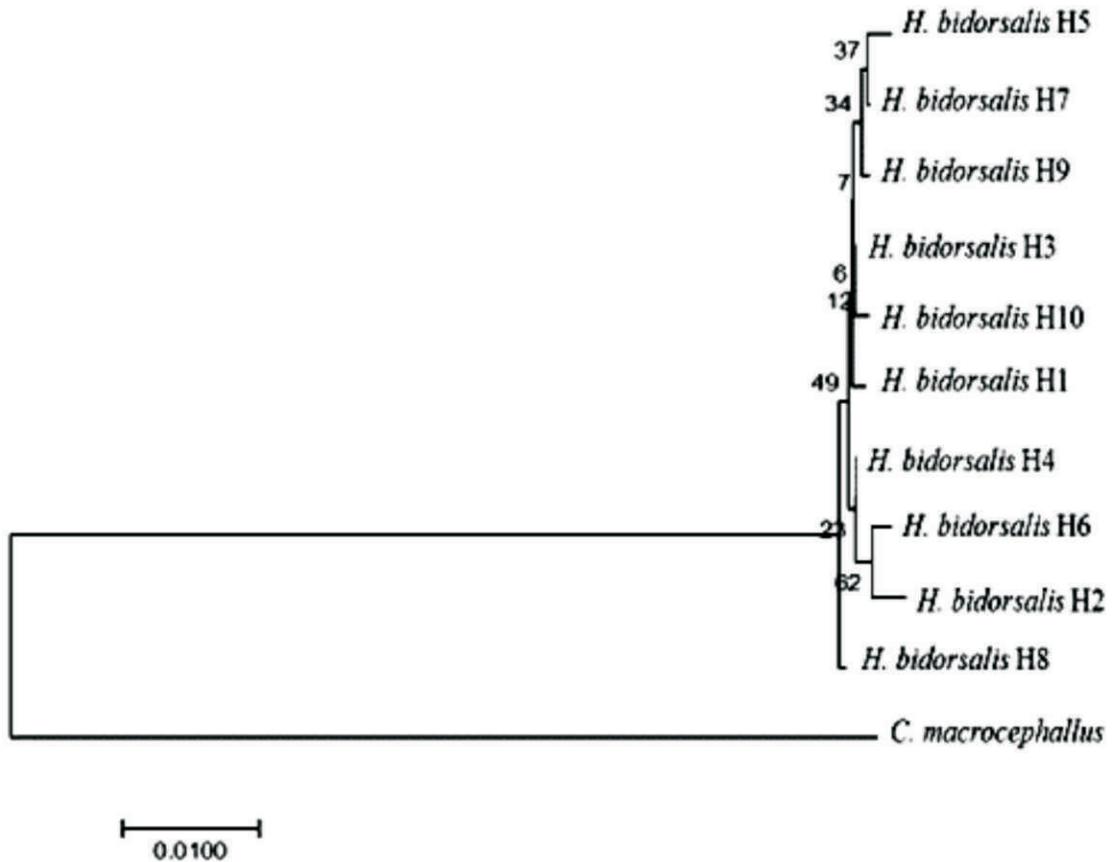


Figure 1: Summary of Neighbour- joining bootstrap tree of *Heterobranchus bidorsalis* cytochrome b gene sequence. The numbers present in the branches are corresponding to bootstrap values based on 1000 replications.

Discussion

The result obtained from this study after sequence blast on Genebank revealed that no sequence lodged is labeled *Heterobranchus bidorsalis* and are distantly related with the *Heterobranchus spp.* earlier lodged by other authors. However, after re-examining the identification and sampling history of the specimens used, it was suspected that the results revealed either that the *Heterobranchus spp.* sequences previously lodged in NCBI GenBank databases could originally be specimens of hybridized species that have been accidentally utilized in cataloging the barcodes and *Cyt b* sequence. Therefore, correct species labelling, morphological

taxonomy and voucher documentation should be prioritized in case that reassessment of spurious data is necessary (Ward *et al.*, 2005). More so hybridization of *H. bidorsalis* with *H. longifilis* (Akinwandeet *al.*, 2009) and *H. bidorsalis* with other clariidae family members for aquaculture purposes is increasingly popular (Aluko *et al.*, 1999; Aluko, 1999; Aluko and Olufeagba 1999; Ipinjolu *at al.*, 2013). However, divergent population of *H. bidorsalis* may exist.

Conclusion

The findings revealed ten haplotypes with seven distinct to river Niger population. This accession could serve as indices for future investigation of

how these variations could help in the description of its character and accessions of the fish's genetic resources, through gene sequence submitted in Genetic data base. The data will help to protect their valuable wild resource and contribute to their recovery. The genetic characterization by the use of mitochondrial DNA Cyt b marker provides baseline information that can support future analysis of *H. bidorsalis* populations and genetic improvement programmes.

References

- Abdul Aziz A., Z. Tarmizi, N. M. Ali, N.A. Ariffin, M.D.D. Abdullah, L.L.Wong, S.Y. Giat, T.N. Aimi, M. Jaafar, A.M.A. Bolong, A.A.A. Kasim and S. Sheriff, (2015). Mitochondrial DNA Diversity of Terubok (*Tenuulosa toli*) from Daro and Mukah, Sarawak Inferred by Partial Cytochrome b (Cyt-B). *J. Fish. Aquat. Sci.*, 10(2): 92-101.
- Adewumi, A.A. and V.F. Olaleye (2011). Catfish culture in Nigeria: progress, prospects and problems. *African Journal of Agricultural Research*, 6(6): 1281-1285.
- Akinwande, A.A., F.O. Moody and S.O. Umar (2009). Growth performance and survival of pure and interspecific *Heterobranchus spp.* hybrids. *Journal of Tropical Agriculture*, 2: 199-203.
- Aluko, P.O. (1998). Growth characteristics of the parental, F₁, F₂ and backcross generation of the hybrids between *Heterobranchus longifilis* and *Clarias angularis*. *West Africa J. Biol. Sci.*, 8: 16-21.
- Aluko, P.O., D. Woru, A. Aremu and O. Popoola (1999). Effect of induced triploid on the growth of *Heterobranchus* and *Clarias* hybrids. *National Institute for Freshwater Fisheries Research. Annual Report*. pp 42-44.
- Aluko, P.O. (1999). Genetic approaches in catfish hybridization: Effort of NIFFR, Nigeria. *National Institute for Freshwater Fisheries Research. Newsletter* 16, 2: 3-8.
- Aluko, P.O. and S.O. Olufeagba (1999). Genetics and food security in Nigeria in the 21st century. Genetically modified fishes, risks, and prospects. *Genetic Society of Nigeria Annual Conference*. Pp 205-206.
- Avise J.C. (2004). Molecular Markers, Natural History and Evolution. 2nd ed. Sinauer Associates, Sunderland, 684 pp.
- Bilio M. (2007). Controlled reproduction and domestication in aquaculture, the current state of the art. Part I–IV. *Aquaculture Europe*. 32 (1).
- Dingley K.H, E.A., Ubick, J.S. Vogel, K.W. Haas (2005). DNA isolation and sample preparation for quantification of adduct levels by accelerator mass spectroscopy. *Methods Mol. Biol.* 2005.
- Excoffier L. and H. Lischer (2011). Arlequin version 3.5: An integrated software for population genetics data analysis. Genetics and Biometry Lab., Univ. of Geneva, Switzerland. 174p.
- Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39, 783–791. doi:10.2307/2408678.
- Fu Y.X., (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925.
- Habib M., W.S. Lakra, V. Mohindra, P. Khare, A.S. Barman, A. Singh, K.K. Lal, P. Punia and A.A. Khan (2010). Evaluation of cytochrome b mtDNA sequences in genetic diversity studies of *Channa marulius* (Channidae: Perciformes). *Mol. Biol. Rep.* Springer Science+Business Media B.V. 2010. 6pp.
- Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41:95–98.
- Hurst, C.D., S.E. Bartlett, W.S. Davidson and I.J. Bruce, (1999). The complete mitochondrial DNA sequence of the Atlantic salmon, *Salmo salar*. *Gene*, 239: 237-242.
- Ipinjolu, J.K., M.Y. Abubakar, I. Magawata and M.I. Buko (2013). Reproductive, survival and growth performance of intergeneric cross of exotic Dutch *Clarias*, *Heterobranchus bidorsalis*, and *Heterobranchus longifilis* in Sokoto North-west Nigeria. *AAFL Bioflux*, 6(6): 11.
- Jukes, T.H. and C.R. Cantor (1969). Evolution of protein molecules. In: H.N. Munro (ed.) *Mammalian Protein Metabolism*, pp 21-132, Academic Press, New York.
- Kumar S., G. Stecher and K. Tamura (2015). MEGA 7; molecular evolutionary Genetics Analysis version 7.0 for bigger dataset. *Molecular Biology and Evolution*.
- Librado, P. and Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451–1452.
- Nei, M. and S. Kumar (2000). *Molecular Evolution and*

- Phylogenetics*. Oxford University Press, Oxford, New York, 333 pp.
- Nei, M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, 512pp.
- Ponzoni, R.W., H.L. Khaw, N.H. Nguyen and A. Hamzah(2010). Inbreeding and effective population size in the Malaysian nucleus of the GIFT strain of Nile tilapia (*Oreochromis niloticus*). *Aquaculture*, 302: 42-48.
- Reed, W., J. Burchard, A.J. Hopson, J. Jenness and I.Yaro (1967). *Fish and Fisheries of Northern Nigeria*. Gaskiya Corporation, Zaria. 226 p.
- Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457-462.
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123:585-595.
- Tamura, K. and M. Nei (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10(3):512-526.
- Thompson J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin and D.G. Higgins (1997). The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 25(24): 4876-4882 DOI10.1093/nar/25.24.4876.
- Xiao, W., Y. Zhang and H. Liu (2001). Molecular systematics of Xenocyprinae (Teleostei: Cyprinidae): taxonomy, biogeography and coevolution of a special group restricted in East Asia. *Molecular Phylogenetics and Evolution*, 18:163-173.
- Yisa, M. and S.O. Olufeagba (2005). An exposition on field identification of clariid catfishes as an important tool in fish breeding and genetics. *Proceedings of the 19th Annual Conference of the Fisheries Society of Nigeria (FISON)*, 29 Nov-03 2004, Ilorin, Nigeria. Pp189-193.



AJFARM

www.theajfarm.com

*ABUBAKAR, M.Y.¹, IPINJOLU, J.K.¹, ESA, Y.B.³ MAGAWATA, I.¹ AND HASSAN, W.A.²
 African Journal of Fisheries and Aquatic Resources Management

Volume 3, 2018

ISSN: 2672-4197 (Prints)

ISSN: 2672-4200 (Online)

Pp 1-9