Mojekwu Tonna Onyekachukwu, Int. Journal of Advances in Agricultural Science & Technology,
Vol.9 Issue.10, October-2022, pg. 10-19

ISSN: 2348-1358
Impact Factor: 6.901
NAAS Rating: 3.77

Evaluation of Perfect Microsatellites in Nile Tilapia (*Oreochromis niloticus*) Genome

Mojekwu Tonna Onyekachukwu
Biotechnology Department, Nigerian Institute For Oceanography and Marine Research (NIOMR)
E-mail: tonyystone@yahoo.com

DOI: 10.47856/ijaast.2022.v09i10.002

Abstract: Microsatellites or simple sequence repeats (SSRs) consist of a sizable part of genomes and play a crucial role in the function of genes and the organization of the genome. The complete availability of a genome sequence for Nile tilapia (*Oreochromis niloticus*) provides the possibility of accomplishing a genome-wide analysis of SSRs in this species. I analyzed the abundance and density of perfect SSRs in the Nile tilapia genome and observed a sum of 252,047 microsatellites with 1–6 bp nucleotide motifs. This indicates that about 2.7 % of the Nile tilapia whole genome sequence (927.77Mb) is made up of perfect SSRs, with an average length of 135.68bp/Mb. The average density and frequency of perfect SSRs were 271.69 loci/Mb and 5834.46 bp/Mb, respectively. The six classes of perfect SSRs proportional distribution within the Nile tilapia genome were not even. Dinucleotide repeats (40.13 %) with a total count of 101145 of an average length of 26.11 bp happen to be the most abundant class of SSRs, while the percentages of mononucleotide, trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide repeats was 31.88 %, 11.98 %, 11.52 %, 4.22 %, and 0.26 %, accordingly. The various classes of SSRs repeat differ in their number of repeats with the highest being 95. My results indicate that 21 motifs contain the prevalent categories with a frequency above 1 locus/Mb: A, AAC, AAG, AATAG, AATTC, AC, AG, AGAT, AT, ATCT, ATG, ATGG, ATT, ATTT, C, CCT, CTG, CTTT, GT, GTTT.
Keywords Frequency, Genome, Microsatellite, Nile tilapia, Repeat motif

Introduction

Microsatellites are tandemly repeated DNA sequences, usually referred to as SSRs (Simple Sequence Repeats) or "STR" (short tandem repeats) of about 1-6 bp lengths per unit (Tautz and Renz 1984). Microsatellites are co-dominant, Mendelian inherited, easily typed, and very polymorphic, all properties which make them ideal genetic markers utilised for biogeography, microevolutionary pedigree analysis, population genetics studies, and genome analysis (Anane-Taabeah *et al.*, 2019; Guichoux *et al*., 2011; Tibihika *et al.*, 2020).

The Nile tilapia, *Oreochromis niloticus* (Linnaeus, 1758) is native to north, west, and central Africa and adapted to freshwater and brackish water at optimal temperatures between 28°C to 42°C (Trewavas, 1983; Teichert-Coddington *et al.*, 1997). It is the most translocated cichlid fish species across the world for aquaculture due to its economic profitability to farmers; characterised by the fast growth rate, early sexual maturity, a high degree of parental care, ability to spawn multiple broods in a season, and high fecundity associated with its large body size (Trewavas, 1983; Ojuok *et al*., 2007). However, its escape from the aquaculture

ISSN: 2348-1358
Impact Factor: 6.901
NAAS Rating: 3.77

system has resulted in hybridization concerns (Shechonge *et al*., 2018; Zengeya *et al*., 2015, 2017).

Tilapia-specific SSR markers have been used by some researchers to study the population and evolution genetics of *Oreochromis niloticus* (Anane-Taabeah *et al.*, 2019; Bezault *et al.*, 2011; Bradbeer *et al.*, 2019; Tibihika *et al*., 2018; 2020; Fuerst *et al.*, 2001; Tesfaye *et al*., 2021; Wasonga *et al*., 2017). Carleton et al. (2002) found $(CA)_n$ microsatellite in *Oreochromis niloticus* with the number of repeats ranging from four to 45 with an average of 19. These studies contributed enormous useful information for the assessment, management, and protection of *O. niloticus* as a genetic resource.

The accessibility of an absolute *O. niloticus* genome sequence has made it feasible to perform genome-wide analysis (Brawand *et al*., 2014). However, there seem to be no reports on the abundance and density of microsatellites (1–6 bp) repeats in the *O. niloticus* genome. I screened the entire *O. niloticus* genome sequence to study the density and distribution of perfect microsatellites, to facilitate the understanding of the structure of the *O. niloticus* genome, and to build up a foundation for the isolation and identification of more *O. niloticus*-specific SSRs.

Materials and Methods

Presently, there are up to five genome assemblies of *O. niloticus* from the NCBI database, however, the first reported assembly from Broad Institute, USA was used for this study (https://www.ncbi.nlm.nih.gov/data-hub/genome/?taxon=8128). The complete *Oreochromis niloticus* genome sequence (GCA_000188235.2) with a total length of 927,696,114 bp was downloaded (Brawand *et al.*, 2014) in FASTA file format to generate SSRs data. MSDB 2.4.3 (Microsatellite Search and Building Database) (http://msdb.biosv.com/) (Lianming *et al*., 2012) was used to scan the entire *O. niloticus* genome for abundance and density of perfect SSRs, using the ‘‘perfect’’ search mode with “all motifs” in the type of motif column selected. I identified six classes of microsatellites: mono-, di-, tri-, tetra-, penta-, and hexa-, nucleotide SSR motifs at a minimum repeat number of 12, 7, 5, 4, 4, and 4, respectively. The flanking sequence length was limited to 200 bp. The statistics of the Microsatellite were chosen using the ‘‘whole’’ mode, showing that the program will give rise to a single statistical Excel file for all sequence files as a whole. Repeats with unit patterns being complements were viewed as one type for the purpose of statistical analysis,. For instance, ATTT denotes AAAT in separate reading frames or on the complementary strand. The SPSS 19.0 software was used to carry out the data analysis and mapping. For ease of comparison among different repeat categories, the relative density, [SSR length (in bp) per Mb of the sequence examined], and the relative frequency, (SSR number per Mb of the sequence analyzed), were assessed.

ISSN: 2348-1358
Impact Factor: 6.901
NAAS Rating: 3.77

Results

Density and frequency of six classes of microsatellites

A sum of 252.047 SSRs was identified in the *O. niloticus* genome assembly after scanning the genome sequence for six classes of SSRs (Table 1). The total and average lengths were 5412605 bp and 135.68, respectively. The relative density and frequency were 5834.46 bp/Mb and 271.69 loci/Mb, respectively. About 2.7 % of the *O. niloticus* whole genome (927.77Mb) was occupied by the perfect SSRs. The counts, density, frequency, length, and percentages of the six classes of perfect SSRs are outlined in Table 1. Dinucleotides were the most abundant type, with the highest relative density (2846.889 bp/Mb) and frequency (109.03 loci/Mb), accounting for 40.13 % of all SSRs, followed by mono-nucleotides (31.88 %), tri-nucleotides (11.98 %), tetra-nucleotides (11.52 %), penta-nucleotides (4.22 %), and hexa-nucleotides (0.26 %).

Abundance and repeat numbers for different categories of microsatellite

Mononucleotide repeats

The leading mononucleotide repeat category was Poly (A) [or poly (T)], with 74,905 loci accounting for 93.2 % of the mononucleotide SSRs. The density, frequency, and total length, of poly (A) were 615.99 bp/Mb, 40.38 loci/Mb, and 571,450 Mb, respectively, and the average length was 15.25 bp (Table 2). Nevertheless, poly (C) [or poly (G)] accounted for only 6.78 % of the sum of mononucleotide SSRs, showing a far less abundance than poly (A) [or poly (T)]. The abundance of poly (C) was also lower (namely 2.95 loci/Mb and 39.56 bp/Mb, accordingly). The mononucleotide repeat numbers ranged from 12 to 95 times. The predominant repeat times ranged from 12 to 31, which numbered 79,701 accounting for 99.17 % of the total count of SSRs that are mononucleotide (Fig. 1A).

Dinucleotide repeats

The dinucleotide repeats are the most abundant SSRs class including AC, AG, AT, CT, and GT categories. Results showed that GT and AC have the highest frequencies (41.29loci/Mb and 40.68 loci/Mb, respectively). AT, AG, and CT had middle frequencies of 13.39, 6.85, and 6.66 loci/Mb respectively (Table 2). These categories of SSRs numbered 101,008 and accounted for 99.86 % of the total number of dinucleotide repeats. The lowest frequency of 0.15 loci/Mb and numbered 137 was observed with the GC repeat. The repeat times of dinucleotide repeats ranged from 7 to 48 times. However, the predominate repeat times ranged from 7 to 30 which numbered 100058 and accounted for 99 % of the total count of SSRs that are dinucleotide (Fig. 1B).

Trinucleotide repeats

The trimer repeats analysis including AAC, AAG, AAT, AGC, AGG, ATC, ATG, ATT, CCT, CTG, CTT, and GTT, showed that the frequencies of ATT and AAT (6.32 and 6.32 loci/Mb) were the highest. Two categories of AAC and GTT had middle frequencies that were 2.31 loci/Mb and 2.24 loci/Mb (Table 2), respectively. The lower frequencies ranged

from 1. 27 to1.94 loci/Mb. The repeat times of trinucleotide SSRs ranged between 5 and 32 times. But 5–19 repeat times were predominant and numbered 30067 and accounted for 99.5 % of the total count of trinucleotide SSRs (Fig. 1C).

Tetranucleotide repeats

Analysis of frequencies and densities of each tetrameric repeat revealed that 10 categories were predominant across the genome, and had frequencies of more than 1 loci/Mb (Table 2). The frequencies of AAAT and ATTT were highest with 3.95 loci/Mb and 3.38 loci/Mb while ATGG, ATCC, GTTT, and AAAC were at the middle range of 2.86 loci/Mb, 2.76 loci/Mb, 2.71 loci/Mb and 2.57 loci/Mb accordingly. The frequencies of the lower tetrameric repeats ATCT, AGAT, CTTT, and AAAG were 1.43 loci/Mb, 1.35 loci/Mb, 1.07 loci/Mb, and 1.01 loci/Mb. The SSRs range of tetranucleotide repeat times was between 4 and 24 times but 4–10 repeat times were predominant, numbered 26,960, and accounted for 93 % of the total count of trinucleotide SSRs (Fig. 1D).

Pentanucleotide repeats

In pentanucleotide repeats categories, AATTC, AATTG, and AATAG had a higher frequency of 1.37 loci/Mb, 1.25 loci/Mb, and 1.04 loci/Mb, and density of 32.62 bp/Mb, 29.73 bp/Mb and 29.85 bp/Mb (Table 2), respectively. The frequencies of the remaining categories were <0.63 loci/Mb. The repeat times of pentanucleotide SSRs ranged between 4 and 20 times. However, repeats ranging between 4 and 8 times were predominant, numbered 9,796, and accounted for 92 % of the total count of pentanucleotide SSRs (Fig. 1E).

Hexanucleotide repeats

The frequencies of all hexanucleotide repeat categories ranged between 0.03 loci/Mb and 0.00 loci/Mb and were lower than that of the above five types of repeats. The repeat times of hexanucleotide SSRs ranged between 4 and 15 times. Moreover, the predominate repeat times ranged between 4 and 6 and numbered 658 which is 98.80 % of the total hexanucleotide SSRs count (Fig. 1F).

Table 1 Count, length, frequency, density, and percentage of six types of perfect microsatellites in the *O. niloticus* genome sequence

Nucleotide	Total Counts	Total Length(bp)	Average Length(bp)	Frequency(loci/Mb)	Density(bp/Mb)	%
mononucleotide	80361	1216211	15.13	86.62	1311.002	31.88
dinucleotide	101145	2641048	26.11	109.03	2846.889	40.13
trinucleotide	30205	595320	19.71	32.56	641.719	11.98
tetranucleotide	29038	659180	22.70	31.3	710.556	11.52
pentanucleotide	10632	283980	26.71	11.46	306.113	4.22
hexanucleotide	666	16866	25.32	0.72	18.181	0.26
Total	**252,047**	**5,412,605**	**135.68**	**271.69**	**5834.46**	

ISSN: 2348-1358
Impact Factor: 6.901
NAAS Rating: 3.77

Table 2 Count, length, frequency, and density percentage of different categories of SSRs (frequency above 1 locus/Mb in increasing order) in the Nile tilapia genome sequence.

Motif	Total Counts	Total Length(bp)	Average Length(bp)	Frequency(loci/Mb)	Density(bp/Mb)
A	37464	571450	15,25	40,38	615,99
T	37441	571251	15,26	40,36	615,77
C	2740	36698	13,39	2,95	39,56
G	2716	36812	13,55	2,93	39,68
AC	37742	1058712	28,05	40,68	1141,23
GT	38309	1078568	28,15	41,29	1162,63
AT	12419	258786	20,84	13,39	278,96
AG	6356	123432	19,42	6,85	133,05
CT	6182	119452	19,32	6,66	128,76
ATT	5862	129342	22,06	6,32	139,42
AAT	5798	129420	22,32	6,25	139,51
AAC	2146	39336	18,33	2,31	42,4
GTT	2081	38433	18,47	2,24	41,43
CCT	1799	31155	17,32	1,94	33,58
AGG	1793	31095	17,34	1,93	33,52
ATG	1592	31059	19,51	1,72	33,48
ATC	1454	28362	19,51	1,57	30,57
AAG	1578	28413	18,01	1,7	30,63
CTT	1548	28131	18,17	1,67	30,32
CTG	1216	21006	17,27	1,31	22,64
AGC	1182	20394	17,25	1,27	21,98
ATTT	3665	64996	17,73	3,95	70,06
AAAT	3140	56096	17,86	3,38	60,47
ATGG	2649	64376	24,3	2,86	69,39
ATCC	2563	61676	24,06	2,76	66,48
GTTT	2510	48980	19,51	2,71	52,8
AAAC	2388	46284	19,38	2,57	49,89
ATCT	1329	59380	44,68	1,43	64,01
AGAT	1249	57560	46,08	1,35	62,05
CTTT	988	22568	22,84	1,07	24,33
AAAG	934	21892	23,44	1,01	23,6
AATTC	1270	30265	23,83	1,37	32,62
AATTG	1163	27580	23,71	1,25	29,73
AATAG	962	27690	28,78	1,04	29,85

ISSN: 2348-1358
Impact Factor: 6.901
NAAS Rating: 3.77

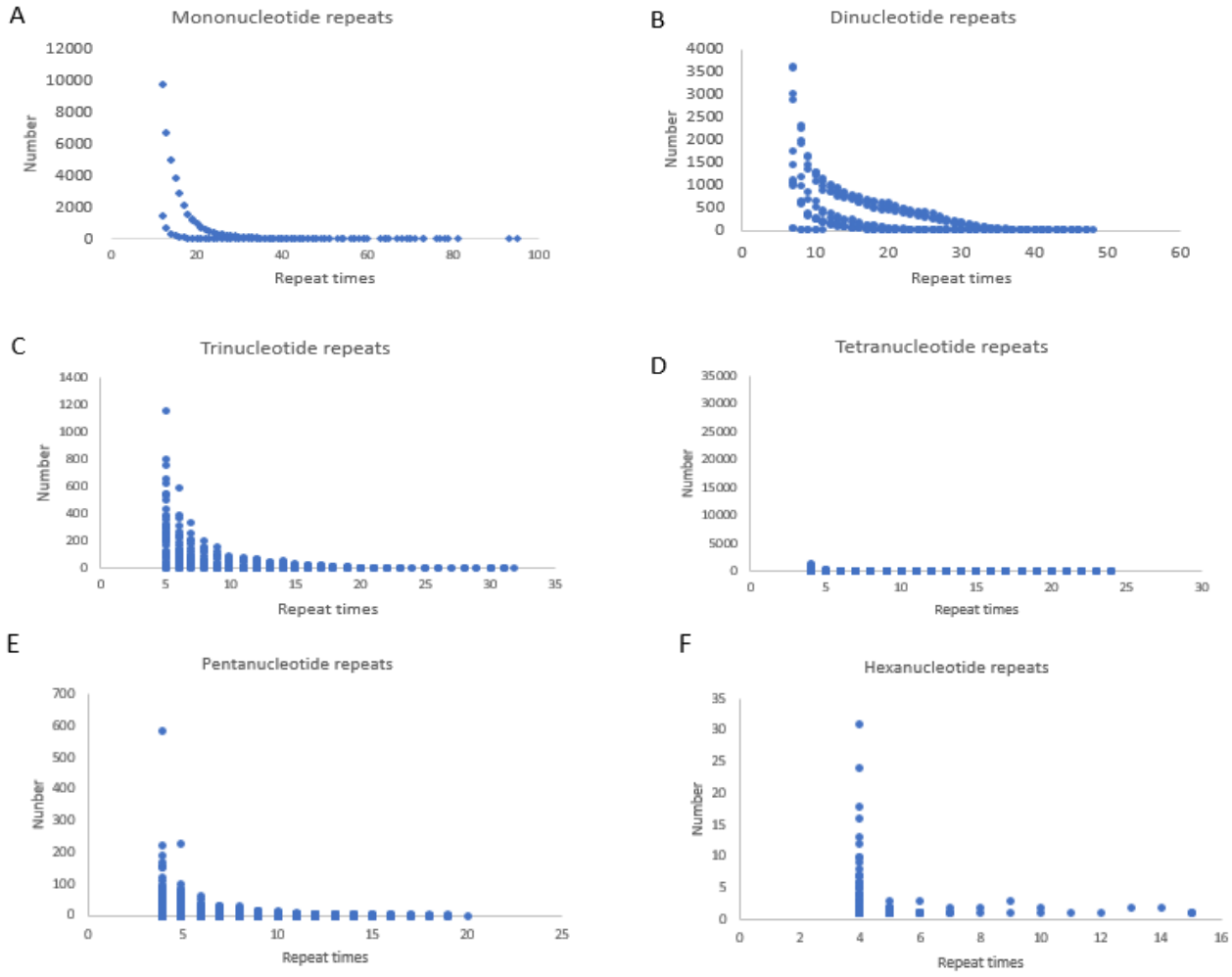

Fig. 1 Repeat times of different types of SSRs in the Nile tilapia genome.

ISSN: 2348-1358
Impact Factor: 6.901
NAAS Rating: 3.77

Discussion

An examination of perfect SSR for the entire *Oreochromis niloticus* genome sequence applying bioinformatics methodology has not been reported. We first scrutinize the abundance of perfect SSRs made up of 1–6 bp motifs in *O. niloticus* genomic sequence. In this research, about 2.7 % of the *O. niloticus* genome comprised perfect SSRs from mono- to hexa-nucleotide repeats. This percentage is similar to previous reports on the genomes of humans (3 %) (Subramanian et al., 2003), mosquitoes (2.14 %) (Yu et al., 2005), and mice (2.85 %) (Tong et al., 2006) but higher than that of yak (0.47 %) (Zhijie, 2015), cattle (0.48 %) (Qi et al., 2013), sheep (0.48 %) (Qi et al., 2013) and chicken (0.49 %) (Huang et al., 2012). The different bioinformatics software tools applied in various studies for SSRs identification, the search criteria variation and the database size could account for the percentage differences.

There was an uneven proportional distribution of the six classes of perfect SSRs in the *O. niloticus* genome. Dinucleotide repeats, accounting for the largest proportion (40.13 %) in six types of SSRs, had the highest frequency (109.03loci/Mb) and maximum density (2846.889bp/Mb), followed by mononucleotide, trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide repeats. Hexanucleotide repeats had the lowest frequency (0.72loci/Mb) and minimum density (18.181bp/Mb) (Table 1). This pattern is similar to findings in cattle, chicken, human, sheep, and yak genomes, (Subramanian *et al.,* 2003; Huang *et al.,* 2012; Qi *et al.,* 2013; Zhijie, 2015), but is different from that of drosophila, mosquito, mouse, silkworm, and zebrafish (Katti *et al.,* 2001; Li *et al.,* 2004; Yu *et al.,* 2005; Tong *et al.,* 2006' Coe, T. *et al.,* 2009). Selection for or against mono-, di-, trimers, tetra-, penta- and hexamers repeats could account for this difference observed in the abundance.

This current investigation showed that the density and number of some categories of a repeat are more than others within each type of repeat. For instance, the mononucleotide repeats, Poly (A) [or Poly (T)] showed a big over-representation, accounting for 93.2 % of the total number of mononucleotide SSRs categories. Likewise, in the other five classes of SSRs, thirteen categories including GT, AC, AAT, ATT, AAAT, ATTT, ATGG, ATCC, GTTT, AAAC, AATTC, AATTG, and AATAG in the *O. niloticus* genome were the predominant repeats, with a normal frequency higher than 1.00 loci/Mb (Table 2). The large amounts of repeats might be affected by their secondary structures and effect on DNA replication. Possibly a mutation must have occurred during SSR evolution, at the poly (A) stretches existing in the genome to produce the A-rich repeats. Furthermore, there are different repeat times for the different categories of SSRs.For instance, the mononucleotide SSR's repeat times ranged mainly between 12 and 31, the dinucleotide SSRs ranged between 7 and 30 times, trinucleotide SSRs ranged between 5 and 19 times, and the repeats ranged for tetranucleotide, pentanucleotide, and hexanucleotide SSRs were between 4–10, 4–8, and 4–6 (Fig. 1), respectively. This study is consistent with findings that nucleotide sequences with higher AT content possess more SSRs than those of higher GC content (Schlotterer 1998).

ISSN: 2348-1358
Impact Factor: 6.901
NAAS Rating: 3.77

This is consistent with our observation of higher AT-rich content of SSRs present in the *O. niloticus* genome.

Studies should be emphasized, the difference in abundance of different classes of SSRs in coding and non-coding regions of in *O. niloticus* genome (i.e. exon, intron, and intergenic regions). Moreover, Some studies showed that SSRs play an important role in the structure and function of the genome and may be associated with some diseases (Hefferon *et al.*, 2004; Campregher *et al.*, 2010).

Acknowledgements

I wish to appreciate my lovely wife and children, Kyrian Chike Anene, Roy Motloutsi, NIOMR and members of the Molecular Ecology and Evolution programme of the University of Pretoria, for the financial, moral, time, and technical support given to me.

References

[1] Anane-Taabeah, G., Frimpong, E. A., & Hallerman, E. (2019). Aquaculture-mediated invasion of the genetically improved farmed tilapia (GIFT) into the lower volta basin of Ghana. *Diversity*, *11*, 188.

[2] Bezault, E., Balaresque, P., Toguyeni, A., Fermon, Y., Araki, H., Baroiller, J. F., & Rognon, X. (2011). Spatial and temporal variation in population genetic structure of wild Nile tilapia (*Oreochromis niloticus*) across Africa. *BMC Genetics*, *12*, 102.

[3] Bradbeer, S. J., Harrington, J., Watson, H., Warraich, A., Shechonge, S., Smith, A., Tamatamah, R., Ngatunga, B. P., Turner., G. F., & Genner, M. J. (2019). Limited hybridization between introduced and critically endangered indigenous tilapia fishes in northern Tanzania. *Hydrobiologia, 832*, 257-268. https://doi.org/10.1007/s10750-018-3572-5.

[4] Brawand, D., Wagner, C. E., Li, Y. I., … Di Palma, F. (2014).The genomic substrate for adaptive radiation in African cichlid fish. *Nature*. 513-375. https://doi.org/10.1038/nature13726

[5] Campregher, C., Scharl, T., Nemeth, M., Honeder, C., Jascur, T., Boland, C.R., & Gasche, C. (2010) The nucleotide composition of microsatellites impacts both replication fidelity and mismatch repair in human colorectal cells. *Human Molecular Genetics*, *19*, 2648–2657.

[6] Carleton, K. L., Streelman, J. T., Lee, B.-Y., Garnhart, N., Kidd, M., & Kocher, T. D. (2002). Rapid isolation of CA microsatellites from the tilapia genome. *Animal Genetics, 33(2),* 140–144.

[7] Coe, T. S., Hamilton, P. B., Griffiths, A. M., Hodgson, D. J., Wahab, M. A., & Tyler, C. R. (2009). Genetic variation in strains of zebrafish (*Danio rerio*) and the implications for ecotoxicology studies. Ecotoxicology (London, England) 18, 144–150.

[8] Fuerst, P., Mwanja, W., & Kaufman, L. (2000). The genetic history of the introduced Nile tilapia of Lake Victoria (Uganda–E. Africa): the population structure of *Oreochromis niloticus* (Pisces: Cichlidae) revealed by DNA microsatellite markers. In: Tilapia Aquaculture in the 21st Century Proceedings from the Fifth International Symposium on Tilapia in Aquaculture; 2000.

[9] Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Le´ger, P., Lepais, O., Lepoittevin, C., Malausa, T., Revardel, E., Salin, F., & Petit, R. J. (2011). Current trends in microsatellite genotyping. *Molecular Ecology Resource, 11*, 591–611.

[10] Hefferon, T.W., Groman, J. D., Yurk, C. E., & Cutting, G. R. (2004). A variable dinucleotide repeat in the CFTR gene contributes to phenotype diversity by forming RNA secondary structures that alter splicing. Proceedings of the National Academy of Sciences of the United States of America, 101, 3504–3509.

ISSN: 2348-1358
Impact Factor: 6.901
NAAS Rating: 3.77

[11] Huang, J., Du., L. M., Li, Y. Z., Li, W, J., Zhang, X.Y., & Yue, B. S. (2012). Distribution regularities of microsatellites in the Gallus gallus genome. *Sichuan Journal of Zoology, 31*, 358–363 (in Chinese with English abstract).

[12] Katti, M. V., Ranjekar, P. K., Gupta, V. S. (2001). Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Molecular Biology of Evolution, 18*,1161–1167.

[13] Li, B., Xia, Q. Y., Lu, C., Zhou, Z.Y., & Xiang, Z. H. (2004) Analysis on frequency and density of microsatellites in the coding sequence of several eukaryotic genomes. *Genomics, Proteomics Bioinformatics, 2*, 24–31.

[14] Lianming, D., Yuzhi, L., Xiuyue, Z., & Bisong, Y. (2012). MSDB: A user-friendly program for reporting distribution and building databases of microsatellites from genome sequences. *Journal of Heredity, 104(1),*154–157 doi:10.1093/jhered/ess082 [Downloaded may 4, 2017]

[15] Ojuok, J. E., Njiru, M., Ntiba, M. J., & Mavuti, K. M. (2007). The effect of overfishing on the life-history strategies of Nile tilapia, *Oreochromis niloticus* (L.) in the Nyanza Gulf of Lake Victoria, Kenya. *Aquatic Ecosystem Health & Management, 10*, 443-448.

[16] Qi, W. H., Jiang, X. M., Xiao, G. S., Huang, X. Y., & Du, L. M. (2013). Seeking and bioinformatics analysis of microsatellite sequence in the genomes of cow and sheep. *Acta Veterinaria et Zootechnica Sinica, 44*,1724–1733 (in Chinese with English abstract)

[17] Schlotterer, C. (1998). Genome evolution: are microsatellites really simple sequences? Current Biology, 8,132–134

[18] Shechonge, A., Ngatunga, B. P., Tamatamah, R., Bradbeer, S. J., Harrington, J., Ford, A. G. P., Turner, G. F., & Genner, M. J. (2018). Losing cichlid fish biodiversity: genetic and morphological homogenization of tilapia following colonization by introduced species. *Conservation Genetics, 19*, 1199–1299.

[19] Subramanian, S., Mishra, R. K., & Singh, L. (2003). Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biology, 4*, R13.

[20] Tautz, D., & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research, 12*, 4127–4138.

[21] Teichert-Coddington, D. R., Popma, T. J., & Lovshin, L. L. (1997). Attributes of tropical pond-cultured fish, pgs 183-198. In: Enga H. S. & Boyd, C. E.(eds.), Dynamics of pond Aquaculture. *CRC Press*, USA: Boca Raton, Florida.

[22] Tesfaye, G., Curto, M., Meulenbroek, P., Englmaier, G. K., Tibihika, P. D., Alemayehu, E., Getahun, A., & Meimberg, H. (2021). Genetic diversity of Nile tilapia (*Oreochromis niloticus*) populations in Ethiopia: insights from nuclear DNA microsatellites and implications for conservation. *BMC Ecology and Evolution, 21***,** 113. https://doi.org/10.1186/s12862-021-01829-2.

[23] Tibihika, P. D., Curto, M., Alemayehu, E., Waidbacher, H., Masembe, C., Akoll, P., & Meimberg, H. (2020). Molecular genetic diversity and differentiation of Nile tilapia (*Oreochromis niloticus*, L. 1758) in east African natural and stocked populations. *BMC Evolutionary Biology*, *20*, 16. https://doi.org/10.1186/s12862-020-1583-0.

[24] Tibihika, P. D., Curto, M., Dornstauder-Schrammel, E., Winter, S., Alemayehu, E., Waidbacher, H., & Meimberg, H. (2018). Application of microsatellite genotyping by sequencing (SSR-GBS) to measure the genetic diversity of the East African Oreochromis niloticus. Conservation Genetics, 20, 357–372. https://doi.org/10.1007/s10592-018-1136-x

[25] Trewavas, E. (1983). Tilapiine fishes of the genera Sarotherodon, Oreochromis, and Danakilia (Vol 9, p. 1-583). London, UK: British Museum (Natural History). https://doi.org/10.5962/bhl.title.123198

[26] Tong, X. L., Dai, F. Y, Li, B., Yu, Q. Y., Xia, Q. Y., & Lu, C. (2006). Microsatellite repeats in the mouse: abundance, distribution, and density. *Current Zoology, 52*, 138–152.

[27] Https://www.ncbi.nlm.nih.gov/data-hub/genome/?taxon=8128.

[28] Yu, Q.Y., Li, B., Li, G. R., Fang, S. M., Yan, H., Tong, X. L., Qian, J. F., Xia, Q. Y., & Lu, C. (2005) Abundance and distribution of microsatellite in the entire mosquito. *Prog Biochem Biophys, 32*, 435–441.

[29] Wasonga, A. G., Daniel, W. A., & Brian, O. (2017). Interspecific hybridization of Tilapiines in Lake Victoria, Kenya. *Journal of Fisheries Livestock Production, 5*, 235 DOI: 10.4172/2332-2608.1000235.

ISSN: 2348-1358
Impact Factor: 6.901
NAAS Rating: 3.77

[30] Zengeya, T. A., Booth, A. J., & Chimimba, C. T. (2015). Broad Niche Overlap between Invasive Nile Tilapia *Oreochromis niloticus* and Indigenous Congenerics in southern Africa: Should We be Concerned? *Entropy, 17*, 4959 – 4973.

[31] Zengeya, T., Ivey, P., Woodford, D. J., Weyl, O., Novoa, A., Shackleton, R., Richardson, D., & Van Wilgen, R. (2017). "Managing conflict-generating invasive species in South Africa: Challenges and trade-offs", *Bothalia, 47*, no. 2. http://dx.doi.org/10.4102/abc.v47i2.2160

[32] Zhijie, Ma (2015). Genome-wide characterization of perfect microsatellites in yak (Bos grunniens). *Genetica, 143*, 515–520. DOI 10.1007/s10709-015-9849-y.

Authors Biography

I obtained Ph.D. in Genetics from the Department of Biochemistry, Genetics, and Microbiology of the University of Pretoria, South Africa. My specialties as a Molecular Biologist include Genomics (gene capture based on SNPs), Molecular Ecology, Evolution, Conservation, and invasive species detection (introgressive hybridization). My main research interest is developing a gene capture panel from the available Cichlids genome based on SNPs for monitoring the introgressive hybridization of invasive species, a valuable resource to both fish breeders and conservation biologists in detecting hybrids. Other interests include improving the biodiversity knowledge of the *Oreochromis* genus using a DNA barcoding approach and writing research grant proposals. I know Bayesian networks, Gene marker, Galaxy, Linux, Mega, MS word packages, Excel, R, Multi-dimensional methods, and SPSS. I know evolutionary and phylogenetic tools like BEAST, Fig Tree, MEGA, STRUCTURE, Tracer, and methods such as Bayesian phylogenetic inference, parsimony, UPGMA, and distance matrix.

