ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12423

Evaluation of the genetic diversity of microsatellite markers among four strains of *Oreochromis niloticus*

M. A. D. Dias*, R. T. F. de Freitas*, S. E. Arranz†, G. V. Villanova† and A. W. S. Hilsdorf‡

*Depto. de Zootecnia, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil. †Facultad de Ciencias Bioquímicas y Farmaceuticas, Instituto de Biología Molecular y Celular de Rosario (IBR-CONICET) and Area Biología, Universidad Nacional de Rosario, Rosario, Argentina. ‡Núcleo Integrado de Biotecnologia, Universidade de Mogi das Cruzes, Mogi das Cruzes, São Paulo, Brazil.

Summary

Different strains of Nile tilapia can be found worldwide. To successfully use them in breeding programs, they must be genetically characterized. In this study, four strains of Nile tilapia – UFLA, GIFT, Chitralada and Red-Stirling – were genetically characterized using 10 noncoding microsatellite loci and two microsatellites located in the promoter and first intron of the *growth hormone* gene (*GH*). The two microsatellites in the *GH* gene were identified at positions −693 to −679 in the promoter [motif $(ATTCT)_8$] and in intron 1 at positions +140 to +168 [motif $(CTGT)_7$]. Genetic diversity was measured as mean numbers of alleles and expected heterozygosity, which were 4 and 0.60 (GIFT), 3.5 and 0.71 (UFLA), 4.5 and 0.57 (Chitralada) and 2.5 and 0.42 (Red-Stirling) respectively. Genetic differentiation was estimated both separately and in combination for noncoding and *GH* microsatellites markers using Jost's D_{EST} index. The UFLA and GIFT strains were the least genetically divergent ($D_{EST} = 0.10$), and Chitralada and Red-Stirling were the most ($D_{EST} = 0.58$). The UFLA strain was genetically characterized for the first time and, because of its unique origin and genetic distinctness, may prove to be an important resource for genetic improvement of Nile tilapia. This study shows that polymorphisms found in coding gene regions might be useful for assessing genetic differentiation among strains.

Keywords genetic variability, growth hormone, Nile tilapia, population genetics, STR, tilapia strains

Introduction

Tilapias, *Oreochromis* ssp., comprise the second-most cultivated group of fishes in the world. According to estimates by the United Nations Food and Agriculture Organization (FAO), tilapia production will increase by 30% by 2030 (World Bank, 2013). In Brazil, Nile tilapia (*O. niloticus*) is the most cultivated species in inland aquaculture with an annual production of 169 306 tons in 2013, representing 43% of aquaculture production (IBGE, 2014).

Genetic variability is an important parameter to be considered in animal genetic breeding programs. The screening of molecular markers has been used to genetically characterize strains and establish crossbreeding schemes in many terrestrial and aquatic farm animals (Melo *et al.* 2006; Moreira *et al.* 2007; Briñez *et al.* 2011; Davids *et al.* 2012). Microsatellites, or short tandem repeats (STRs), are among the most informative markers because of their codominant and highly polymorphic natures (O'Connell & Wright 1997). These markers are an important tool for identifying levels of genetic differences among strains and for monitoring inbreeding levels in broodstocks (Romana-Eguia *et al.* 2004; Melo *et al.* 2008). The STRs used in such assessments are generally considered to be selectively neutral (Oliveira *et al.* 2006).

Monitoring the presence of STRs within genes related to economically important phenotypes is another approach for assessing associations of STR polymorphisms with performance in farming systems (De-Santis & Jerry 2007; Mojekwu & Anumud 2013). Such associations should be evaluated in genetic improvement programs (Dahm & Geisler 2006; Mojekwu & Anumud 2013).

The *growth hormone* gene (*GH*) is responsible for regulating somatic growth and anabolic processes, such as cell division and muscle growth (De-Santis & Jerry 2007), and also plays a role in osmoregulation (Sakamoto & McCormick, 2006). Among fish, the presence of an STR in *GH* has

Address for correspondence

A. W. S. Hilsdorf, Universidade de Mogi das Cruzes, Núcleo Integrado de Biotecnologia, Av. Dr. Cândido Xavier de Almeida Souza, 200, 08780-911, Mogi das Cruzes, SP, Brazil.
E-mail: wagner@umc.br

Accepted for publication 04 January 2016

© 2016 Stichting International Foundation for Animal Genetics

been described in different species, such as *Lates calcarifer* (Yue *et al.* 2001), *Sparus aurata* (Almuly *et al.* 2005) and *Dicentrarchus labrax* (Quéré *et al.* 2010). In tilapia, *GH* is present in two forms: *GH1* (Ber & Daniel 1992) and *GH2* (Ber & Daniel 1993), and a STR was detected within *GH2* (Yue & Orban 2002).

Considering the importance of *GH* for growth metabolism and to better quantify the genetic variability of tilapia strains, this study aimed to (i) establish an *in silico* database of the tilapia *GH1* sequence for the presence of STR markers, (ii) screen microsatellite polymorphisms in the promoter and intron 1 region of the *GH1* gene in different strains of tilapia and (iii) characterize and evaluate genetic variability within and among four tilapia strains using noncoding STR markers and an STR in the *GH* gene.

Material and methods

Sources of genetic material and DNA extraction

A total of 96 individuals of the following four strains of *O. niloticus* were evaluated in this study. The 26 individuals of the Thai Chitralada strain (CHIT), originating from the Asian Institute of Technology, were randomly sampled from a larger group of broodstock held at Indústria Brasileira do Peixe, Ltda. (IBP), Brazil. The Red-Stirling strain (REDS), originating from the University of Stirling (Scotland), was imported to Brazil in 2001; the 20 individuals used herein were also randomly sampled from a larger broodstock kept in the IBP hatchery. The GIFT strain, originating from the Genetic Improvement of Farmed Tilapias (GIFT) project, was brought to Brazil in 2005 (Lupchinski-Júnior *et al.* 2008); the 25 GIFT individuals were randomly taken from two families maintained at the Federal University of Lavras (UFLA), Brazil. The UFLA strain is a local strain, originating from the first Nile tilapia importation to Brazil from the Ivory Coast, Africa, in 1971 (Nugent 1988) and transferred to UFLA aquaculture facilities in 1977. Since then, the stock has been under a mass selection breeding program (A.J.L. Almeida, UFLA, Brazil, personal communication).

Fin clips were sampled, and total genomic DNA was extracted using the methods described by Taggart *et al.* (1992), except that STE buffer (0.1 M NaCl, 0.05 M Tris-HCl and 0.01 M EDTA) was made with a lower concentration of EDTA. The DNA concentration and quality were assessed by electrophoresis through a 1% agarose gel, in which 5 μl of genomic DNA solution was applied and a lambda DNA/HindIII (Invitrogen) molecular weight marker was used. Finally, DNA was diluted to 100 ng/μl and stored at −20 °C.

STR marker identification, amplification and genotyping

First, the TANDEM REPEATS FINDER program version 4.07b (Benson 1999) was used to identify the STR region in the tilapia *GH1* gene (GenBank accession no. M97766-1, using default parameters). Then, primers were designed to target the STR flanking regions in the *GH1* gene promoter and intron 1 sequences using the PRIMER3 PLUS program (Untergasser *et al.* 2007). An alignment of genomic DNA sequences was performed for the *GH1* and *GH2* tilapia genes using the MEGA5 program (Tamura *et al.* 2011) to check for sequence homology with the primers. A total of 10 unlinked noncoding STR loci (*UNH828, UNH829, UNH009, UNH005, UNH103, UNH104, UNH123, UNH203, UNH866* and *GM672*) were selected from the linkage map produced by Lee & Kocher (1996). Primer sequences are shown in Table S1.

PCRs were performed in a 20-μl reaction volume containing 50 ng of template DNA, 2 μM of each primer, 0.25 mM of dNTPs, 0.5 U of Taq DNA polymerase, 1 μl of 10× reaction buffer and 1.5 mM of $MgCl_2$. Samples for all loci were subjected to an initial denaturation step at 94 °C for 3 min, followed by 35 cycles of 94 °C for 40 s, an appropriate annealing temperature for 30 s and 72 °C for 1 min and followed by a final extension at 72 °C for 10 min. The annealing temperatures used for each locus were as follows: 54 °C for *UNH 005, UNH 009, UNH 104* and *UNH 123*; 56 °C for *UNH 103* and *UNH 203*; 58 °C for *UNH 828, UNH 829, UNH 866* and *GM 672*; and 60 °C for STR promoter and intron.

PCR was performed as described by Schuelke (2000) using an IRDye®700-labeled universal M13 primer. PCR products were genotyped in a Matrix Plus gel of denaturing polyacrylamide at 6.5%, and DNA was detected using a Li-Cor DNA Analyzer 4300 (IR2) automated sequencer. Allele sizes were estimated using a Saga^{GT} Client (Li-color Biosciences) by comparing the samples to molecular weight standards (50–350 bp ladder) on the Li-Cor DNA Analyzer 4300 (IR2).

Data analysis

To assess the kinship among individuals within strain, we compared multilocus genotypes between all pairs of tagged fish using COANCESTRY version 1.0.0 (Wang 2011). We used the triadic likelihood estimator TrioML, which shows the lowest variation needed to identify occurrence of sibling–sibling or parent–offspring pairs in the dataset. The program CERVUS 3.0 (Kalinowski *et al.* 2007) was used to estimate the number of alleles per locus (A), the polymorphic information content (PIC), the observed (H_o) and expected (H_e) heterozygosities, Hardy–Weinberg equilibrium (HWE) probability (P_{HW}) and deficit of intrapopulation heterozygosity (F_{is}) and also to detect the presence of null alleles per locus and strain. Linkage disequilibrium was examined using the log likelihood ratio statistic (G-test) within GENEPOP version 4.3 (Rousset 2008) with 10 000 dememorizations, 100 batches and 5000 iterations per batch. The D_{EST} genetic differentiation index (Jost 2008) among strains was

© 2016 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12423

estimated using the DEMETICS package implemented in the R software package (R Development Core Team 2010). A UPGMA dendrogram was constructed based on D_{EST} genetic distances using the hclust function of the MASS pack of the R software package (R Development Core Team 2010).

Results

Relatedness structure within strains

The average relatedness TrioML estimators (and variances) obtained were 0.077 (0.013) for CHIT, 0.1412 (0.044) for GIFT, 0.120 (0.023) for REDS and 0.126 (0.033) for UFLA strains (Figs S1–S4). The relatedness estimator values showed low overall levels of kinship among individuals within the four strains, indicating low frequencies of sibling–sibling or parent–offspring pairs in the dataset ($r \geq 0.25$).

In silico assessment of STRs in *GH1*

The complete sequence of the *GH1* gene deposited in GenBank (accession no. M97766-1; Ber & Daniel 1993) is 3486 bp with promoter-binding elements CAAT box and TATA box, exons 1–6, introns 1–5 and a regulatory polyA signal-sequence tail. In the current study, 22 tandem repeat sequences were found and evaluated for copy number, consensus size and alignments with low entropy scores. From these sequences, two STR tracts were identified at positions −693 to −679 (in the promoter region, known as the STR promoter, with motif ATTCT) and +140 to +168 (in the intron 1 region, known as the STR intron, with motif CTGT). Primer sequences were designed specifically for the STR promoter and for the STR intron candidate markers (Table S1). The STR promoter primers were specific to the *GH1* gene, whereas the primers designed for the STR intron annealed to both *O. niloticus GH* genes in this region.

Allelic diversity and Hardy–Weinberg equilibrium in *GH1* and noncoding STR loci

We screened for variability of GH-associated and noncoding microsatellite DNA loci in samples from four major *O. niloticus* strains. Genetic variability at the two STR loci in the *GH1* gene of the four strains of Nile tilapia is shown in Table 1. Analysis of the two STR loci of *GH1* revealed a total of nine alleles, with a higher level of polymorphism in the UFLA, GIFT and CHIT strains. The presence of STR promoter allele 191 was identified only in the UFLA and CHIT strains, and alleles 186 and 206 were not observed in the REDS strain. Among the three alleles found in the STR intron, allele 198 was found only in the CHIT strain at a frequency of 0.08%. Tests showed no evidence of scoring errors attributable to null alleles within loci and strains. The average STR promoter PIC for the CHIT, GIFT and UFLA strains (0.71) was greater than that for the REDS strain (0.35) (Table 1). In the STR intron, the total average PIC was 0.38 and only slight variation was observed among strains.

The average observed and expected heterozygosities per strain at the STR promoter locus ranged from 0.32 to 0.82 and from 0.39 to 0.90 respectively (Table 2). Significant departures from HWE were not observed at the STR promoter among strains except for the CHIT strain, in which an excess of heterozygotes was observed ($F_{is} = -0.22$). Conversely, in the STR intron, the UFLA and GIFT strains showed significant departures from HWE with a significant homozygote excess observed ($F_{is} = 0.71$ and 0.42 respectively). Inbreeding was apparent within the UFLA strain ($F_{is} = 0.30$) followed by the GIFT ($F_{is} = 0.24$) and REDS ($F_{is} = 0.11$) strains. The CHIT strain did not show significant inbreeding levels for the STR intron.

Genetic variability was estimated for the 10 noncoding STR loci among the four strains of Nile tilapia (Tables 3 and 4). A total of 88 alleles with high rates of polymorphism

Table 1 Summary of allelic variation at two STR loci within the *GH1* gene in *Oreochromis niloticus*: number of alleles, number of private alleles (in brackets), polymorphic information content (superscript) and total number of alleles per marker in all strains. Alleles and frequencies within the respective strains are shown in the body of the table.

Locus and alleles	GIFT (n = 25)	UFLA (n = 20)	CHIT (n = 26)	REDS (n = 20)	Total
STR promoter	6 (0)$^{0.77}$	5 (0)$^{0.70}$	6 (0)$^{0.56}$	3 (0)$^{0.35}$	6
206	0.04	0.08	0.07	0.00	–
201	0.16	0.15	0.02	0.18	–
196	0.38	0.33	0.61	0.76	–
191	0.00	0.12	0.14	0.00	–
186	0.22	0.17	0.05	0.00	–
181	0.20	0.15	0.11	0.06	–
STR intron	2 (0)$^{0.37}$	2 (0)$^{0.36}$	3 (1)$^{0.44}$	2 (0)$^{0.35}$	3
206	0.38	0.52	0.08	0.34	–
202	0.62	0.48	0.29	0.66	–
198	0.00	0.00	0.63	0.00	–
Average	4 (0)$^{0.57}$	3.5 (0)$^{0.53}$	4.5 (0.5)$^{0.50}$	2.5 (0)$^{0.35}$	4.5

GIFT, Genetic Improvement of Farmed Tilapias strain; UFLA, Universidade Federal de Lavras strain; CHIT, Chitralada strain; REDS, Red-Stirling strain.

© 2016 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12423

Table 2 Summary statistics for genetic diversity at STR loci within the *GH1* gene in four Nile tilapia strains: number of analyzed individuals (n); observed (H_o) and expected heterozygosity (H_e); probability values of concordance with Hardy–Weinberg expectations (P_{HW}); inbreeding coefficient (F_{is}).

Locus	GIFT (n = 25)	UFLA (n = 20)	CHIT (n = 26)	REDS (n = 25)
STR promotor				
H_o	0.76	0.82	0.73	0.32
H_e	0.72	0.90	0.60	0.39
P_{HW}	0.42	0.26	0.05**	0.20
F_{is}	0.05	−0.10	−0.22	0.19
STR intron				
H_o	0.28	0.15	0.58	0.44
H_e	0.48	0.51	0.53	0.46
P_{HW}	0.08*	0.00**	0.35	0.59
F_{is}	0.42	0.71	−0.11	0.04
Average (SE)				
H_o	0.52 (0.34)	0.48 (0.47)	0.66 (0.10)	0.38 (0.08)
H_e	0.60 (0.17)	0.75 (0.27)	0.56 (0.05)	0.43 (0.05)
F_{is}	0.24 (0.26)	0.30 (0.57)	-0.17 (0.07)	0.11 (0.10)

*$P < 0.05$; **$P < 0.01$.
GIFT, Genetic Improvement of Farmed Tilapias strain; UFLA, Universidade Federal de Lavras strain; CHIT, Chitralada strain; REDS, Red-Stirling strain; SE, standard error.

among the strains were identified using the 10 noncoding STR loci, except for the *UNH-005* locus (as allele 159 was fixed in the UFLA and GIFT strains) and the *UNH-866* locus (as allele 167 was fixed in the REDS strain and as private alleles *171* and *177* were found in the UFLA and CHIT strains respectively).

The CHIT strain showed the highest mean number of alleles ($n = 7.5$) followed by the GIFT ($n = 5.9$), UFLA ($n = 5.6$) and REDS ($n = 3.9$) strains. Deviation from expected heterozygosity relative to that observed (0.20) was higher in the UFLA strain, which suggests a high level of inbreeding (mean $F_{is} = 0.35$). The GIFT strain showed the highest average expected heterozygosity (0.72), whereas the REDS strain showed the lowest (0.62). Regarding the presence of private alleles, the CHIT strain showed the highest frequency, with nine private alleles distributed among seven of the 10 assessed loci. Additionally, the UFLA strain had six private alleles (at the *UNH-103*, *UNH-123*, *UNH-203* and *GM-672* loci), and the REDS strain had five private alleles (in the *UNH-828*, *UNH-829*, *UNH-005* and *UNH-009* loci). The GIFT strain had a private allele, only at the *UNH-123* locus.

The overall average PIC value for noncoding microsatellites was 0.62, which indicates that the loci screened were statistically informative for estimating genetic variability both within and among strains. The locus with the highest average PIC was *UNH-103* (0.79), followed by *UNH-828* (0.77), *UNH-009* (0.74), *UNH-829* (0.72), *UNH-203* (0.71), *UNH-123* (0.70), *GM-672* (0.66), *UNH-104* (0.61), *UNH-005* (0.30) and *UNH-866* (0.13).

Significant departures from HWE ($P < 0.05$) were detected in 17 of 37 tests. Only the *UNH-009* locus showed no departure from HWE among the four strains. F_{is} values suggest inbreeding across the four strains: The UFLA strain was the most inbred strain ($F_{is} = 0.35$), followed by the GIFT ($F_{is} = 0.21$), CHIT ($F_{is} = 0.15$) and REDS ($F_{is} = 0.19$) strains. Linkage disequilibrium tests for the *GH*-associated loci yielded significant values for GIFT (0.015), REDS (0.003) and UFLA (0.0007) and a non-significant value for CHIT (0.405). This result indicates that there is non-random association of allelic variation in the GIFT, REDs and UFLA strains and, for CHIT, that the two loci segregate independently.

Genetic divergence among the four tilapia strains

The amount of differentiation among strains is shown in Table 5. There was no significant difference between the UFLA and GIFT strains when considering only the STR loci present in the *GH1* gene. However, we were able to detect differences at these loci in comparisons with the other strains. For noncoding STR loci, all comparisons among the four strains showed highly significant levels of genetic differentiation.

Table 3 Summary of genetic variation at 10 noncoding STR loci in *Oreochromis niloticus*: number of alleles, number of private alleles (in brackets), polymorphic information content (superscript) and total number of alleles per marker in all strains.

Locus	GIFT (n = 25)	UFLA (n = 20)	CHIT (n = 26)	REDS (n = 20)	Total
UNH-828	7 (0)$^{0.79}$	9 (0)$^{0.75}$	10 (0)$^{0.79}$	4 (1)$^{0.55}$	10
UNH-829	5 (0)$^{0.66}$	5 (0)$^{0.75}$	7 (0)$^{0.68}$	7 (2)$^{0.78}$	7
UNH-005	1 (0)^{0}	1 (0)^{0}	2 (1)$^{0.23}$	2 (1)$^{0.37}$	2
UNH-009	6 (0)$^{0.73}$	6 (0)$^{0.80}$	7 (2)$^{0.75}$	5 (2)$^{0.57}$	9
UNH-103	11 (0)$^{0.85}$	11 (2)$^{0.85}$	10 (1)$^{0.72}$	6 (0)$^{0.63}$	14
UNH-104	4 (0)$^{0.56}$	5 (0)$^{0.67}$	6 (2)$^{0.74}$	2 (0)$^{0.37}$	7
UNH-123	10 (1)$^{0.64}$	7 (2)$^{0.76}$	11 (1)$^{0.85}$	6 (0)$^{0.60}$	15
UNH-203	8 (0)$^{0.56}$	4 (1)$^{0.76}$	11 (1)$^{0.87}$	3 (0)$^{0.43}$	12
UNH-866	2 (0)$^{0.11}$	2 (0)$^{0.13}$	4 (1)$^{0.21}$	1 (0)^{0}	4
GM672	5 (0)$^{0.67}$	6 (1)$^{0.65}$	7 (0)$^{0.77}$	3 (0)$^{0.49}$	8
Average	5.9 (0.1)$^{0.56}$	5.6 (0.6)$^{0.61}$	7.5 (0.9)$^{0.66}$	4.1 (0.6)$^{0.48}$	8.8

GIFT, Genetic Improvement of Farmed Tilapias strain; UFLA, Universidade Federal de Lavras strain; CHIT, Chitralada strain; REDS, Red-Stirling strain.

© 2016 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12423

Table 4 Summary statistics of genetic diversity at 10 noncoding STR loci in four Nile tilapia strains: number of analyzed individuals (n); observed (H_o) and expected heterozygosity (H_e); probability values of concordance with Hardy–Weinberg expectations (P_{HW}); inbreeding coefficient (F_{is}).

Locus	GIFT (n = 25)	UFLA (n = 20)	CHIT (n = 26)	REDS (n = 25)
UNH-828				
H_o	0.72	0.95	0.89	0.60
H_e	0.83	0.80	0.89	0.63
P_{HW}	0.11**	0.06**	0.25	0.47
F_{is}	0.13	−0.19.	−0.08	0.05
UNH-829				
H_o	0.64	0.55	0.88	0.88
H_e	0.72	0.81	0.74	0.83
P_{HW}	0.21	0.01**	0.03**	0.33
F_{is}	0.12	0.32	−0.20	−0.07
UNH-005				
H_o	–	–	0.15	0.36
H_e	–	–	0.15	0.51
P_{HW}	–	–	0.94	0.14
F_{is}	–	–	−0.06	0.29
UNH-009				
H_o	0.72	0.750	0.68	0.60
H_e	0.78	0.85	0.79	0.63
P_{HW}	0.29	0.19	0.15	0.43
F_{is}	0.08	0.12	0.14	0.05
UNH-103				
H_o	0.92	0.70	0.65	0.56
H_e	0.88	0.86	0.76	0.67
P_{HW}	0.41	0.01**	0.11	0.09**
F_{is}	−0.05	0.21	0.14	0.17
UNH-104				
H_o	0.28	0.50	0.65	0.17
H_e	0.62	0.73	0.79	0.51
P_{HW}	0.00**	0.02**	0.07	0.00**
F_{is}	0.55	0.32	0.18	0.67
UNH-123				
H_o	0.24	0.45	0.63	0.60
H_e	0.70	0.81	0.88	0.67
P_{HW}	0.00**	0.00**	0.00**	0.28
F_{is}	0.66	0.44	0.28	0.10
UNH-203				
H_o	0.36	0.65	0.40	0.36
H_e	0.62	0.80	0.90	0.54
P_{HW}	0.00**	0.06	0.00**	0.04**
F_{is}	0.42	0.19	0.56	0.34
UNH-866				
H_o	0.04	0.15	0.12	–
H_e	0.12	0.14	0.22	–
P_{HW}	0.06	0.92	0.01**	–
F_{is}	0.66	−0.06	0.47	–
GM-672				
H_o	0.28	0.30	0.73	0.52
H_e	0.73	0.71	0.82	0.58
P_{HW}	0.00**	0.00**	0.36	0.31
F_{is}	0.62	0.58	0.11	0.11
Average (SE)				
H_o	0.47 (0.29)	0.56 (0.25)	0.58 (0.27)	0.52 (0.20)
H_e	0.67 (0.22)	0.72 (0.24)	0.69 (0.27)	0.62 (0.10)
F_{is}	0.35 (0.28)	0.21 (0.24)	0.15 (0.23)	0.19 (0.22)

*P < 0.05; **P < 0.01.

GIFT, Genetic Improvement of Farmed Tilapias strain; UFLA, Universidade Federal de Lavras strain; CHIT, Chitralada strain; REDS, Red-Stirling strain; SE, standard error.

Finally, STRs present at the *GH1* and noncoding STR loci were evaluated in combination. This approach improved our estimates of genetic divergence among the strains (Table 5 and Fig. 1). The UFLA and GIFT strains were least divergent followed by the CHIT and REDS strains.

Discussion

STR polymorphisms in the *GH1* gene

Given the importance of GH in metabolism and its association with growth performance, this study aimed to identify STRs in the promoter region and intron 1 of the *GH1* tilapia gene, which to date has not been described in the literature. The STR promoter and STR intron 1 loci found in the *GH1* gene were polymorphic both within and among the four tilapia strains. Differences observed in population parameters for these loci could have been caused by responses related to selection programs that used different selection methods but which shared the same goal: improvement of growth rate as well as founder effects and breeding history of these strains.

A total of six STRs were reported by Yue & Orban (2002) in the *insulin-like growth factor II* (*IGFII*), *GH2*, *prolactin 1* and *insulin* genes of the Mozambique tilapia (*O. mossambicus*) and Nile tilapia (*O. niloticus*). Genotyping of 24 fish of each species revealed that all six STRs were polymorphic in both species. In the Mozambique tilapia, the number of alleles ranged from 3 to 17 (M = 9.8), whereas in the Nile tilapia, it ranged from 4 to 21 (M = 10.5). H_e was similar between species, 0.44 to 0.95 (M = 0.79) for Mozambique tilapia and 0.52 to 0.96 (M = 0.73) for Nile tilapia. The authors concluded that these STRs could be used to map genes and to detect segregation of quantitative trait loci that are important economically and in evolutionary studies. The STR identified by Yue & Orban

Table 5 Pairwise genetic differentiation using D_{EST} indices among four strains of Nile tilapia.

	UFLA	GIFT	CHIT	REDS
STRs in *GH1* gene				
UFLA	–	0.00	0.09**	0.40**
GIFT		–	0.16**	0.11**
CHIT			–	0.13**
REDS				–
Noncoding STR loci				
UFLA	–	0.09**	0.28**	0.40**
GIFT		–	0.30**	0.51**
CHIT			–	0.51**
REDS				–
Together				
UFLA	–	0.10**	0.32**	0.44**
GIFT		–	0.33**	0.47**
CHIT			–	0.58**
REDS				–

**P < 0.01.

GIFT, Genetic Improvement of Farmed Tilapias strain; UFLA, Universidade Federal de Lavras strain; CHIT, Chitralada strain; REDS, Red-Stirling strain.

© 2016 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12423

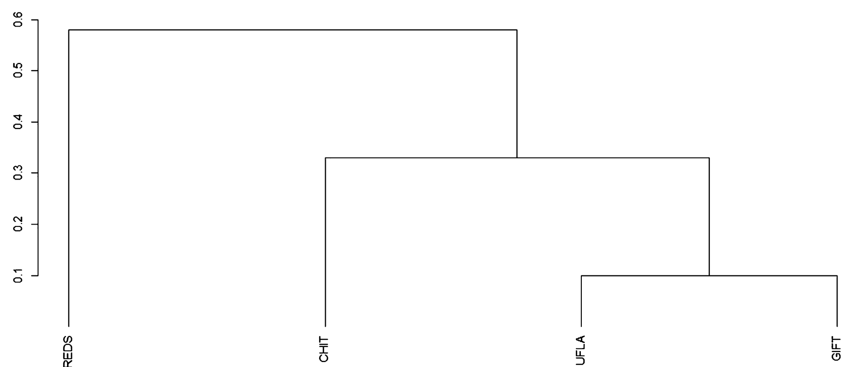

Figure 1 UPGMA dendrogram of four strains of Nile tilapia based on Jost's genetic differentiation index (D_{EST}).

(2002) in *GH2* [(CTGT)$_7$; in intron 1] was similar to the variant found in the intron 1 region of *GH1* in this present study. The total variation observed in the *GH2* gene (GenBank accession no. M97765) for the Mozambique and Nile tilapia species was as follows: allele numbers, 3 and 5; H_e, 0.44 and 0.55; and H_o, 0.29 and 0.50 respectively. The primers designed in this study for the intron 1 region were not specific for the *GH1* gene, and this was because the homology of *GH1* was >99% with the *GH2* tilapia gene (Ber & Daniel 1993). This fact has not diminished the potential importance of the STR intron as a source of variability among populations (Table 1). The prospect of a STR intron locus described above confirms the homology between *GH1* and *GH2*, given that the presence of three or four alleles was not observed in any individual during genotyping. The STR promoter was not found in the *GH2* sequence through analysis using the TANDEM REPEATS FINDER program analysis.

Genetic characterization and divergence of the four tilapia strains

The need to improve the genetic quality of Nile tilapia is essential to the future of tilapia farming worldwide (Li *et al.* 2006). Our findings characterize the genetic characteristics of *GH1* and noncoding STR markers for four tilapia strains. Although the GH-associated microsatellites were not as variable as those at noncoding loci, likely due to founder effects, breeding history and linkage to a locus that likely was subject to selection, they were useful to genetically distinguish the strains used herein and could prove useful as marker for within-strain marker-assisted selection.

The UFLA strain has undergone decades of mass selection in a region with a temperate wet climate, well-defined seasons and an annual average winter temperature of 16.5 °C (Dantas *et al.* 2007). Performance tests of the UFLA strain showed growth potential in intensive farming conditions (Freato *et al.* 2012). As a result, local adaptation may make the UFLA strain an important genetic resource for inclusion in further breeding programs. The UFLA and GIFT strains were more similar genetically, most likely because the GIFT strain was developed from wild African populations of Nile tilapia (Eknath *et al.* 1993) and the UFLA strain also originated from wild African populations without additional crossbreeding.

To understand adaptive mechanisms and identify causes of variation among populations, it is critical in breeding programs to seek out new strains with new market features. Ndiwa *et al.* (2014) compared the native strains from Loboi swamp, which had adapted to tolerate higher temperatures (spring strains), with other wild populations of *O. niloticus* using 16 STR loci. These authors found substantial genetic divergence between the spring strains and wild populations, and they therefore concluded that the protection of spring strains was imperative to prevent the loss of this genetic resource.

To genetically characterize strains or wild populations by STR, it is important to use loci from different linkage groups with a high level of polymorphism (Kocher *et al.* 1998). The use of less polymorphic loci should be considered in cases in which populations have private alleles (Ndiwa *et al.* 2014) or where there is or might be linkage with quantitative trait loci. In this present study, the *UNH-005* and *UNH-866* loci were used, even though these were less polymorphic, because they showed private alleles.

Rutten *et al.* (2004), while evaluating four strains of tilapia [CHIT, GIFT, IDRC (International Development Research Centers) and GOTT] at 14 STR loci, found an average number of alleles per locus that was similar to those reported here but with greater variation in the number of alleles per locus (between 5 and 20, with an average of 11 alleles). Moreira *et al.* (2007), while characterizing the CHIT, REDS and reciprocal F_1 strains at five STR loci, found variation ranging from seven to eight alleles per locus. The strain with the lowest variability indices was REDS; this is because this strain at formation was likely subject to founder effect and thus lost variability. However, this strain is important in crossbreeding programs because of its attractive color, which takes advantage of gains by heterosis and alleles that were adapted to culture under adverse environmental conditions (Shikano & Taniguchi 2003). Romana-Eguia *et al.* (2004), while comparing groups of wild tilapia and red-colored hybrids, found higher variability than those reported above. A difference of 2.5 alleles was observed between the group of wild strains (UFLA, CHIT

© 2016 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12423

and GIFT) and the REDS strain, whereas Romana-Eguia *et al.* (2004) reported 4.8 alleles. The GIFT strain used in both studies had an average deviation of 4.5 alleles. The observed and expected heterozygosity indices followed the same patterns. One explanation for this difference in results could be the number of STR markers, as the linkage groups and GIFT families being screened were not the same. Regardless, in both studies, it was possible to distinguish the level of variability between strains.

Rutten *et al.* (2004) observed similar values of H_e in the CHIT (0.62), GIFT (0.70), IDRC (0.71) and GOTT (0.70) strains, indicating that the selection schemes on known strains, such as CHIT and GIFT, changed the allelic and genotypic proportions depending on the number of founders and, mainly, on the effective founder number. H_e values for CHIT (0.79) were smaller than those for the REDS strain and their reciprocal F_1 hybrids ($M = 0.93$) (Moreira *et al.* 2007), in contrast to the values reported in this work, in which the REDS H_e was the lowest among the strains studied.

Tambasco *et al.* (2000) suggested that heterozygosity indices should not be considered in genetic mapping and kinship analyses because, in their study, the *BM1224* locus showed an excess of homozygotes and was also the locus with higher polymorphic information content (PIC = 0.685). In our study, an overall average PIC value of 0.68 was estimated, which indicated that the STRs used herein showed sufficient diversity for genetic differentiation among the strains (Botstein *et al.* 1980).

Overall levels of inbreeding for the 10 noncoding STR loci estimated in the four strain samplings by inbreeding (F_{is}) and relatedness (r) coefficients showed moderate to high levels of inbreeding. This was expected likely because of the founder effects when establishing the four strain stocks. Despite the inbreeding found by the two estimators across the four strains, the level of kinship among all individual pairs within the samplings evaluated in the current study was lower than that for full sibs ($r < 0.25$). It is important to mention that the performance of these two marker-based estimators of relatedness and inbreeding fluctuates and has been under scrutiny by different authors (Wang 2014; Taylor 2015). We have no evidence of inbreeding depression in these strains.

In the first quantification, the *GH1* gene STR loci were used independent of the genomic STR, given the importance of the *GH1* gene in muscle growth metabolism and for possible future applications in breeding programs (Poompuang & Hallerman 1997; Quéré *et al.* 2010). No differentiation was observed between the UFLA and GIFT strains, irrespective of the index used. The other comparisons were statistically significant using a D_{EST} that ranged from moderate (UFLA × CHIT, GIFT × REDS and CHIT × REDS) to high (UFLA × REDS and GIFT × CHIT) levels of differentiation.

The levels of differentiation, taking only the noncoding STR into consideration, were quantified. Again, higher values of genetic differentiation were obtained using the D_{EST} index. This could have been caused by the private alleles described above. Thus, moderate genetic differentiation between the UFLA and GIFT strains ($D_{EST} = 0.09$) was observed.

Genetic evaluations using STR loci in the *GH* gene (STR promoter and STR intron) and noncoding regions (10 STRs) together showed an increased level of genetic differentiation among strains. Although the STR promoter and STR intron loci were markers associated with coding loci, they improved the genetic differentiation of the tilapia strains that were studied, most likely because of the presence of private alleles that were discussed previously. Quéré *et al.* (2010) analyzed the genetic structure of four populations of European sea bass (*Dicentrarchus labrax*) at four functional loci (in the *IGF1*, *GH* and *somatolactin* genes), both individually and together. They concluded that the joint use of all loci improved the distinction of the populations that were evaluated.

The importance of conservation and the use of locally adapted genetic resources in aquaculture are pivotal to the success of fish breeding programs to develop more productive, resistant and cheaper production strains (Bartley *et al.* 2009). Because of the worldwide distribution of tilapia, their genetic resources have not yet been fully documented or assessed for use in aquaculture, and many of these resources have become threatened and irreversibly lost (Eknath & Hulata 2009; Hallerman & Hilsdorf 2014). The presence of STR polymorphisms in genes, such as *GH1*, might have the potential for genetic improvement (Poompuang & Hallerman 1997). Thus, further studies of associations between polymorphisms and growth rate should be carried out to evaluate these loci for inclusion in marker-assisted selection programs.

Despite the recent accumulation of studies based on SNPs (Williams *et al.* 2010; van Bers *et al.* 2012), the genetic characterization of strains using STRs either in noncoding regions or in genes has proven to be efficient and remains very useful for characterizing the genetic differences among strains in a breeding program. Another important finding described herein was the genetic characterization of the UFLA strain as a valuable genetic resource for breeding programs in Brazil. Results of this present study suggest that tilapia genetic improvement programs should incorporate a continuous process of selection using adapted genetic resources to generate more locally productive strains.

Acknowledgements

We would like to thank the São Paulo Research Foundation (FAPESP # 11/23752-2) for granting funds to support this project, and CNPq and CAPES for funding Marco Aurélio Dessimoni Dias's scholarship (#142832/2010-3 and #CAPG-BA no: 052-11 respectively). Also we are deeply in debt to Adilson and all the staff of Indústria Brasileira do

© 2016 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12423

Peixe, Ltda. for their assistance during the field work. This work was developed as part of the full requirements for the doctoral dissertation of M.A.D. Dias in Animal Sciences at the Federal University of Lavras (UFLA).

References

Almuly R., Poleg-Danin Y., Gorshkov S., Gorshkova G., Rapoport B., Soller M., Kashi Y. & Funkenstein B. (2005) Characterization of the 5′ flanking region of the growth hormone gene of the marine teleost, gilthead sea bream *Sparus aurata*: analysis of a polymorphic microsatellite in the proximal promoter. *Fisheries Science* **71**, 479–90.

Bartley D.M., Nguyen T.T.T., Halwart M. & Silva S.S. (2009) Use and exchange of aquatic genetic resources in aquaculture: information relevant to access and benefit sharing. *Reviews in Aquaculture* **1**, 157–60.

Benson G. (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27**, 573–80.

Ber R. & Daniel V. (1992) Structure and sequence of the growth hormone encoding gene from *Tilapia nilotica*. *Gene* **113**, 245–50.

Ber R. & Daniel V. (1993) Sequence analysis suggests a recent duplication of the growth hormone encoding gene in *Tilapia nilotica*. *Gene* **125**, 143–50.

van Bers N.E., Crooijmans R.P., Groenen M.A., Dibbits B.W. & Komen J. (2012) SNP marker detection and genotyping in tilapia. *Molecular Ecology Resources* **12**, 932–41.

Botstein D., White R.L., Skolnick M. & Davis R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**, 314.

Briñez R.B., Caraballo X.O. & Salazar M.V. (2011) Genetic diversity of six populations of red hybrid tilapia, using microsatellites genetic markers. *Revista MVZ Cordoba* **16**, 2491–8.

Dahm R. & Geisler R. (2006) Learning from small fry: the zebrafish as a genetic model organism for aquaculture fish species. *Marine Biotechnology* **8**, 329–45.

Dantas A.A.A., Carvalho L.G. & Ferreira E. (2007) Classificação e tendências climáticas em Lavras, MG. *Ciência e Agrotecnologia* **31**, 1862–6.

Davids A.H., Cloete S.W.P., Bester-van der Merwe A.E., Muchadeyi F., Slabbert R. & Dzama K. (2012) Genetic variation within and among three ostrich breeds, estimated by using microsatellite markers. *South African Journal of Animal Science* **42**, 156–63.

De-Santis C. & Jerry D.R. (2007) Candidate growth genes in finfish – where should we be looking? *Aquaculture* **272**, 22–38.

Eknath A.E. & Hulata G. (2009) Use and exchange of genetic resources of Nile tilapia (*Oreochromis niloticus*). *Reviews in Aquaculture* **1**, 197–213.

Eknath A.E., Tayamen M.M.M., Palada de Vera J.C. *et al.* (1993) Genetic improvement of farmed tilapias: the growth performance of eight strains of *Oreochromis niloticus* tested in different farm environments. *Aquaculture* **111**, 171–88.

Freato T.A., Freitas R.T.F., Pimenta M.E.S.G., Oliveira G.R., Reis Neto R.V. & Mattos B.O. (2012) Evaluation of Nile tilapia strains cultivated in cages under different feeding programmes. *Brazilian Journal of Animal Science*, **41**, 1332–6.

Hallerman E. & Hilsdorf A.W.S. (2014) Conservation genetics of tilapias: seeking to define appropriate units for management. *The Israeli Journal of Aquaculture – Bamidgeh* **66**, IJA_66.2014.1043.

IBGE, 2014. *Produção da Pecuária Municipal*. Instituto Brasileiro de Geografia e Estatística, Ministério do Planejamento, Orçamento e Gestão. 41, 108 pp.

Jost L. (2008) Gst and its relatives do not measure differentiation. *Molecular Ecology* **17**, 4015–26.

Kalinowski S.T., Taper M.L. & Marshall T.C. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16**, 1099–106.

Kocher T.D., Lee W.J., Sobolewska H., Penman D. & McAndrew B. (1998) A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). *Genetics* **148**, 1225–32.

Lee W.J. & Kocher T.D. (1996) Microsatellite DNA markers for genetic mapping in *Oreochromis niloticus*. *Journal of Fish Biology* **49**, 169–71.

Li G., He X.J., Hu G.C., Cai W.Q., Deng X.W. & Zhou P.Y. (2006) Improving growth performance and caudal fin stripe pattern in selected F_6–F_8 generations of GIFT Nile tilapia (*Oreochromis niloticus* L.) using mass selection. *Aquaculture Research* **37**, 1165–71.

Lupchinski-Júnior E., Vargas L., Povh J.A., Ribeiro R.P., Mangolim C.A. & Barrero N.M.L. (2008) Avaliação da variabilidade das gerações G_0 e F_1 da linhagem GIFT de tilapia do Nilo (*Oreochromis niloticus*) por RAPD. *Acta Scientiarum Animal Science* **30**, 233–40.

Melo D.C., Oliveira D.A.A., Ribeiro L.P., Teixeira C.S., Sousa A.B., Coelho E.G.A., Crepaldi D.V. & Teixeira E.A. (2006) Caracterização genética de seis plantéis comerciais de tilápia (*Oreochromis*) utilizando marcadores microssatélites. *Arquivo Brasileiro de Medicina Veteterinária e Zootecnia* **58**, 87–93.

Melo D.C., Oliveira D.A.A., Seerig A. & Carvalho D.C. (2008) Aplicações práticas de marcadores microssatélites na caracterização genética e identificação de plantéis de tilápia. *Revista Brasileira de Reprodução Animal* **32**, 220–4.

Mojekwu T.O. & Anumud C.I. (2013) Microsatellite markers in aquaculture: application in fish population genetics. *Journal of Environmental Science, Toxicology and Food Technology* **5**, 43–8.

Moreira A.A., Hilsdorf A.W.S., Silva J.V. & Souza V.R. (2007) Variabilidade genética de duas variedades de tilápia nilótica por meio de marcadores microssatélites. *Pesquisa Agropecuária Brasileira* **42**, 521–6.

Ndiwa T.C., Nyingi D.W. & Agnese J.F.N. (2014) An important natural genetics resources of *Oreochromis nilotius* (Linnaeus, 1758) threatened by aquaculture activities in Loboi drainage, Kenya. *PLoS ONE* **9**, 106972.

Nugent C.. (1988) The status of wild and cultured tilapia genetic resources in various countries. In: *Tilapia Genetic Resources for Aquaculture* (Ed. by R.S.V. Pullin), pp. 22–4. ICLARM Conference Proceedings 16, Manila, Philippines.

O'Connell M. & Wright J.M. (1997) Microsatellite DNA in fishes. *Reviews in Fish Biology and Fisheries* **7**, 331–63.

Oliveira E.J., Pádua J.G., Zucchi M.I., Vencovsky R. & Vieira M.L.C. (2006) Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology* **29**, 294–307.

Poompuang S., Hallerman E.M. (1997) Toward detection of quantitative trait loci and marker-assisted selection in fish. *Reviews in Fisheries Science* **5**, 253–77.

© 2016 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12423

Quéré N., Guinand B., Kuhl H., Reinhardt R., Bonhomme F. & Desmarais E. (2010) Genomic sequences and genetic differentiation at associated tandem repeat markers in *growth hormone*, *somatolactin* and *insulin-like growth factor-1* genes of the sea bass, *Dicentrarchus labrax*. *Aquatic Living Resources* **23**, 285–96.

R Development Core Team.2010. *R: a Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available at: http://www.R-project.org. Accessed on 15 June 2014.

Romana-Eguia M.R.R., Ikeda M., Basiao Z.U. & Taniguchi N. (2004) Genetic diversity in farmed Asian Nile and red hybrid tilapia stocks evaluated from microsatellite and mitochondrial DNA. *Aquaculture* **236**, 131–50.

Rousset F. (2008) GENEPOP'007: a complete reimplementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103–6.

Rutten M.J.M., Komen H., Deerenberg R.M., Siwek M. & Bovenhuis H. (2004) Genetic characterization of four strains of Nile tilapia (*Oreochromis niloticus* L.) using microsatellite markers. *Animal Genetics* **35**, 33–7.

Sakamoto T. & McCormick S.D. (2006) Prolactin and growth hormone in fish osmoregulation, *General and Comparative Endocrinology*, **147**, 24–30.

Schuelke M. (2000) An economic method for the fluorescent labelling of PCR fragments. *Nature Biotechnology* **18**, 233–4.

Shikano T. & Taniguchi N. (2003) DNA markers for estimation of inbreeding depression and heterosis in the guppy *Poecilia reticulata*. *Aquaculture Research* **34**, 905–11.

Tambasco D.D., Alencar M.M., Coutiho L.L., Tambasco A.J., Tambasco M.D. & Regitano L.C.A. (2000) Caracterização molecular de animais da raça Nelore utilizando microssatélites e genes candidatos. *Revista Brasileira de Zootecnia* **29**, 1044–9.

Tamura K., Peterson D., Peterson N., Stecher G., Nei M. & Kumar S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**, 2731–9.

Taggart, J.B., Hynes, R.A., Prodohl, P.A. & Fergusson, A. (1992) A simplified protocol for routine total DNA isolation from salmonid fishes, *Journal of Fish Biology*, **40**, 963–5.

Taylor R.L. (2015) The use and abuse of genetic marker-based estimates of relatedness and inbreeding. *Ecology and Evolution* **5**, 3140–50.

Untergasser A., Nijveen H., Rao X., Bisseling T., Geurts R. & Leunissen J.A.M. (2007) PRIMER3PLUS, an enhanced web interface to Primer3. *Nucleic Acids Research* **35**, W71–4.

Wang J. (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources* **11**, 141–5.

Wang J. (2014) Marker-based estimates of relatedness and inbreeding coefficients: an assessment of current methods. *Journal of Evolutionary Biology* **27**, 518–30.

Williams L.M., Ma X., Boyko A.R., Bustamante C.D. & Oleksiak M.F. (2010) SNP identification, verification, and utility for population genetics in a non-model genus. *BMC Genetics* **11**, 32.

World Bank 2013. Fish to 2030: prospects for fisheries and aquaculture. World Bank Report Number 83177-GLB. Washington D.C., USA.

Yue G.H. & Orban L. (2002) Microsatellites from genes show polymorphism in two related *Oreochromis* species. *Molecular Ecology Notes* **2**, 99–100.

Yue G.H., Li Y. & Orban L. (2001) Characterization of microsatellites in the *IGF-2* and *GH* genes of Asian seabass (*Lates calcarifer*). *Marine Biotechnology* **3**, 1–3.

Supporting information

Additional supporting information may be found in the online version of this article.

Figure S1 Relatedness evaluation: Chitralada strain.

Figure S2 Relatedness evaluation: GIFT strain.

Figure S3 Relatedness evaluation: Red-Stirling strain.

Figure S4 Relatedness evaluation: UFLA strain.

Table S1 Loci, base pairs, microsatellite motif, linkage group, primer sequence and accession number on GenBank used for microsatellite genotyping of Oreochromis niloticus.