Journal of Scientific Research & Reports

20(3): 1-15, 2018; Article no.JSRR.42950
ISSN: 2320-0227

Single Nucleotide Polymorphisms and Haplotype Analyses in Tilapia Fish Inferred from mtDNA D-loop and Cyt-b Regions

E. V. Ikpeme[1], O. U. Udensi[2], E. E. Ekerette[1*] and M. O. Ozoje[1,3]

[1]Animal Genetics and Genomics Unit, Department of Genetics and Biotechnology, University of Calabar, PMB 1115, Calabar, Nigeria.
[2]Plant Resource Management and Genomics Unit, Department of Genetics and Biotechnology, University of Calabar, PMB 1115, Calabar, Nigeria.
[3]Department of Animal Breeding and Genetics, Federal University of Agriculture, PMB 2240, Abeokuta, Nigeria.

Authors' contributions

This work was carried out in collaboration between all authors. Authors EVI and OUU designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Authors EEE and MOO managed the analyses of the study. Author EEE managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JSRR/2018/42950
Editor(s):
(1) Dr. Mahmoud Nasr, Sanitary Engineering Department, Faculty of Engineering, Alexandria University, Egypt.
(2) Dr. Mario A. Pagnotta, Department of Science and Technologies for Agriculture, Forestry, Nature and Energy (DAFNE), Tuscia University, Italy.
(3) Dr. Leszek Labedzki, Professor, Institute of Technology and Life Sciences, Kujawsko-Pomorski Research Centre, Poland.
Reviewers:
(1) Haifa El-Hentati, Tunisia.
(2) Bensu Karahalil, Gazi University, Turkey.
(3) Zafer Bulut, Selcuk University, Turkey.
Complete Peer review History: http://www.sciencedomain.org/review-history/26611

Original Research Article

Received 29 May 2018
Accepted 21 August 2018
Published 11 October 2018

ABSTRACT

Objective: This research was aimed at analysing single nucleotide polymorphisms and haplotypes on D-loop and Cyt-b regions of the mitochondrial DNA of tilapia fish.
Methods: Fifteen and thirteen tilapia fish were obtained from two populations, south-south (Domita farm) and south-west (Odeda farm). DNA extraction from fish tissue was done using Quick-gDNA^{TM} mini Prep kit after which PCR amplification was carried out. Sequencing of the two mtDNA regions were done using forward primer 5'- GGATTYTAACCCYTRCCCC- 3' and reverse 3'-AGTAAGTCAGGACCAAGCC-5' for D-loop and 5'-GGATTTTAACCCTTACCCC-3' and 3'-

**Corresponding author: E-mail: ekemeks4life@yahoo.com;*

AGTAAAGTCAGGACCAAGCC-5' for Cyt-b region. Statistical analyses were carried out on the aligned sequenced data using MEGA version 6.06, Dnasp 5.1, Codon code aligner 6.06 as well as NETWORK 4.6.1.1.

Results: mtDNA polymorphism was highest in the D-loop of South-South (SS) population with 176 polymorphic sites, while South-West (SW) population had 162 polymorphic sites translating to 176, 162 and 144 SNPs with higher non-synonymous substitutions than synonymous substitutions. Haplotype diversities (Hd) were 1.00 ± 0.024 and 1.00 ± 0.030 while nucleotide diversities were 0.168 ± 0.086 and 0.161 ± 0.084 for D-loop of SS and SW populations, respectively. For Cyt b region, haplotype and nucleotide diversities were 0.91 ± 0.003 and 0.051 ± 0.016. Positive selection was more on mtDNA D-loop of tilapia sampled from SS than those from the SW as well as Cyt-b region of tilapia fish from SS. 28 haplotypes were identified among the tilapia from SS and SW with no shared haplotypes while 9 haplotypes were identified from the Cyt-b region with haplotypes 4, 5, 6 and 7 shared between species. Median-joining network analysis revealed population-based clustering pattern. The demographic expansion was not significant using Tajima's D and Fu's F statistics.

Conclusion: Higher SNPs were revealed in mtDNA D-loop when compared with mtDNA Cyt-b region of tilapia fish.

Keywords: SNPs; haplotype; mtDNA D-loop and Cyt-b; tilapia fish; variation.

1. INTRODUCTION

Tilapia fish is the most widely harvested and consumed fish after carp globally [1]. Commercially, tilapia fish is highly consumed with a global harvest of approximately, 4,677,613 tonnes in 2013 [2]. For emphasis, China is the world's largest producer of tilapia with about 1,600,000 tonnes production, while in Africa; Egypt is the largest producer of about 800,000 tones [2]. Nigeria has an estimated population of over 160 million with a coastline measuring approximately 853 kilometres. This vast coastline according to Osagie [3] can be harnessed for tilapia fish farming, which might probably have the capacity to make a significant contribution to agriculture.

According to Oyakhilomen and Zibah [4], the demand-supply gap of fish is about 1.8million tonnes. This was premised upon the report that the annual demand for fish in Nigeria to be about 2.66million tonnes. Tilapia fish has received little or no research attention comparatively, its importance as rich protein, potassium, phosphorus, vitamin B12 as well as low-fat content [5,6] notwithstanding. The other frightening important issue is the genetic erosion in this species of fish orchestrated by indiscriminate and over-exploitation from the wild by fishermen in the bid to bridging the demand-supply gap. The implication, therefore, is that if efforts are not intensified, especially in research geared towards domesticating, conserving and integrating tilapia farming into the agricultural programme in Nigeria, it might spell doom.

However, the complaint expressed by both breeders and farmers is poor growth, which affects market competition. To mitigate this shortcoming, a robust diversity analysis becomes imperative, which will be aimed at screening and selecting superior genotypes of tilapia fish for breeding and improvement. Undoubtedly, allelic differences between and among organisms of the same species have become an integrated part in agricultural/breeding programmes, which is aimed at selecting breeding stock and identifying extinct-threatened species for possible conservation measures [7]. Myriads of techniques and methods have been developed and adopted for genetic analysis such as morphometric-based technique. Unfortunately, though simple and direct for the identification and characterisation of fish stock [8,9], its reliability has been marred by environmental factors [10].

The specificity and resolution power of other molecular markers such as RAPD, SSR, ISSR, AFLP, notwithstanding, DNA sequence data have been reported to be more informative in explaining genetic relatedness among species of organisms. Out of the two major genomic DNAs – nuclear and mitochondrial, mitochondrial DNA (mtDNA) have been used in studying stock structures invertebrates such as fish [11,12,13], birds [14,15], reptiles [16,17] as well as mammals [18,19]. Habib et al. [20] reviewed that mtDNA variation is being adopted as a dependable tool for determining genetic diversity within and among species. The choice of mtDNA is hinged on the fact that vertebrates show more variations in mitochondrial genome than nuclear

DNA due to rapid mutation rate and copy number per cell [21,22] as well as being inherited maternally [23].

Two regions along the mtDNA have attracted research interest, especially in constructing phylogenetic relationship among or between species. These are cytochrome –b (cyt-b) and D-loop regions. Cyt–b region of mtDNA has been used widely in genetic diversity studies of many animal species widely fish [22,24,25,26]. Similarly, the hyper-variable/control region (D-loop) of mtDNA flanked by tRNApro and tRNAphe genes in the mitochondrial genome has been used for characterising tilapia species [13].

According to Chambers et al. [27], phylogenetic analysis has been used in tracing the origin and evolution of species, prediction of physiological, biochemical and structural features of sequences. It is also used to assess ancestral history and relationships [28]. There are several landmarks on the DNA that can be utilised for the identification and characterisation of species of organisms, especially Single Nucleotide Polymorphisms (SNPs). SNP-based research is geared towards studying the genetic differences between species for the prediction of phenotypes and phylogeny. Giving that when SNPs occur inside a gene, they create different variants or alleles of that gene and the sequences tend to be transmitted unchanged across generation; in this present study, using SNPs markers on D-loop and cyt-b regions of mtDNA, genetic diversity, polymorphisms, and haplotype in tilapia fish sampled from two populations in Nigeria were investigated.

2. MATERIALS AND METHODS

2.1 Sample Collection

Twenty-eight matured tilapia fish were collected from two populations; 15 samples from Domita fish farm in Akwa Ibom State at approximately 5°1'4"N, 7°59'52"E for SS population and 13 samples from Odeda fish farm in Ogun State at approximately 7°13′00″N, 3°31′00″E for SW population (Fig. 1). The Sampling protocol was according to Samaradivakara et al. [9] with modifications.

2.2 DNA Extraction from Fish Tissue

To ensure homogeneity, tissue samples were cut off from the dorsal region of the left-hand side of each fish and preserved in 95% ethanol of good analytical grade for DNA extraction and analyses. Extraction of mtDNA and purification from the tissue samples was carried out in Biotechnology Laboratory Unit of Animal Science Department, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria using Quick-gDNA^{TM} MiniPrep kit (Zymo Research, USA). Beta-mecaptoethanol was added to the lysis buffer to a final solution of 500µl per 100ml according to the manufacturer's instruction. Approximately, 25mg of each fish tissue was homogenised in 500 µl of genomic lysis buffer in Eppendorf tube. The lysate was centrifuged at 10000 × g for 5 minutes using a Centurion Scientific microcentrifuge (Model: C2015, USA). The supernatant was carefully transferred to a Zymo-Spin^{TM} column in a collection tube and centrifuged at 10000 × g for 1minute. Both the flow through and the collection tube were then discarded before transferring the Zymo-Spin^{TM} column to a new collection tube. 200µl of DNA pre-wash buffer was added to the spin column and centrifuged at 10000 × g for 1 minute. 500µl of g-DNA wash buffer was added to the spin column and further centrifuged at 10000 × g for another 1 minute. The spin column was transferred to a clean microcentrifuge tube, and ≥50µl of DNA elution buffer was added to the spin column. This was incubated for 2-5 minutes at room temperature and afterwards centrifuged at top speed (14000 × g) for 30 seconds to elute the DNA. The eluted DNA was stored at -20ºC for further analysis.

2.3 Polymerase Chain Reaction (PCR) Amplification

PCR amplification was carried out in STABVIDA Laboratory, Quinta de torre, Portugal. The primers Fish-comum-D-loop Forward (5'-GGATTYTAACCCYTRCCCC- 3') and Czilli Reverse (3'- AGTAAAGTCAGGACCAAGCC 5') for D-loop region and Czilli Forward (5' –GGATTTTAACCCTTACCCC-3') and Czilli Reverse (3'– AGTAAAGTCAGGACCAAGCC –5') for cyt-b region of the mtDNA were used. PCR for each mtDNA region was carried out using 25µl reaction volume containing 1µlgenomic DNA, 2mM $MgCl_2$, 200µM of dNTP, 2.5µl of 10xPCR buffer, 1µM of each primer and two units of STABVIDA proprietary Taq polymerase. This was performed using the GeneAmp® PCR system (9700 thermal cycler, USA) with the cycling condition of initial denaturation step at 95ºC for 5 minutes, followed by 25 cycles of denaturation at 94ºC for 40

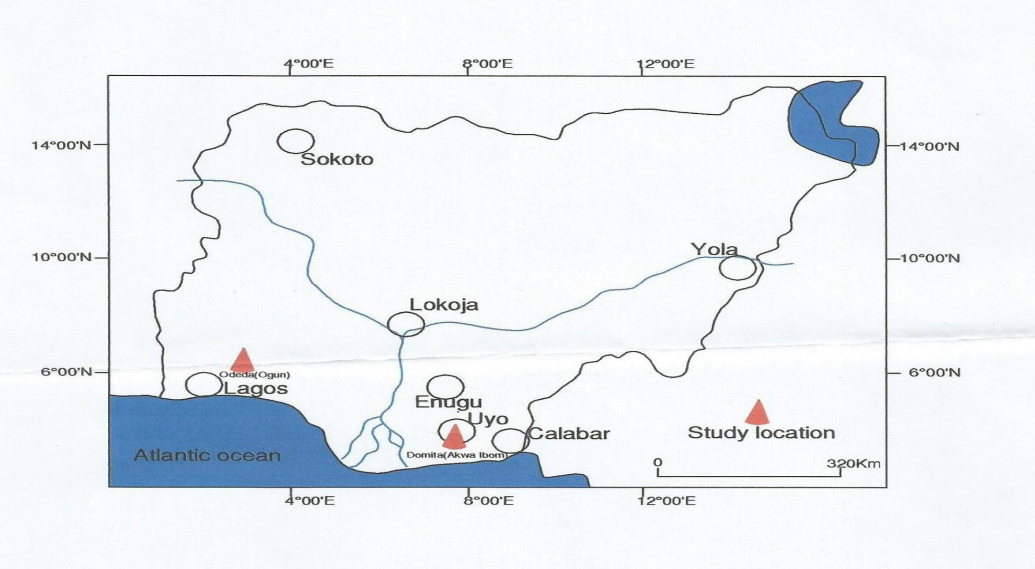

Fig. 1. A map showing the sampling locations

seconds, annealing at 54ºC for 45 seconds, extension at 72ºC for 1 minute and final extension at 72ºC for 7 minutes . PCR products were purified using exofast protocol according to the manufacturer.

2.4 Sequencing of D-loop and Cytochrome Regions of mtDNA

D-loop and cyt-b regions of the mtDNA were sequenced for all tissue-DNA extracts using the primers (5'-GGATTYTAACCCYTRCCCC-3') and Czilli Reverse (3'-AGTAAAGTCAGGACCAAGCC 5') for D-loop region and Czilli Forward (5' – GGATTTTAACCCTTACCCC -3') and Czilli Reverse (3' – AGTAAAGTCAGGACCAAGCC – 5') for cyt-b region. Sequencing reaction was performed in STABVIDA Laboratory, Quinta de torre, Portugal with AB13730×L sequencer using 20µl reaction comprising approximately, 20ng of purified PCR product as template DNA, 8µl of Big Dye Terminator Reaction Mix (dNTPs, ddNTPs, buffer, enzyme and $MgCl_2$), 8µl of deionized water, 2µl of primer programmed as 25cycles at 96ºC for 10 seconds, 60ºC for 5 seconds and 60ºC for 4 minutes. The length of the sequenced fragments ranged from 950 to 1366bp for D-loop and 1241 to 1401bp for cyt- b.

2.5 Statistical Analyses

Viewing and editing of the sequences were carried out using Bioedit software version 7.2.5 [29]. Multiple sequence alignment and determination selection types were performed using Molecular Evolutionary Genetics Analysis (MEGA) version 6.06 [30] excluding all the gaps. CodonCode Aligner version 6.06 was used to analyse the mutation of SNPs in the aligned sequences while NETWORK 4.6.1.1 was used for network analysis of the different haplotypes [31]. DnaSP 5.1 was adopted to test the demographic expansion and to estimate the mtDNA polymorphism in the populations.

3. RESULTS

3.1 Genetic Diversity of Tilapia Fish in the Studied Populations

The accession numbers for mtDNA D-loop sequence for SS and SW tilapia population are MF385001 and MF385002 while Cyt b sequence for SS tilapia population was MF384326 for your reference. The total number of the aligned site was 745 for D-loop and 1022 for Cyt-b region. mtDNA polymorphism was highest in the D-loop of SS population with 176 polymorphic sites, while SW population had 162 polymorphic sites. The lowest polymorphism was observed in Cyt-b of SW population with 144 polymorphic sites. Parsimony information site detected were 99 (SS D-loop), 124 (SW D-loop) and 112 for SS Cyt-b. Haplotype diversities (Hd) were 1.00 ± 0.024 and 1.00 ± 0.030 while nucleotide diversities were 0.168 ± 0.086 and 0.161 ± 0.084 for D-loop of SS and SW populations, respectively. For Cyt b

region, haplotype and nucleotide diversities were 0.91 ± 0.003 and 0.051 ± 0.016. However, Cyt-b region of mtDNA showed more sequence conservation of 85.9% than D-loop, which had 76.3% (SS D-loop) and 78% (SW D-loop), respectively (Table 1).

3.2 Mutation Analysis of SNPs

The mtDNA D-loop of tilapia population from south-west revealed 176 SNPs resulting in 155 non-synonymous and 21 synonymous mutations while mtDNA D-loop of tilapia population from SS detected 162 SNPs, which resulted in 148 non-synonymous and 14 synonymous mutations, respectively (Table 2). In the SS population, there was a high prevalence of lysine substitution, which occurred 23 times out of which 21 non-synonymous and 2 synonymous mutations were produced. Asparagine substitution occurred 20 times, resulting in non-synonymous mutation, while valine substitution occurred two times resulting in non-synonymous mutations. There was a high rate of purine to pyrimidine (pyrimidine to purine) substitution, which resulted in 108 transversions when compared with purine to purine (pyrimidine to pyrimidine) substitution, which resulted in 68 transition mutation. For the South West tilapia population, isoleucine and leucine substitution had the same prevalence rate of 16, while the former resulted in 13 synonymous and 3 synonymous, the later resulted in 15 non-synonymous and 1 synonymous mutation. It was also observed that methionine had the least prevalence occurring two times which were all non-synonymous. There was a high rate of purine to pyrimidine (pyrimidine to purine) substitution resulting in 90 transversions when compared with purine to purine (pyrimidine to pyrimidine) substitution resulting in 72 transition mutations.

From the mtDNA cyt- b sequence of the SS tilapia population,144 SNPs were detected resulting in 137 non-synonymous and 7 synonymous mutations. Leucine had the highest substitution prevalence of 25, which gave all non-synonymous mutation followed by Serine with a substitution prevalence of 21 while methionine was the least substituted amino acid. There was a high rate of purine to pyrimidine (pyrimidine to purine) substitution resulting in 91 transversion in comparison with purine to purine (pyrimidine to pyrimidine) substitution, which resulted in 53 transition mutation (Table 3).

3.3 Selection Analysis

Selection analysis revealed positive and negative selection in the populations. The mtDNAD-loop of tilapia from SS population revealed non-synonymous to synonymous substitution rate of 56.93 (47 positive site index) and -45.834 (26 negative site index), while mtDNA D-loop of tilapia sampled from SW population was 42.511 (50 positive site index) and -34.140 (28 negative site index). However, mtDNA cyt-b of tilapia sampled from SS population had 33.391 and -18.543 non-synonymous to synonymous substitution for 47 positive and 23 negative selection site indexes, respectively (Table 4). Comparatively, mtDNA D-loop of tilapia sampled from SS had more positive selection than mtDNA D-loop of tilapia from south-west population as well as mtDNA cyt-b region.

Table 1. Genetic diversity parameters among tilapia fish from the studied populations

Diversity indices	SS D-loop	SW D-loop	SS Cyt b
Number of sequences	15	13	15
Number of sites	745	745	1022
Monomorphic sites	568	574	878
Polymorphic sites	176	162	144
Singleton variable site	77	38	29
Parsimony information site	99	124	112
Number of haplotypes	15	13	9
Haplotype diversity (Hd)	1.00 ± 0.024	1.00 ± 0.030	0.91 ± 0.003
Nucleotide diversity (π)	0.168 ± 0.086	0.161 ± 0.084	0.051 ± 0.016
Average number of pairewise differences	52.07	49.14	52.114
Sequence conservation	0.763(76.3%)	0.780(78%)	0.859 (85.9%)
Minimum number of recombination	42	18	1

SS D-loop (South-South mitochondrial D-loop region)
SW D-loop (South-West mitochondrial D-loop region)
SS cyt-b (South-South mitochondrial cytochrome b region

Table 2. Mutation analysis of single nucleotide polymorphism (SNPs) in mtDNA D-loop of tilapia fish from south- south and south-west, Nigeria

S/N	mtDNA D- loop (SS)				mtDNA D-loop (SW)			
	SNP	**Amino acid change**	**Syn/non-syn**	**Mutation types**	**SNP**	**Amino acid change**	**Syn/non-syn**	**Mutation types**
1	68A>G	Ser23Asn	non-syn	Transition	2C>T	Ala1Val	Non-syn	Transition
2	71A>C	Gln24Pro	Non-syn	Transversion	6T>G	Tyr2STP	Non-syn	Transversion
3	117G>A	Lys39Lys	Syn	Transition	13T>G	Tyr5Asp	Non-syn	Transversion
4	155T>A	Leu52STP	Non-syn	Transversion	14A>G	Tyr5Cys	Non-syn	Transition
5	181G>C	Lys60Asn	Non-syn	Transversion	16C>G	His6Glu	Non-syn	Transversion
6	205T>A	Ser68Arg	Non-syn	Transversion	17A>G	His6Arg	Non-syn	Transition
7	226A>C	Gln75His	Non-syn	Transversion	18C>G	His6Glu	Non-syn	Transversion
8	232G>C	Ser77Ser	Syn	Transversion	22T>A	STP8Lys	Non-syn	Transversion
9	243A>C	Gln81Pro	Non-syn	Transversion	24A>T	STP8Tyr	Non-syn	Transversion
10	254T>C	STP85Gln	Non-syn	Transition	26T>A	Phe9Tyr	Non-syn	Transversion
11	256G>A	STP85Gln	Non-syn	Transition	30G>A	MET10Ile	Non-syn	Transition
12	257T>C	STP86Gln	Non-syn	Transition	34A>T	Asn12Tyr	Non-syn	Transversion
13	295T>C	Gly98Gly	Syn	Transition	35A>G	Asn12Ser	Non-syn	Transversion
14	420T>A	Val140Asp	Non-syn	Transversion	41C>G	Ser14Cys	Non-syn	Transversion
15	486T>C	Val162Ala	Non-syn	Transition	50G>A	STP17Tyr	Non-syn	Transversion
↓								
176	**176SNPs**	**176**	**155/21**	**108/68**	**162SNPs**	**162**	**148/14**	**90/72**

Table 3. Mutation analysis of single nucleotide polymorphism (SNPs) in mtDNA D-loop and Cyt-b of tilapia fish from south- south, Nigeria

S/N	MtDNA D- loop				mtDNACyt- b			
	SNP	Amino acid change	Syn/non-syn	Mutation types	SNP	Amino acid change	Syn/non-syn	Mutation types
1	68A>G	Ser23Asn	non-syn	Transition	4A>G	Lys2Gly	Non-syn	Transition
2	71A>C	Gln24Pro	Non-syn	Transversion	5A>G	Lys2Gly	Non-syn	Transition
3	117G>A	Lys39Lys	Syn	Transition	14T>G	Ile5Arg	Non-syn	Transversion
4	155T>A	Leu52STP	Non-syn	Transversion	24T>G	Ser8Arg	Non-syn	Transversion
5	181G>C	Lys60Asn	Non-syn	Transversion	26T>G	Phe9Cys	Non-syn	Transversion
6	205T>A	Ser68Arg	Non-syn	Transversion	28T>G	Cys10Gly	Non-syn	Transversion
7	226A>C	Gln75His	Non-syn	Transversion	39T>G	Phe13Leu	Non-syn	Transversion
8	232G>C	Ser77Ser	Syn	Transversion	40A>T	Asn14Tyr	Non-syn	Transversion
9	243A>C	Gln81Pro	Non-syn	Transversion	45G>A	Gln15Gln	Syn	Transition
10	254T>C	STP85Gln	Non-syn	Transition	73T>G	Leu25Val	Non-syn	Transversion
11	256G>A	STP85Gln	Non-syn	Transition	135C>A	Arg45Arg	Syn	Transversion
12	257T>C	STP86Gln	Non-syn	Transition	815T>C	Phe272Ser	Non-syn	Transition
13	295T>C	Gly98Gly	Syn	Transition	817C>G	Arg273Gly	Non-syn	Transversion
14	420T>A	Val140Asp	Non-syn	Transversion	818G>T	Arg273Leu	Non-syn	Transversion
15	486T>C	Val162Ala	Non-syn	Transition	821C>G	Ala274Gly	Non-syn	Transversion
176	**176SNPs**	**176**	**155/21**	**108/68**	**144SNPs**	**144**	**137/7**	**91/53**

Table 4. Selection analysis of tilapia fish from South-South and South-West Nigeria

Populations	Selection type	d_N	d_S	d_N-d_S	Site index	P-value
	Positive	65.784	8.854	56.93	47	0.005
SS D-loop	Negative	21.047	66.885	-45.834	26	0.009
	Neutral	0.000	0.000	0.000	146	0.000
	Positive	51.152	8.641	42.511	50	0.006
SW D-loop	Negative	18.593	52.733	-34.140	28	0.008
	Neutral	0.000	0.000	0.000	136	0.000
	Positive	42.599	9.206	33.391	47	0.007
SS cyt b	Negative	9.100	27.648	-18.548	23	0.009
	Neutral	0.000	0.000	0.000	251	0.000

SS D-loop (South-South mitochondrial D-loop region)
SW D-loop (South-West mitochondrial D-loop region)
SS cyt b (South-South mitochondrial cytochrome b region)
d_N(non- synonymous)
d_S (synonymous)

3.4 Network Analysis of Haplotype Variations

For haplotype analysis, sequence data from mtDNA D-loop of tilapia sampled from SS and SW populations were pooled together where28 haplotypes were identified having no shared haplotypes between the samples (Table 5). The median-joining network analysis, however, revealed an interesting result as haplotype 1-15 associated with D-loop of SS samples were clustered together while haplotype 16-28 associated with D-loop of SW samples had two groups (Fig. 2). Similarly, 9 haplotypes were identified from 15 samples sequenced for mtDNA cyb-b of tilapia samples from S/S population. Haplotype 5 was shared by 4 (SS5, SS10, SS11 and SS14) individuals with a frequency of 0.308. Three major groups identified - group 1 (haplotype 1, 2, 4, 5, 6 & 9), group 2 (haplotype 7 & 8) and group 3 (haplotype 3), respectively (Fig. 3).

3.5 Demographic Expansion

Tajima's D values were -1.080, -0.83, and 0.673, while the Fu's Fs values were -1.68, -1.206, and 0.584 for samples in S-S D-loop, S-W D-loop and S-S cyt-b, respectively (Table 5). With the probability values, it showed that the Tajima's D and Fu's Fs statistics were not significant (Table 6).

4. DISCUSSION

Genetic diversity analyses using any technique is usually geared towards the improvement of species of interest as well as mitigate genetic erosion. Obviously, the different techniques have different resolution and informativeness, which is very cardinal in achieving these goal(s). The extent of diversity found in a species/population if harnessed properly could be exploited for the effective management, conservation and improvement of species.

Table 5. Haplotype variation in mtDNA D-loop and Cyt-b of tilapia fish from South-South Nigeria

D-loop			Cyt b		
Haplotypes	**Samples**	**Frequency**	**Haplotypes**	**Samples**	**Frequency**
H_1	SS1	1 (0.036)	H_1	SS1	1 (0.067)
H_2	SS2	1 (0.036)	H_2	SS2	1 (0.067)
H_3	SS3	1 (0.036)	H_3	SS3	1 (0.067)
H_4	SS4	1 (0.036)	H_4	SS4 and SS12	2 (0.133)
H_5	SS5	1 (0.036)	H_5	SS5,SS10, SS11and SS14	4 (0.267)
H_6	SS6	1 (0.036)	H_6	SS6and SS9	2 (0.133)
H_7	SS7	1 (0.036)	H_7	SS7and SS8	2 (0.133)
H_8	SS8	1 (0.036)	H_8	SS13	1 (0.067)
H_9	SS9	1 (0.036)	H_9	SS15	1 (0.067)
H_{10}	SS10	1 (0.036)			
H_{11}	SS11	1 (0.036)			
H_{12}	SS12	1 (0.036)			
H_{13}	SS13	1 (0.036)			
H_{14}	SS14	1 (0.036)			
H_{15}	SS15	1 (0.036)			
H_{16}	SW1	1 (0.036)			
H_{17}	SW2	1 (0.036)			
H_{18}	SW3	1 (0.036)			
H_{19}	SW4	1 (0.036)			
H_{20}	SW5	1 (0.036)			
H_{21}	SW6	1 (0.036)			
H_{22}	SW7	1 (0.036)			
H_{23}	SW8	1 (0.036)			
H_{24}	SW9	1 (0.036)			
H_{25}	SW10	1 (0.036)			
H_{26}	SW11	1 (0.036)			
H_{27}	SW12	1 (0.036)			
H_{28}	SW13	1 (0.036)			

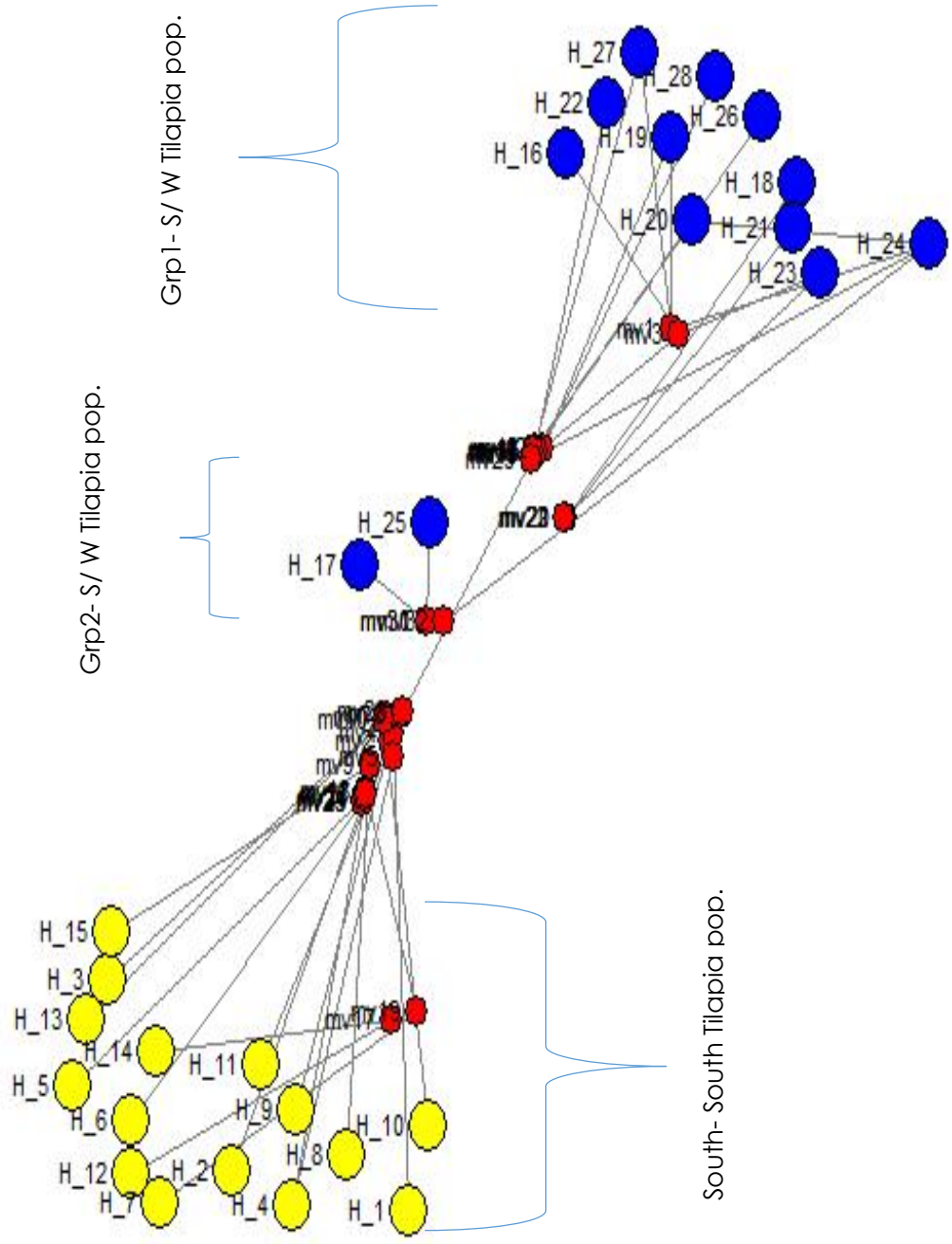

Fig. 2. Median joining network analysis of the 28 haplotypes identified on mtDNA D-loop of tilapia fish from South-South and South-West populations

Table 6. Demographic expansion indices of tilapia fish in the studied population

Parameters	SS D-loop	SW D-loop	SS cyt b
Tajima's D	-1.080 (p>0.10)	-0.83(p>0.10)	0.673 (p>0.01
Fu's F	-1.681(p>0.10)	-1.206 (p>0.10)	0.584 (p>0.01)

SS D-loop (South-South mitochondrial D-loop region)
SW D-loop (South-West mitochondrial D-loop region)
SS cyt-b (South-South mitochondrial cytochrome b region)

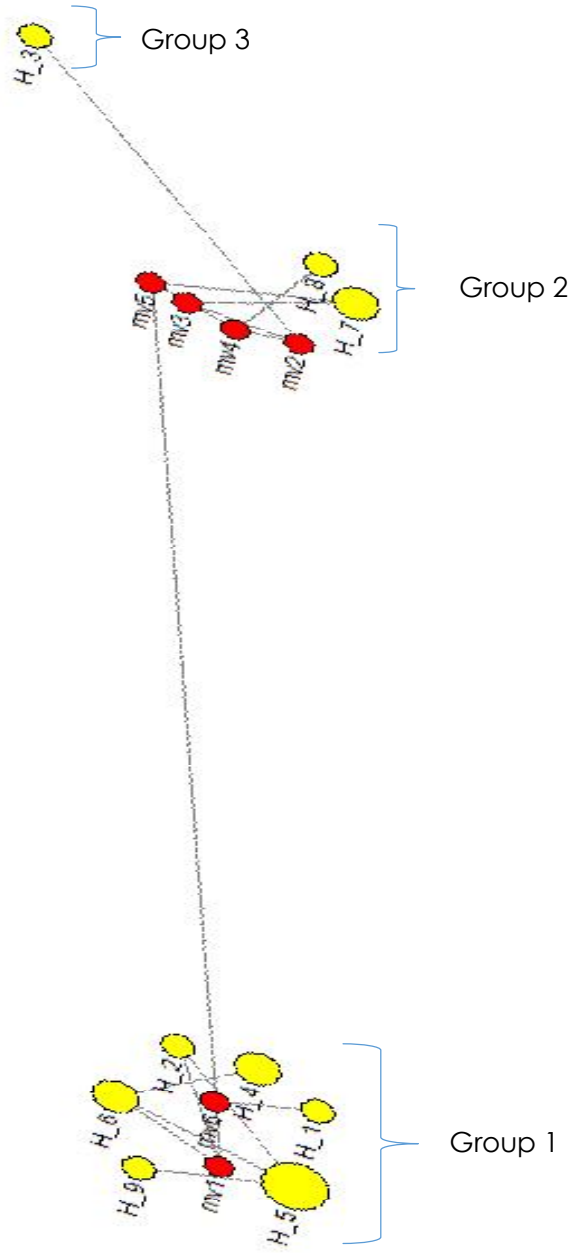

Fig. 3. Median joining network analysis of the 9 haplotypes identified on mtDNA cyt-b of tilapia fish from South-South population

Of the myriad of biological data used for genetic diversity analysis, sequence data have proven to be more informative in resolving genetic differences and relatedness among species. Of interest are the sequence data from mitochondrial DNA. According to Habib et al. [20], mtDNA variation is being adopted as a dependable tool for determining genetic diversity within and among species. This study becomes imperative owing to the observed genetic erosion in tilapia fish orchestrated by the continuous over-exploitation from the wild by local fishermen as well as low research interest, which has hampered breeding, cultivation and improvement.

Our result showed that the polymorphic sites for mtDNA D-loop (SS), D-loop (SW) and Cyt-b

(S/S) were 176, 162 and 144, respectively with parsimony information site of 99, 124 and 112. There were also variations in the haplotype and nucleotide diversity. Agbebi et al. [32] reported a haplotype diversity of 1.00 in four tilapia species using mtDNA D-loop sequence. Abdul et al. [22] reported haplotype and nucleotide diversity of 0.232 and 0.00321 in Terubok fish using Cyt-b sequence, which was lower than the present result using Cyt-b sequence. The implication, therefore, is that different fish species have varying haplotype and nucleotide diversity using the same region of the mtDNA. The relatively high haplotype and nucleotide diversity observed in this study might suggest high molecular differences within and between the populations [33]. It is already obvious that the D-loop region of the mtDNA is highly variable when compared with the Cyt-b region. The number of recombination processes (42) and 18 in the mtDNA D-loop as compared with 1 for Cyt-b region might have increased the observed polymorphism.

Conserved sequences are similar or identical sequences in nucleic acids, proteins or polysaccharides across species. The implication of sequence conservation is that a sequence has been maintained by evolution speciation, notwithstanding. Sequence conservation of 85.9% for mtDNA Cyt-b region might have contributed to the reduced polymorphic sites available in the mtDNA region. Going by the report of [34], D-loop mtDNA region has high evolution rates.

The informativeness of mtDNA is hinged on the rapid mutation rate [19,21] and it's maternally inheritance fashion [23]. Single nucleotide polymorphism (SNP) represents the most widespread type of sequence variation in genome, which has emerged as valuable genetic markers for revealing the evolutionary history of populations [35]. Lorenc et al. [36] pointed out that SNPs are becoming the dominant form of molecular markers for genetic and genomic analyses. Understandably, the occurrence of SNPs inside a gene creates different variants or alleles of the gene, which tend to be inherited or transmitted unchanged across generations. 176 SNPs, 162 SNPs and 144 SNPs were detected from the three tilapia populations studied, which corresponds with the number of polymorphic sites earlier reported. For all the SNPs detected, the non-synonymous substitution was higher than synonymous substitution giving rise to higher positive selection. What it does therefore suggest is that evolutionary distance based on non-synonymous substitutions is expected to be greater than synonymous substitutions. According to Pennings and Hermission [37], there could be the occurrence of a selective sweep, which reduces or eliminates the variation among the nucleotides near a mutation in DNA. This might be due to a beneficial alleles having recently reached fixation as a result of strong positive natural selection.

The observed differences in the number of SNPs detected in the D-loop of the tilapia fish sampled from S/S and S/W could be explained from the premise that there might have been the occurrence of rare or previously non-existing allele whose prevalence has increased in the population. The implication is that genetic variants on the DNA of the beneficial allele become more prevalent [38]. Understandably, organisms that have been genetically manipulated are subjected to artificial selective pressure as well as forced- adaptation to the new environment, which will provide a baseline from which different varietals could have emerged through selective sweeps. However, this was not the case for the plant materials used were landraces [39].

It might be important to highlight that the identification of rare alleles within conserved sequences can be complied with assessing the risk of diseases. As was out earlier that the non-synonymous substitutions were greater than synonymous substitutions. What this might suggest is the possibility of altering the biological functionality due to the alteration of the amino acid sequence of the protein. According to Barry et al. [40], a haplotype is a set of SNPs on one chromosome that tends to always occur together or a group of genes in an organism that are inherited together from a single parent. The interesting observation in our result is the fact that there was no sharing of a haplotype for the mtDNA D-loop, the population notwithstanding. This is a variant with the Cyt-b region of mtDNA where H_4, H_5, H_6 and H_7 were shared. This might have led to the observed haplotype diversity of 0.91. Ekerette et al. [41] reported two clusters of tilapia fish based on population (SS and SW). This also played out in the median-joining network analysis of the 28 haplotypes identified in mtDNA D-loop. This might further explain the fact that there are differences in the two populations.

For the demographic expansion analysis, Tajima's D and Fu's F were carried out, which distinguishes between a DNA sequence evolving randomly and one evolving under a non-random process; computing a standardized measure of the total number of segregating sites or DNA sites that are polymorphic in the sampled DNA and the average number of mutations between pairs in the sample [42,43]. We report negative Tajima's D value for mtDNA D-loop tilapia fish samples (-1.080 and -0.83) while mtDNA Cyt-b region had a positive Tajima's D value of 0.673. It has been reported that a negative Tajima's D signifies an excess of low frequency polymorphisms relative to expectation, indicating population size expansion. For emphasis, when Tajima's D = 0, it means that observed variation is similar to expected variation while Tajima's D < 0 implies that rare alleles are present at high frequencies, having fewer haplotype. However, when Tajima's D > 0, it implies sudden population contraction. The result obtained for mtDNA D-loop showed that there might have to be a recent selective sweep giving rise to lower average heterozygosity than the number of segregating lines while for Cyt-b region where Tajima's D value (0.673) reveals sudden population contraction. The implication, therefore, is that the more segregating sites, the more DNA sites that will be polymorphic. This may have informed polymorphic sites detected in the mtDNA D-loop.

5. CONCLUSION

The results put together revealed higher nucleotide and haplotype diversity, more SNPs detected, positive selection and unshared haplotype in mtDNA D-loop when compared with mtDNA Cyt-b region of tilapia fish, giving rise to more polymorphism. By implication, tilapia fish within these populations could be employed in selective breeding and genetic improvement, especially from the south- south population.

ETHICAL APPROVAL

The research was performed on biological material derived from fish obtained from fish farms. After obtaining the tissue for analysis, the meat was standard for consumption. Therefore, our research did not require the approval of the Animal Experimentation committee.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Food and Agriculture Organization. Fishery and Aquaculture Statistics. Rome: FAO Publication; 2009.
2. Fitzsimmons K, Brunno C, Loc T. Tilapia global supply and demand in 2014. Austria, Adelaide; 2004.
3. Osagie C. Aquaculture as path to thriving agriculture; 2012.
   Available:http://www.jsdafrica.com/Jsda/Vol15No2Spring2013B/PDF/Fishery%20Production%20and%20Economic%20Growth%20in%20Nigeria.Oyinbo%20Oyakhilomen.pdf.
4. Oyakhilomen O, Zibah RG. Fishery production and economic growth: Pathway for sustainable development. Journal of Sustainable Development. 2013;15:2.
5. Mjoun K, Kurt A. Tilapia: Profile and economic importance. South Dakota: South Dakota Cooperative Extension Service; 2010.
6. Job BE, Antai EE, Inyang-Etoh AP, Otogo GA, Ezekiel HS. Proximate composition and mineral contents of cultured and wild tilapia (*Oreochromis niloticus*) (Pisces: Cichlidae) (Linnaeus, 1758). Pakistan Journal of Nutrition. 2015;14(4):195-200.
7. De-Silva MPKSK. Genetic diversity of genetically improved farmed tilapia (GIFT) brood stocks in Sri Lanka. International Journal of Scientific Research and Innovation Technology. 2015;2(3):66-76.
8. Hockaday S, Beddow TA, Stone M, Hancock P, Ross LG. Using truss networks to estimate the biomass of *Oreochromis niloticus* and to investigate shape characters. Journal of Fish Biology. 2000;57:981-1000.
9. Samaradivakara SP, Hirimuthugoda NY, Gunawardana RHANM, Illeperuma RJ, Fernandopulle ND, De-Silva AD, Alexander PABD. Morphological variation of four tilapia populations in selected reservoirs in Sri Lanka. Tropical Agricultural Research. 2012;23(2):105-116.
10. Wimberger PH. Plasticity of fish body shape, the effects of diet, development, family and age in two species of geophagus (Pisces: Cichlidae). Biology Journal of Linnean Society. 1992;45:197-218.
11. Castro JA, Antonia P, Misericordia R. Mitochondrial DNA: A tool for populational

genetics studies. International Microbiology. 1988;1:327–332.
12. Arnason E. Mitochondrial cytochrome *b* DNA variation in the high-fecundity atlantic cod: trans-atlantic clines and shallow gene genealogy. Genetics. 2004;166: 1871-1885.
13. Abdel-Hamid ZG, Heba AM, El-Kader A, Aboelhassan DM, Mahrous KF. Genetic diversity in Egyptian tilapia species using PCR-RFLP of D-loop mitochondrial DNA gene. Research Journal of Pharmaceutical, Biological and Chemical Sciences. 2014;5(6):469-475.
14. Mila B, Girman DJ, Kimura M, Smith TB. Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American song bird. Proceedings of the Royal Society London Series. 2000;267(1447):1033-1040.
15. Zink RM, Barrowclough GF, Atwood JL, Blackwell-Rago RC. Genetics, taxonomy and conservation of the threatened California gnatcatcher. Conservation Biology. 2000;14(5):1394-1405.
16. Riberson A, Sotiriou E, Miaud C, Andreone F, Taberlet P. Lack of genetic diversity in *Salamandra lanzai* revealed by cytochrome b gene sequences. Copeia. 2002;20:229-232.
17. Shanker K, Ramadevi J, Choudhaury BC, Singh L, Aggarawal RK. Phylogeny of olive ridley turtles (*Lepidochelys olivacea*) on the east coast of India: implications for conservation theory. Molecular Ecology. 2004;13(7):1899-1909.
18. Menotti-Raymond M, O'Brien SJ. Dating the genetic bottleneck of the African cheetah. Proceedings of the National Academy of Science. 1993;90(8):3172-3176.
19. Farghadani R, Arman AB. Nucleotide variation of the mitochondrial *cytochrome b* gene in the Malay population. Romanian Journal of Legal Medicine. 2015;23:57-60.
20. Habib M, Lakra WS, Vindhya M, Praveen K, Barman AS, Akanksha S, Kuldeep KL, Peyush P, Asif AK. Evaluation of cytochrome b mtDNA sequences in genetic diversity studies of *Channa marulius* (Channidae: Perciformes). Mloecular Biology Reports. 2010;6(3):41-49.
21. Ahmed M, El-mezawy A. Detection of species-specific genetic markers in farm animals by RFLP analyses of cytochrome b gene. Biochemistry in Animal Husbandry. 2005;21(3-4):1-11.
22. Abdul HAA, Zakirah T, Nabilah MA, Nur AA, Muhd DDA, Wong LL, Seah YG, Tun NAMJ, Abol MAB, Awang AAK, Shahreza MS. Mitochondrial DNA diversity of terubok (*Tenualosa toli*) from Daro and Mukah, Sarawak inferred by partial cytochrome b (Cyt-b). Journal of Fisheries and Aquatic Science. 2015;10(2):92-101.
23. Brown JR, Bechenbach AT, Smith MJ. Intraspecific DNA sequence variation of the mitochondrial control region of the white sturgeon (*Acipenser transmontanus*). Molecular Biology and Evolution. 1993; 10(2):326-341.
24. Norfatimah MY, Abdulah MNS, Othman AS, Patimah I, Jamsari AFJ. Genetic variation of *Lates calcarifer* in Peninsular Malaysia based on the cytochrome b gene. Aquaculture Research. 2009;40:1742-1749.
25. Pereyra S, Garcia G, Miller P, Oviedo S, Domingo A. Low genetic diversity and population structure of narrownose shark (*Mustelus schmitti*). Fisheries Research. 2010;106:468-473.
26. Tseng MC, Shiao JC, Hung YH. Genetic identification of *Thunnusorientalis, Thunnus thynnus* and *Thunnusmaccoyii* by a cytochrome b gene analysis. Environmental Biology of Fishes. 2011;91: 103-115.
27. Chambers JK. Macdonald LE, Sarau HM, Ames RS, Freeman K, Foley JJ, Zhu Y, McLaughlin MM, Murdock P, McMillan L, Trill J, Swift A, Aiyar N, Taylor P, Vawter L, Naheed S, Szekeres, P, Hervieu G, Scott C, Watson JM, Murphy A, Duzic E, Klein C, Bergsma DJ, Wilson S, Livi P. AG protein-coupled receptor for UDP-glucose. Journal of Biological Chemistry. 2000;15: 10767–10771.
28. Harrison CJ, Langdale JA. A step by step guide to phylogeny reconstruction. The Plant Journal. 2006;45:561–572.
29. Hall TA. Bioedit: a user friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acid Symposium Series. 1999;41:95-98.
30. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution. 2013;30: 2725-2729.
31. Bandelt HJ, Forster P, Rohl A. Median-joining networks for inferring intraspecific

phylogenies. Molecular Biology and Evolution. 1999;16(1):37-48.
32. Agbebi OT, Echefu CJ, Adeosun IO, Ajibade AH, Adegbite EA, Adebambo AO, Ilori MB, Durosaro SO, Ajibike AB. Mitochondrial diversity and time divergence of commonly cultured cichlids in Nigeria. British Biotechnology Journal. 2016;13(2):1-7.
33. Lui RY, Lei CE, Liu SH, Yang GS. Genetic diversity and origin of Chinese domestic goats revealed by complete mtDNA D-loop sequence variation. Asian-Australasian Journal of Animal Science. 2007;20(2): 178-183.
34. Meyer A. Evolution of mitochondrial DNA in fishes. In: Mochachka PW, Mommsen TP. (Eds.). Biochemistry and molecular biology of fishes. New York: Elsevier Press; 1993.
35. Brumfield RT, Beerli P, Nickerson DA, Edwards SV. The utility of ingle nucleotide polymorphisms in inferences of population history. Trends in Ecology and Evolution. 2003;18:249-256.
36. Lorenc TM, Hayashi S, Stiller J, Lee H, Manoli S, Ruperao P, Visendi P, Berkman PJ, Lai K, Batley J, Edwards D. Discovery of Single Nucleotide Polymorphisms in Complex Genomes Using SGSautoSNP. Biology. 2012;1:370-382.
37. Pennings PS, Hermission J. Soft sweeps II – molecular population genetics of adaptation from recurrent mutation or migration. Molecular Biology and Evolution. 2006;23(5):1076-1084.
38. Smith JM, Haigh J. The hitch-hiking effect of a favourable gene. Genetics Research. 1974;23(1):23-35.
39. Gore MA, Chia J-M, Elshire RJ, Sun ES, Hurwitz BL, Peiffer JA, McMullen MD, Grills GS, Ross-Ibarra JW, Doreen H, Buckler ES. A first-generation haplotype map of maize. Science. 2009;326:1115-1117.
40. Barry CC, Moore PD, Ladle R. Biogeography: an ecological and evolutionary approach. 9th edition. Wiley-Blackwell: Hoboken; 2016.
41. Ekerette EE, Ikpeme EV, Udensi OU, Ozoje MO, Etukudo OM, Umoyen AJ, Durosaro SO, Wheto M. Phylogenetics and molecular divergence of tilapia fish [Oreochromis species] using mitochondrial D-loop and cytochrome b regions. American Journal of Molecular Biology. 2018;8(1):39-57.
42. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123(3): 585–95.
43. Fu YX, Li WH. Statistical tests of neutrality of mutation. Genetics. 1993;133(3):693-709.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history/26611