Growth hormone gene polymorphism associated with grow-out performance of *Oreochromis niloticus* strains

Marco A. D. Dias [1], **Júlio S. S. Bueno-Filho** [2], **Rafael V. R. Neto** [3], **Suhaila K. K. Jaser** [4], **Aline A. Lago** [1], **Alexandre W. S. Hilsdorf** Corresp. [4]

[1] Department of Animal Sciences, Federal University of Lavras, Lavras, Minas Gerais, Brasil
[2] Department of Statistics, Federal University of Lavras, Lavras, Minas Gerais, Brasil
[3] Campus Registro, São Paulo State University, Registro, São Paulo, Brasil
[4] Unit of Biotechnology, University of Mogi das Cruzes, Mogi das Cruzes, São Paulo, Brasil

Corresponding Author: Alexandre W. S. Hilsdorf
Email address: wagner@umc.br

Introgressive crossbreeding between Red-Stirling and Chitralada tilapia was carried out to assess the association of STR loci polymorphisms located in the growth hormone (GH1) promoter and intron-I with growth performance. The growth of seven genetic groups was assessed using a mixed linear model. We considered sex, age, and initial weight class of each animal as fixed effects, while genetic group, STR-promoter, and STR-intron were considered as random effects. We found six alleles in the STR-promoter, which were arranged in 18 of the 21 genotypic possible combinations. Alleles 191 and 196 of the STR-promoter correspond to 80% of the assessed individuals' total alleles, and genotypes 196/196, 191/196, and 196/201 were the most frequent. We found only three alleles in the STR-intron. The loci STR-promoter showed a significant association with the animals' weight. Genotypes 181/181, 181/191, and 196/206 showed better average performance, and genotypes 196/201, 201/201, and 191/196 showed the worst performance. The polymorphism in the STR-intron was not associated with fish weight. The proposed approach is unprecedented for STR found in the GH1 gene. The observed association indicates that GH1 gene polymorphism could be used as a tool in tilapia breeding programs.

Growth hormone gene polymorphism associated with grow-out performance of *Oreochromis niloticus* strains

Marco Aurélio Dessimoni Dias[1,4], Júlio Sílvio de Sousa Bueno Filho[2], Rafael Vilhena Reis Neto[3], Suhaila Karim Khalil Jaser[1], Aline de Assis Lago[1,3], Alexandre Wagner Silva Hilsdorf[1,4]

[1]Unit of Biotechnology, University of Mogi das Cruzes, Mogi das Cruzes, SP, Brazil

[2]Department of Statistics, Federal University of Lavras, Lavras, MG, Brazil

[3]São Paulo State University, Campus Registro, Registro, SP, Brazil

[4]Department of Animal Sciences, Federal University of Lavras, Lavras, MG, Brazil

Corresponding author:

Alexandre Hilsdorf

Email: address: wagner@umc.br

ABSTRACT

Introgressive crossbreeding between Red-Stirling and Chitralada tilapia was carried out to assess the association of STR loci polymorphisms located in the growth hormone (GH1) promoter and intron-I with growth performance. The growth of seven genetic groups was assessed using a mixed linear model. We considered sex, age, and initial weight class of each animal as fixed effects, while genetic group, STR-promoter, and STR-intron were considered as random effects. We found six alleles in the STR-promoter, which were arranged in 18 of the 21 genotypic possible combinations. Alleles 191 and 196 of the STR-promoter correspond to 80% of the assessed individuals' total alleles, and genotypes 196/196, 191/196, and 196/201 were the most frequent. We found only three alleles in the STR-intron. The loci STR-promoter showed a significant association with the animals' weight. Genotypes 181/181, 181/191, and 196/206 showed better average performance, and genotypes 196/201, 201/201, and 191/196 showed the worst performance. The polymorphism in the STR-intron was not associated with fish weight. The proposed approach is unprecedented for STR found in the GH1 gene. The observed association indicates that GH1 gene polymorphism could be used as a tool in tilapia breeding programs.

INTRODUCTION

Classical breeding programs are the practice of developing genetic improvements by phenotype culling and pedigree information. Such programs have increased productivity for a majority of existing crop varieties and animal breeds. However, advances in molecular-assisted breeding have enhanced the efficiency with which breeders can select superior phenotypes with the best gene combinations. Genomic regions associated with quantitative traits (QTL) can be mapped using polymorphic DNA markers, such as microsatellites or SNPs (single nucleotide polymorphisms) (Pannier *et al*., 2010). This association can speed up breeding program outcomes and enable early identification of superior individuals with higher precision (Liu & Cordes, 2014).

Candidate genes have been under scrutiny, particularly those influencing growth rate in livestock and aquaculture species (De-Santis & Jerry, 2007). Growth hormone (GH) in particular is an anabolic hormone that regulates post-natal somatic growth processes, such as cellular division and muscular growth, while also modulating metabolism and energy homeostasis (Vijayakumar *et al*., 2011). The GH gene is not structurally conserved in finfishes and has five exons in carp (Chiou, Chen & Cjang, 1990; Ho, Wong & Chan, 1991) and six exons in other teleosts (Ber & Daniel, 1992; Male *et al*., 1992; Ohkubo *et al*., 1996; Almuly *et al*., 2000; Pinheiro *et al*., 2008). In addition, GH is duplicated (GH1 and GH2) in some finfish species, such as salmonids, common carp, and tilapia (Ber and Daniel, 1993; Du, Devlin & Hew, 1993; Figueroa *et al*., 2005). Polymorphisms have also been found in different segments in the GH

gene, and several studies have reported its association with growth in livestock and finfish. De-Santis & Jerry (2007) provide a review on this topic.

Tilapia is currently one of the most important aquaculture species worldwide (FAO, 2016). Different strains have been used in selective breeding programs of Nile tilapia (Eknath *et al*., 1993; Thodesen *et al*., 2013; Ansah, Frimpong & Hallerman, 2014). However, there is lack of understanding about the role of somatotropic axis genes on the growth of Nile tilapia under aquaculture systems. Furthermore, these gene polymorphisms have not been assessed in various Nile tilapia strains. Therefore, we investigated the putative presence of microsatellite variation in a GH promoter of Nile tilapia to assess the possible association of this polymorphism with growth rate in Nile tilapia strains and their recurrent crosses. The result could be targeted as a candidate gene in selective breeding programs for tilapia aquaculture.

Material & Methods

Origin of genetic material

Data were obtained from two genetically distinct strains of *Oreochromis niloticus,* Red-Stirling and Chitralada, and their crossbreeds. Seven genetic groups were obtained by backcross breeding (Lago *et al*., 2017), as shown in Table 1. The progenies from the seven genetic groups were obtained in mating hapas ($8 \times 6 \times 2$ m^3) using one dam with two sires after two-week intervals to synchronize spawning (Tsadik, 2014). Hapas were checked daily to identify reproduction, and larvae (fry) were collected weekly for four weeks. Animals collected in the same week were considered to be from the same age-of-birth group (ABG) for further statistical analyses (Fry collection: 1st week: 05/11/2013; 2nd week: 12/11/2013; 3rd week: 19/11/2013); 4th week: 26/11/2013). The ABGs were recorded to isolate the effect of age on the association

model. Furthermore, 100 fry from each genetic group and CGA were transferred to a 1 × 1 × 1 m^3 hapa based in a greenhouse. After achieving around 10 cm in length, fry were manually sexed and marked with passive integrated transponder (PIT) tags (VERI-TAG/ISSO FDX-B, Microchips-Brazil).

Grow-out performance

Fingerlings of each sex from the seven genetic groups were classified into four weight classes so that dominance behavior could be minimized. They were stocked into eight grow-out floating cages (2 m × 2 m × 1.8 m) with a final stock density of 58 fish/m^3 and separated by sex. All grow-out procedures were described by Lago *et al*. (2017), including water quality parameters and feeding regime. The following measures were taken to estimate the performance: body weight (BWe), standard length (SL), body height (BH), body width (BWi), head height (HH), head length (HL).

DNA Extraction and microsatellite (STR) polymorphism screening

At 190 days of grow-out, 25 tagged individuals of both sexes and genetic groups of a certain age were sampled from each floating cage to fin-clip for GH1 STR-promoter and STR-intron genotyping. Total genomic DNA was extracted using a method based on high salt concentration described by Aljanabi & Martinez (1997). DNA integrity and concentration were assessed by electrophoresis in 1% agarose gel with Lambda/HindIII (Fermentas, Waltman, Massachussets, USA) as a reference using a NanoVue™ spectrophotometer (GE Healthcare,

Little Chalfont, UK). All samples were diluted to a final concentration of 100 ng/µL and stored at −20 °C.

We used the GH1 sequence published by Ber and Daniel (1993) (GeneBank access number M97766) as a reference to assess the presence of STRs. Two STRs were found at position -719 in the promoter (ATTCT motif) and at position +141 in intron-I (CTGT motif). Two primer pairs were designed to amplify the STR regions (STR-promoter: 5'TGTAAAACGACGGCCAGTCCAGCATGTTTGCACTGAGTA3' and 5'GCCTAGCCATGGACACATTTA3'; STR-intron: 5'TGTAAAA CGACGGCCAGTACTGATGCCAGCCATGAACT3' and 5'AGAACAAACGCTGGCTG TCT3').

The amplification conditions were standardized and carried out in a final volume of 20 µl using 10 ng/µL of the isolated DNA; 0.5 UI of Taq DNA polymerase (Thermo Fisher Scientific, Inc.©, SP, Brazil); 1x buffer (100 mM Tris–HCl pH 8.8; 500 mM KCl); 2 mM of $MgCl_2$; 2.5 mM of dNTP; and 10 µM of each primer. The thermal cycles included an initial denaturation step of 60 sec at 95°C, followed by 35 cycles of denaturation at 95°C for 40 sec and annealing at 60°C for 60 sec, and then extension at 72°C for 30 s.

The electrophoresis and genotyping of the STR loci were carried out in a Matrix Plus of denaturing polyacrylamide 6.5% gel and a Li-Cor 4300 DNA Analyzer (IR2, Lincoln, NE, USA) using the IRDye®700 marker and universal M13 tail primer as described by Schuelke (2000). Allele sizes were gauged by interpolating their position relative to molecular weight markers (50-350 bp DNA Sizing Standard IRDye® 700) using the SagaGT Client program (Li-Cor Biosciences, Lincoln, NE, USA). Putative genotyping errors (Pompanon *et al*., 2005) were minimized by semi-automated scoring followed by visual inspection by two independent people.

141

142 **Data analysis**

143 To assess the genetic diversity of the STR-promoter and STR-intron of each genetic
144 group, we estimated the allelic and genotypic frequencies for each locus using HW-QuickCheck
145 (Kalinowski, 2006). MICRO-CHECKER (Van Oosterhout, 2004) software was used to infer any
146 Hardy-Weinberg equilibrium (HWE) deviation due to null alleles. The number of alleles (A) and
147 allelic richness (Ar) were obtained using FSTAT (version 1.2; Goudet, 1995).

148 The association between STRs and growth rate was investigated using a univariate mixed
149 model in each weighing. The fixed effects were the ABG, initial weight class, and sex, while the
150 random effects were the genetic group, STR-promoter, and STR-intron. In a matrix form, the
151 corresponding model is:

152 $$y = X\beta + Z_1 g_{\text{gg}} + Z_2 g_{\text{STR - promoter}} + Z_3 g_{\text{STR - intron}} + e$$

153 where:

154 y is the vector of realized (or possibly transformed) weights;

155 β is the vector of fixed effects;

156 g_{gg} is the vector of genetic group's random effect with distribution $\sim N(0, I\sigma_{\gamma_1})$ and $\gamma_1 =$

157 $\sigma_e / \sigma_{\text{gg}}$;

158 $g_{\text{STR - promoter}}$ is the vector of STR-promoter's random effect with distribution $\sim N(0, I\sigma_{\gamma 2}$

159) and $\gamma_2 = \sigma_e / \sigma_{\text{STR - promoter}}$;

160 $g_{\text{STR - Íntron}}$ is the vector of STR-intron's random effect with distribution $\sim N(0, I\sigma_{\gamma_3})$ and

161 $\gamma_3 = \sigma_e / \sigma_{\text{STR - Íntron}}$;

162 e is the vector of error with distribution $\sim N(0, I\sigma_e)$;

163 *and* X, Z_1, Z_2, and Z_3 represent the design matrices for the mentioned effects, respectively.

164 The lme4 package's functions lmer ( ) and ranef ( ) were used in the R program for
165 Windows to estimate the random effects of the mixed linear model (Bates *et al*., 2015). For
166 consistency of the database, we carried out visual analysis and used the function boxcox ( ) of the
167 MASS package to assess the best scale to normality assumption (Box & Cox, 1964). When a
168 significant association between the STR loci and the animals' weight was found, we ranked the
169 genotypes of each weighing according to their respective phenotypic average. Spearman's
170 correlation coefficient was calculated for the same feature in different measurements to confirm
171 the final classification of genotypes using PROC CORR in the Statistical Analysis System (SAS)
172 computational package (SAS Institute Inc., 2013).

173

174 **RESULTS**

175 **Genetic parameters of the STR-promoter and STR-intron**

176 Six alleles were found in the locus STR-promoter $(CTATT)_n$ at the -719 nucleotide
177 position (alleles 181, 186, 191, 196, 201, and 206). The most frequent alleles among the genetic
178 groups were 191 and 196, which had a frequency of 80%. We identified 18 out of 21 possible
179 genotypic combinations, and the most frequent genotypes where 196/196, 191/196, and 196/201
180 (Table 2).

181 Alleles 186 and 206 were exclusive to the Chitralada strain, and allele 201 was exclusive
182 to the Red-Stirling strain (Table 2). Genotyping errors in the STR-promoter locus were discarded
183 due to the presence of null alleles since there was evidence of null alleles in only the Red-Stirling
184 strain (GG1) and hybrids, of which dams had a higher proportion of Red-Stirling's alleles (GG3
185 and GG7).

The Chitralada (GG2) groups and the hybrids from Chitralada dams except from GG4 were in HWE. Excess homozygotes were observed in the genetic groups (F_{IS} = 0.16) except from the groups GG2 (F_{IS} = 0.00) and GG6 (F_{IS}) = -0.05). This shows that the backcross breeding process has fixed STR-promoter alleles (Table 3).

At the +141 nucleotide position of STR-Intron-I $(CTGT)_n$, there was no evidence of null alleles for any genetic group. At this locus, we identified three alleles (198, 202, and 206). Allele 198 was not observed in the Red-Stirling group (GG1). Alleles 202 and 206 were the most frequent among the genetic groups. The homozygote genotype 198/198 was found in only the GG6 group (Table 2). The GG7 group was the only one that presented deviation of the HWE (Table 3). The Chitralada (GG2) and Red-Stirling (GG1) groups had different results related to HWE, which was due to the number of alleles and the high homozygosity of the GG1 group (*A*= 4 and 5; F_{IS} = 0.20 and 0.24, respectively).

Association between GH1 STR loci and the grow-out performance of tilapia

The fixed effects of sex, ABG, and initial weight class were significant, which were therefore retained in the model. We did not compare the growth performance among genetic groups, but this variable was considered in the model to confirm the hypothesis that the genotype ranking of STR loci was corrected by the genetic group effect. The association analysis was performed in each weighing, and the STR-promoter locus was significantly associated with the animal weight in each weighing ($P < 0.05$).

Alleles 181 and 196 were present in the same proportion in the four STR-promoter genotypes with better performance (Table 4). However, alleles 191 and 201 showed high frequency in the genotypes with lower performance, and allele 201 was exclusive to the Red-

Stirling strain. Segregation of allele 201 was observed among the hybrid genetic groups (Table 2), which had lower performance compared to the Chitralada strain.

Among the genotypes found in the STR-promoter (Table 2), 181/181, 181/191, 196/206, and 196/196 showed the highest performance in all weighings. We verified that the rankings of the four best and four worst genotypes were constant, independently of weighing. The genotypes of low frequency in some of the measurements (186/186; 206/206; 181/186; 186/191; 186/201; 186/206; 201/206) were excluded from the ranking assessment.

The Spearman correlation coefficient estimated for genotypes among weighing measures showed medium to high values. This reflects the relative stability of the phenotypic performance in the genotype ranking of the STR-promoter locus among the assessed traits (Table 5). The correlation can be considered small for only the correlation between loci one and three ($P = 0.06$). STR-Intron polymorphism and growth rate showed no significant relationship through all weighings during the trial. Therefore, the STR genotypes found for this locus were not ranked, and consequently, their correlation was not assessed.

DISCUSSION

Microsatellites have been described in different livestock and fish genes in either promoter, 5′- and 3′-UTRs, or intronic sequences. The presence of STR in the GH gene has also been detected in Mozambique (*Oreochromis mossambicus*) and Nile tilapia (*O. niloticus*), gilthead sea bream (*Sparus aurata*), barramundi (*Lates calcarifer*), yellowtail (*Seriola quinqueradiata*), pufferfish (*Fugu rubripes*), and olive flounder (*Paralichthys olivaceus*) (Chistiakov, Hellemans, Volckaert, 2006; De-Santis & Jerry, 2007).

Our results showed the presence of two STRs (promoter and intron 1) in the GH gene of *O. niloticus*. Although they were described by Ber & Daniel (1993), their polymorphism and putative association with growth rate have not been determined thus far. Almuly *et al*. (2005) identified and characterized the STR (dinucleotide [CA]n in the 5'promoter region) in the GH gene of the gilthead sea bream. This study compared the 5′ flanking region of the GH gene of eight species of finfishes, including tilapia. A CA motif was identified near box transcription factors HFN3, PIT1, and TATA in yellowtail ($[CA]_5$ perfect STR motifs), gilthead sea bream ($[CA]_{13}$ perfect STR motifs), barramundi ($[CA]_{14}$ perfect STR motifs), and flounder ($[CA]_5$ CC $[CA]_3$ compound imperfect STR motifs). In the same 5′ flanking position of the GH promoter gene, two motif repeats [CA] were identified in *O. mossambicus* (Sekkali *et al*., 1999) and *O. niloticus* (Ber and Daniel, 1992). Even though this [CA]n STR is located in the same GH promoter position of other species, there are no studies of likely polymorphism of this STR in both *Oreochromis* tilapias.

To gain knowledge on how an STR variation can affect gene expression, we first need to determine the allelic variation of STR in a given candidate gene. For instance, Yue *et al*. (2001) detected 14 alleles in the insulin-like growth factor II gene (IGF-2) intron and six alleles in the GH gene. Six STRs were detected in the IGF-II, GH2, prolactin-I, and insulin genes of Mozambique tilapia (*O. mossambicus*) and Nile tilapia (*O. niloticus*). They were polymorphic for both species, ranging from 3 to 17 alleles in Mozambique tilapia and 4 to 21 in Nile tilapia (Yue & Orban, 2002). The authors posited that these allelic variations of STR loci could map genes coding for economically important features and QTLs.

The assessed STR located at the -719 position of promoter and the position +141 of intron-I in the two Nile tilapia strains (Red-Stirling and Chitralada) were polymorphic with five

and three alleles, respectively. As a result, 20 genotypes for promoter STR and six for intron-1 could potentially be found among the seven genetic groups. This polymorphism found in an isolated broodstock may be larger when assessing wild individuals and other genetically improved strains of Nile tilapia. Tjian & Maniatis (1994) posited that the transcriptional machinery of eukaryotic genes depends on enhancers and sequence-specific transcription factors of the promoter activated by extracellular signals that regulate the complex patterns of gene expression. Therefore, the modulation of transcriptional relies on an activator protein to bind to DNA that loops around on itself to interact with other proteins to regulate RNA polymerase activity. STR expansion or retraction near these sequence-specific transcription factors may interfere with the differential expression of messenger RNA with ensuing protein synthesis. Phenotypic growth is ultimately dependent on the serum level of the GH (De-Santis & Jerry, 2007).

The model used to assess the putative association between allelic variants and growth rate in tilapia showed significant values for STR (CTATT)n variants in the promoter. The best growth performance was associated with alleles 181 (similar frequency in GG1 and GG2 groups), 186 (exclusive to GG2), and 191 (16% more common in GG2). However, allele 201 (private of GG1) was associated with the worse performance (Table 4). Four genotypes were associated with higher average weight (STR-promoter 181/181, 181/191, 196/206, and 196/196), two of them (181/191, 196/206) were exclusive to GG2, and one (181/181) was exclusive to GG1. The frequency of genotype 196/196 was also associated with higher average weights at different weighings and was ten times higher in the GG2 group, showing the genetic superiority of this backcross breeding (Table 2). This illustrated by one dominant fish (196/206) weighing 109.32 g more than a recessive fish (201/201) at the final weighing (Table 4).

The hypothesis that GH polymorphism (SNP or STR) may be associated with economic traits has been proven in different livestock species (De-Santis and Jerry, 2007). For instance, Komatsu *et al*. (2011) carried out an association study on five polymorphisms of both SNPs and STRs in the GH secretagogue receptor gene with the carcass traits of Wagyu cattle. Two STRs were present: one in the promoter region ((TG)n) and one in the intron-I ((GTTT)n). The authors found a significant association for only the promoter's STR and showed that among the four most frequent alleles, the smaller one (TG)19 was associated with the best performance and higher carcass weight. The authors proposed hypotheses of transcriptional and translational processes to explain the best performance of this allele. To prove these hypotheses, they carried out a simulation of mRNA secondary structures produced by individuals with the alleles $(TG)_{19}$, $(TG)_{22}$, $(TG)_{23}$, $(TG)_{24}$, $(TG)_{29,}$ and $(TG)_{34}$ and with haplotypes TR-SNP's ($(TG)_{19}$-A and $(TG)_{24}$-C). They verified that the activity levels of mRNA secondary structures are affected by the presence of alleles $(TG)_{19}$ and not-$(TG)_{19}$ or between A and C. Therefore, they concluded that allele $(TG)_{19}$ is a potential marker and economically viable for use in breeding programs for Wagyu cattle.

As mentioned, studies on the characterization and polymorphism discovery of somatrotopic axis genes have been published with different finfish species. However, few have been devoted to clarifying the likely association of these genes with quantitative traits (*Paralichthys olivaceus* (Kang *et al*., 2002), *Sparus aurata* (Almuly *et al*., 2005), *Salmo salar* (Gross & Nilsson, 1999), *Oncorhynchus kisutch* (Forbes *et al*., 1994), *Salvelinus alpinus* (Tao & Boulding, 2003), and *Lates calcarifer* (Xu *et al*., 2006). Streelman & Kocher (2002) performed a pioneer study for tilapia. The assessed expression levels of the prolactin-1 (*prl1*) gene with STRs in its gene promoter ($(CA)_n$, gene position -200), which were associated with resistance to higher

salinity in hybrids from *O. mossambicus* (fit for higher salt concentrations) and *O. niloticus* (not fit). The authors observed STR alleles that confer better responses to higher salinities because of different STRs of variable length in the promoter. These may induce different DNA conformations and therefore decrease the synergistic interactions between the proteins and transcriptional factors in the gene promoter.

Two studies on the association between GH SNP polymorphism and trait performance in Nile tilapia have been published. The first one used general linear models to assess the *Pst*I locus polymorphism in intron-I in Chitralada and GIFT strains. A significant correlation was detected with total length, standard length, and body height and width (Blanck *et al*., 2009). The second study carried out a more extensive search for SNPs in the promoter and intron-I of two strains of Nile tilapia. As a result, a significant association was found between the promoter's SNPs and growth rate using a univariate linear mixed model (Jaser *et al.,* 2017).

CONCLUSIONS

According to De-Santis & Jerry (2007), "Candidate genes are commonly targeted based on prior knowledge of their role in the regulation of specific metabolic pathways influencing a particular quantitative trait…. putative genes are first surveyed for polymorphisms, and the statistical association between specific alleles and phenotypic expression of the trait of interest are examined." The outcomes of the present study seem consistent with De-Santis and Jerry´s criteria for selecting candidate genes for a marker-assisted selection program. Specifically, continuous investigations to identify polymorphisms in the GH gene with a wide sampling of different strains may reveal other alleles and composite genotypes to be tested in variable management and environmental conditions. It is important to mention that molecular

methodologies of gene expression such as RT-PCR (Real-Time PCR) or even RNA-Seq (RNA sequencing) would provide better understanding of the interaction between GH expression and growth rate in farmed animals.

ACKNOWLEDGMENTS

We are deeply in debt to Adilson and all the staff of Indústria Brasileira do Peixe, Ltda. for their assistance during the fieldwork. 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

Aljanabi SM, Martinez I. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* **25**:4692-4693 DOI 10.1093/nar/25.22.4692.

Almuly R, Cavari B, Ferstman H, Kolodny O, Funkenstein B. 2000. Genomic structure and sequence of the gilthead seabream (*Sparus aurata*) growth hormone-encoding gene: identification of minisatellite polymorphism in intron I. *Genome* **43**:836-845 DOI 10.1139/g00-051.

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* 71:479-490 DOI 10.1111/j.1444-2906.2005.00991.x.

346 Ansah YB, Frimpong EA, Hallerman EM. 2014. Genetically-Improved Tilapia Strains in Africa:
347 Potential Benefits and Negative Impacts. *Sustainability* **6**:3697-3721 DOI
348 10.3390/su6063697.
349 Bates D, Mächler M, Bolker BM, Walker SC. 2015. Fitting Linear Mixed-Effects Models Using
350 lme4. *Journal of Statistical Software* **67**:1-48 DOI 10.18637/jss.v067.i01.
351 Ber R, Daniel V. 1992. Structure and sequence of the growth hormone encoding gene from
352 Tilapia nilotica. *Gene* **113**:245-250 DOI 10.1016/0378-1119(92)90402-B.
353 Ber R, Daniel V. 1993. Sequence analysis suggests a recent duplication of the growth hormone
354 enconding gene in Tilapia nilotica. *Gene* 125:143-150 DOI 10.1016/0378-1119(93)90321-
355 S.
356 Blanck DV, Gasparino E, Ribeiro RP, Marques DS. 2009. Polimorfismo no gene GH1-PstI
357 associado a características corporais de variedades de tilápia do Nilo. *Pesquisa*
358 *Agropecuaria Brasileira* **44**:599-604 DOI 10.1590/S0100-204X2009000600008.
359 Box GEP, Cox DR. 1964. An analysis of transformations. *Journal of the Royal Statistical*
360 *Society. Series B, Statistical methodology* **26**:211-252.
361 Chiou CS, Chen HT, Cjang WC. 1990. The complete nucleotide sequence of the growth-
362 hormone gene from the common carp (*Cyprinus carpio*). *Biochimica et Biophysica Acta*
363 **1087**:91-94 DOI 10.1016/0167-4781(90)90126-M.
364 Chistiakov DA, Hellemans B, Volckaert FAM. 2006. Microsatellites and their genomic
365 distribution, evolution, function and applications: A review with special reference to fish
366 genetics. *Aquaculture* **255**:1-29 DOI 10.1016/j.aquaculture.2005.11.031
367 De-Santis C, Jerry DR. 2007. Candidate growth genes in finfish: where should we be looking?
368 *Aquaculture* **272**:22-38 DOI 10.1016/j.aquaculture.2007.08.036.

369 Du SJ, Devlin RH, Hew CL. 1993. Genomic structure of growth hormone genes in chinook
370 salmon (*Oncorhynchus tshawytscha*) presence of 2 functional genes, GH-I and GH-Ii, and
371 a malespecific pseudogene, GH-Psi. *DNA and Cell Biology* **12**:739-751 DOI
372 10.1089/dna.1993.12.739.
373 Eknath AE, Tayamen MM, Palada de Vera MS, Danting JC, Reyes RA, Dionisio EE, Gjedrem
374 JB, Pullin RSV. 1993. Genetic improvement of farmed tilapias: The growth performance
375 of eight strains of *Oreochromis niloticus* tested in eleven different environments.
376 *Aquaculture* **111**:171-188 DOI 10.1016/0044-8486(93)90035-W.
377 FAO. 2016. The State of World Fisheries and Aquaculture. FAO Fisheries and Aquaculture
378 Department. Rome, Italy, 200 pp.
379 Figueroa J, San Martín R, Flores C, Grothusen H, Kausel G. 2005. Seasonal modulation of
380 growth hormone mRNA and protein levels in carp pituitary: evidence for two expressed
381 genes. *Journal of Comparative Physiology B* **175**:185-192 DOI 10.1007/s00360-005-0474-
382 4.
383 Forbes SH, Knudsen KL, North TW, Allendorf FW. 1994. One of two growth-hormone genes in
384 coho salmon is sex-linked. *Proceedings of the National Academy of Sciences of the United*
385 *States* **91**:1628-1631.
386 Goudet J. 1995. Fstat version 1.2: a computer program to calculate F-statistics. *Heredity* **86**:485-
387 486.
388 Gross R, Nilsson J. 1999. Restriction fragment length polymorphism at the growth hormone 1
389 gene in Atlantic salmon (*Salmo salar* L.) and its association with weight among the
390 offspring of a hatchery stock. *Aquaculture* **173**:73-80 DOI 10.1016/S0044-
391 8486(98)00470-0.

392 Ho WKK, Wong M-W, Chan APY. 1991. Cloning and sequencing of the grass carp
393 (*Ctenopharyngodon idellus*) growth hormone gene. *Biochimica et Biophysica Acta*
394 **1090**:245-248 DOI 10.1016/0167-4781(91)90110-8.
395 Jaser SKK, Dias MAD, Lago AA, Neto RVR, Hilsdorf AWS. 2017. Single nucleotide
396 polymorphisms in the growth hormone gene of *Oreochromis niloticus* and their
397 association with growth. *Aquaculture Research* **48**:5835-5845 DOI 10.1111/are.13406.
398 Kalinowski ST. 2006. HW-QUICKCHECK: an easy-to-use computer program for checking
399 genotypes for agreement with Hrady-Weinberg expectation. *Molecular Ecology Notes*
400 **6**:974-979 DOI 10.1111/j.1471-8286.2006.01456.x.
401 Kang JH, Lee S-J, Park S-R, Ryu H-Y. 2002. DNA polymorhism in the growth hormone gene
402 and its association with weight in olive flounder *Paralichthys olivaceus*. *Fisheries Science*
403 **68**:494-498 DOI 10.1046/j.1444-2906.2002.00453.x.
404 Komatsu M, Iyoh T, Fujimori Y, Morita M. 2011. Genetic association between GHSR1a 5'UTR-
405 microsatellite and n(C>A) loci and growth and carcass traits in Japanese Black cattle.
406 *Animal Science Journal* **8**2:396–405 DOI 10.1111/j.1740-0929.2010.00852.x.
407 Liu ZJ, Cordes JF. 2014. DNA marker technologies and their applications in aquaculture
408 genetics. *Aquaculture* **238**:1-37 DOI 10.1016/j.aquaculture.2004.05.027.
409 Male R, Nerland AH, Lorens JB, Telle W, Lossius I, Totland GK. 1992. The complete
410 nucleotide sequence of the Atlantic salmon growth hormone I gene. *Biochimica*
411 *et Biophysica Acta* **1130**:345-348 DOI 10.1016/0378-1119(89)90079-6.
412 Ohkubo T, Araki M, Tanaka M, Sudo S, Nakashima, K. 1996. Molecular cloning and
413 characterization of the yellowtail GH gene and its promoter: a consensus sequence for

414 teleost and avian Pit-1/ GHF-1 binding sites. *Journal of Molecular Endocrinology* **16**:63-
415 72 DOI 10.1016/0378-1119(89)90079-6.
416 Pannier L, Hamill RM, Mullen AM, Sweeney T. 2010. Functional genomic approaches to
417 understand the biological pathways underpinning intramuscular fat in beef. *CAB Reviews*
418 *Perspectives in Agriculture Veterinary Science Nutrition and Natural Resources* **5**:1-11.
419 Pinheiro JS, Wolff JLC, Araújo RC, Hilsdorf AWS. 2008. Molecular cloning and sequence
420 analysis of growth hormone cDNA of Neotropical freshwater fish Pacu (*Piaractus*
421 *mesopotamicus*). *Genetics and Molecular Biology* **31**:381-384 DOI 10.1590/S1415-
422 47572008000200037.
423 Pompanon F, Bonin A, Bellemain E, Taberlet P. 2005. Genotyping errors: causes, consequences
424 and solutions. *Nature Reviews Genetics* 6:847-59 DOI 10.1038/nrg1707.
425 SAS Institute Inc. 2013. Base SAS® 9.4 Procedures Guide: Statistical Procedures, 2nd edition.
426 SAS Institute Inc., Carry, NC, USA.
427 Sekkali B, Brim H, Muller M, Argenton F, Bortolussi M, Colombo L, Belayew A, Martial JA.
428 1999. Structure and functional analysis of a tilapia (*Oreochromis mossambicus*) growth
429 hormone gene: activation and repression by pituitary transcription factor Pit-1. *DNA and*
430 *Cell Biology* **18**:489-502 DOI 10.1089/104454999315213.
431 Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature*
432 *Biotechnology* **18:** 233–234 DOI 18:233-234 DOI 10.1038/72708.
433 Streelman JT, Kocher TD. 2002. Microsatellite variation associated with prolactin expression
434 and growth of salt-challenged tilapia. *Physiological Genomics* **9**:1-4 DOI
435 10.1152/physiolgenomics.00105.2001.

436 Tao WJ, Boulding EG. 2003. Association between single nucleotide polymorphism in candidate
437 genes and growth rate in Arctic char (*Salvelinus alpinus* L.). *Heredity* **91**:60-69 DOI
438 10.1038/sj.hdy.6800281.
439 Thodesen J, Rye M, Yu-Xiang W, Kong-Song Y, Bentsen HB, Gjedrem T. 2013. Genetic
440 improvement of tilapias in China: genetic parameters and selection responses in growth,
441 survival and external color traits of red tilapia (*Oreochromis* spp.) after four generations of
442 multi-trait selection. *Aquaculture* 416/417:354-366 DOI
443 10.1016/j.aquaculture.2013.09.047
444 Tjian R, Maniatis T. 1994. Transcriptional activation: a complex puzzle with few pieces. *Cell* **77**:
445 5-8 DOI 10.1016/0092-8674(94)90227-5.
446 Tsadik GG. 2014. Comparison of broodfemales selection methods used at early maturity to
447 improve seed production in Nile tilapia, *Oreochromis niloticus* (L.). Global Research
448 Journal of Fishery Science and Aquaculture 1:1-9.
449 Van Oosterhout C, Hutchinson WF, Wills DPM and Shipley P (2004) MICRO-CHECKER:
450 Software for identifying and correcting genotyping errors in microsatellite data. *Molecular*
451 *Ecology Resources* **4**:535-538 DOI 10.1111/j.1471-8286.2004.00684.x
452 Vijayakumar A, Yakar S, Leroith D. 2011. The intricate role of growth hormone in metabolism.
453 *Frontiers in Endocrinology* **2**:1-11 DOI 10.3389/fendo.2011.00032.
454 Xu YX, Zhu ZY, Lo LC, Wang CM, Lin G, Feng F, Yue GH. 2006. Characterization of two
455 parvalbumin genes and their association with growth traits in Asian seabass (*Lates*
456 *calcarifer*). *Animal Genetics* **37**:266-268 DOI 10.1111/j.1365-2052.2006.01423.x.

Yue G, 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 DOI 10.1007/s101260000043.

Yue GH, Orban L. 2002. Microsatellites from genes show polymorphism in two related *Oreochromis* species. *Molecular Ecology Notes* **2**: 99-100 DOI 10.1046/j.1471-8286.2002.00159.x.

Table 1(on next page)

Genetic groups developed by backcross breeding

[a]REDS = Red-Stirling strain; [b]CHIT = Chitralada strain.

Table 1 - Genetic groups developed by backcross breeding

Male	Female	Genetic Group	Recoding
REDS	REDS	REDS[a]	GG1
CHIT	CHIT	CHIT[b]	GG2
CHIT	REDS	1/2 CHIT:REDS	GG3
REDS	CHIT	1/2 REDS:CHIT	GG4
GG4	CHIT	3/4 (GG4):CHIT	GG5
GG5	CHIT	7/8 (GG5):CHIT	GG6
CHIT	GG4	3/4 CHIT: GG4	GG7

[a]REDS = Red-Stirling strain; [b]CHIT = Chitralada strain.

1

Table 2(on next page)

Allelic and genotypic frequencies of the loci STR-promoter and STR-intron of each genetic group: GG1 (REDS); GG2 (CHIT); GG3 (½ CHIT:REDS); GG4 (½ REDS:CHIT); GG5 (¾ (GG4):CHIT); GG6 (7/8 (GG5):CHIT); GG7 (¾ CHIT:(GG4)).

Table 2 - Allelic and genotypic frequencies of the loci STR-promoter and STR-intron of each genetic group: GG1 (REDS); GG2 (CHIT); GG3 (½ CHIT:REDS); GG4 (½ REDS:CHIT); GG5 (¾ (GG4):CHIT); GG6 (7/8 (GG5):CHIT); GG7 (¾ CHIT:(GG4)).

	Genetic Groups						
Alleles	GG1	GG2	GG3	GG4	GG5	GG6	GG7
STR-promoter allelic frequency							
181	0.06	0.07	0.06	0.03	0.17	0.19	0.07
186	0	0.02	0	0	0	0.04	0
191	0.02	0.18	0.08	0.23	0.07	0.28	0.17
196	0.59	0.71	0.82	0.66	0.64	0.38	0.74
201	0.33	0	0.02	0.09	0.05	0.08	0
206	0	0.03	0.01	0	0.07	0.05	0.01
STR-intron allelic frequency							
198	0	0.1	0.03	0.05	0.04	0.15	0.03
202	0.36	0.3	0.38	0.48	0.26	0.25	0.4
206	0.64	0.6	0.59	0.46	0.69	0.61	0.56
Genótipos	STR-promoter genotypic frequency						
181/181	0.03	0	0,02	0.02	0.05	0.01	0.01
191/191	0	0,02	0,05	0.08	0	0.05	0.1
196/196	0.42	0,52	0,72	0.46	0.44	0.14	0.66
201/201	0.14	0	0,01	0	0.01	0	0
206/206	0	0	0,01	0	0	0	0
181/186	0.01	0	0,09	0	0	0.01	0
181/191	0	0,04	0	0	0.03	0.14	0.01
181/196	0	0,08	0	0.02	0.18	0.12	0
181/201	0.02	0,01	0	0	0.01	0.02	0
181/206	0	0	0	0	0.01	0.05	0
186/191	0	0	0	0	0	0.02	0
186/196	0	0,04	0	0	0	0.04	0
186/201	0	0	0	0	0	0.01	0
191/196	0.02	0,25	0,06	0.25	0.07	0.21	0.17
191/201	0.01	0	0	0.05	0.02	0.05	0
191/206	0	0,02	0	0	0.02	0.03	0
196/201	0.35	0	0,03	0.12	0.05	0.08	0.01
196/206	0	0,02	0,01	0	0.11	0.02	0.03
STR-intron genotypic frequency							
202/202	0.1	0,1	0.12	0.25	0,06	0.03	0.11
206/206	0.4	0,35	0.35	0.19	0,47	0.36	0.24
198/202	0	0,05	0.05	0.02	0,02	0.08	0.02
198/206	0	0,15	0.01	0.09	0,07	0.19	0.05
202/206	0.5	0,35	0.47	0.45	0,38	0.34	0.58

1

Table 3(on next page)

SSummary statistics for genetic diversity at STR-promoter and STR-intron loci of the backcross breeding between the Red-Stirling and Chitralada strains.

ap-value<0.01; NS= not significant: GG1=REDS; GG2=CHIT; GG3=½ CHIT:REDS; GG4=½ REDS:CHIT; GG5=¾ (GG4):CHIT; GG6=7/8 (GG5):CHIT; GG7=¾ CHIT: (GG4). number of alleles (A), allelic richness (Ar), null alleles estimate (An), mean estimate of Nei's gene diversity (hs), polymorphic information content (PIC), observed heterozygosity (Ho), expected heterozygosity (He), Hardy-Weinberg equilibrium (PHW), inbreeding coefficient FIS).

Table 3 - Summary statistics for genetic diversity at STR-promoter and STR-intron loci of the backcross breeding between the Red-Stirling and Chitralada strains: number of alleles (A), allelic richness (Ar), null alleles estimate (An), mean estimate of Nei's gene diversity (hs), polymorphic information content (PIC), observed heterozygosity (Ho), expected heterozygosity (He), Hardy-Weinberg equilibrium (PHW), inbreeding coefficient F_{IS}).

Population Parameters	Genetic Groups							
	GG1	GG2	GG3	GG4	GG5	GG6	GG7	Mean
STR-promoter								
A	5	6	5	4	5	6	5	2.14
A_r	5	5.91	5	4	5	6	4.94	5.12
H_o	0.42	0.47	0.18	0.44	0.5	0.79	0.3	0.44
H_e	0.54	0.47	0.32	0.51	0.54	0.75	0.42	0.51
P_{HW}	0.00[a]	0.51[NS]	0.0[a]	0.02[a]	0.11[NS]	0.15[NS]	0.00[a]	-
F_{is}	0.22	0	0.44	0.14	0.07	-0.05	0.29	0.16
STR-intron								
A	2	3	3	3	3	3	3	2.86
A_r	2	3	3	3	3	3	3	2.86
H_o	0.51	0.55	0.53	0.56	0.47	0.6	0.65	0.55
H_e	0.46	0.54	0.51	0.55	0.45	0.55	0.52	0.51
P_{HW}	0.14[NS]	0.45[NS]	0.38[NS]	0.43[NS]	0.31[NS]	0.11[NS]	0.00[a]	-
F_{is}	-0.11	-0.02	-0.04	-0.02	-0.05	-0.09	-0.25	-0.08

[a]p-value<0.01; NS= not significant: GG1=REDS; GG2=CHIT; GG3=½ CHIT:REDS; GG4=½ REDS:CHIT; GG5=¾ (GG4):CHIT; GG6=7/8 (GG5):CHIT; GG7=¾ CHIT: (GG4).

1

Table 4(on next page)

Genotypes ranking found for the STR-promoter, related to the weight in a recurrent cross plan between the Chitralada and Red-Stirling strains.

1average weight of individuals with the genotype; *number of individuals with the genotype.

Table 4 - Genotypes ranking found for the STR-promoter, related to the weight in a recurrent cross plan between the Chitralada and Red-Stirling strains.

	Measurement			
	1°	2°	3°	4°
Genotypes	Weight			
181/181	2 (49.71[1])[14*]	2 (69.78)[14]	3 (117.00)[14]	3 (202.00)[13]
181/191	1 (49.77)[22]	4 (70.71)[21]	6 (121.61)[18]	2 (202.52)[17]
181/196	3 (46.67)[60]	6 (66.71)[59]	11 (110.07)[56]	5 (179.64)[55]
181/201	9 (39.57)[7]	5 (59.29)[7]	4 (97.14)[7]	7 (158.00)[7]
181/206	5 (47.17)[6]	7 (67.33)[6]	9 (107.33)[6]	11 (163.00)[6]
186/196	4 (44.00)[7]	9 (62.57)[2]	7 (102.71)[7]	4 (166.00)[7]
191/191	6 (44.03)[29]	8 (64.03)[29]	5 (111.09)[23]	10 (185.00)[23]
191/196	10 (45.76)[100]	14 (66.56)[97]	14 (117.40)[90]	6 (195.14)[87]
191/201	12 (42.55)[11]	10 (61.91)[11]	10 (104.50)[10]	12 (171.50)[10]
191/206	11 (48.57)[7]	11 (68.00)[7]	8 (111.67)[3]	8 (175.00)[3]
196/196	8 (44.48)[321]	1 (64.70)[317]	1 (112.31)[290]	9 (187.80)[280]
196/201	14 (41.20)[54]	13 (56.44)[54]	13 (92.04)[45]	13 (154.70)[43]
196/206	7 (44.61)[18]	3 (68.11)[18]	2 (130.37)[16]	1 (222.12)[16]
201/201	13 (42.00)[13]	12 (50.17)[12]	12 (70.55)[11]	14 (112.80)[10]

1 [1]average weight of individuals with the genotype; *number of individuals with the genotype

Table 5(on next page)

Spearman's correlation (under the diagonal) and significance probability against the hypothesis of null correlation (above the diagonal) among the genotypic ranking of the STR-promoter locus in each weighing.

[a]p-value<0.01; [b]p-value<0.05.

Table 5 - Spearman's correlation (under the diagonal) and significance probability against the hypothesis of null correlation (above the diagonal) among the genotypic ranking of the STR-promoter locus in each weighing.

	1°	2°	3°	4°
1°	-	0.00[a]	0.06[b]	0.00[a]
2°	0.67	-		0.01[a]
3°	0.45	0.85	-	0.03[a]
4°	0.66	0.59	0.52	-

[a]p-value<0.01; [b]p-value<0.05.

1