ALEX JÚNIO DA SILVA CARDOSO

DIGITAL IMAGE ANALYSIS AS A TOOL FOR PHENOTYPING IN NILE TILAPIA SELECTIVE BREEDING

Thesis submitted to the Animal Science Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

Adviser: Paulo Sávio Lopes

Co-advisers: Carlos Antonio Lopes de Oliveira
Leonardo Bonato Felix

VIÇOSA – MINAS GERAIS
2022

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T

C268d
2022

Cardoso, Alex Júnio da Silva, 1993-
Digital image analysis as a tool for phenotyping in Nile tilapia selective breeding / Alex Júnio da Silva Cardoso. – Viçosa, MG, 2022.
1 tese eletrônica (78 f.): il. (algumas color.).

Texto em inglês.
Inclui anexo.
Orientador: Paulo Sávio Lopes.
Tese (doutorado) - Universidade Federal de Viçosa, Departamento de Zootecnia, 2022.
Inclui bibliografia.
DOI: https://doi.org/10.47328/ufvbbt.2022.338
Modo de acesso: World Wide Web.

1. Tilápia (Peixe) - Melhoramento genético. 2. Imagens digitais - Análise. 3. Redes neurais (Computação). 4. Morfometria. I. Lopes, Paulo Sávio, 1958-. II. Universidade Federal de Viçosa. Departamento de Zootecnia. Programa de Pós-Graduação em Zootecnia. III. Título.

CDD 22. ed. 639.3774

Bibliotecário(a) responsável: Bruna Silva CRB6/2552

ALEX JÚNIO DA SILVA CARDOSO

DIGITAL IMAGE ANALYSIS AS A TOOL FOR PHENOTYPING IN NILE TILAPIA SELECTIVE BREEDING

Thesis submitted to the Animal Science Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

APPROVED: April 18, 2022.

Assent:

Alex Júnio da Silva Cardoso
Author

Paulo Sávio Lopes
Adviser

ACKNOWLEDGMENTS

I am grateful to the Federal University of Viçosa (UFV), the Department of Animal Science (DZO), and the Animal Science Graduate Program for the opportunity granted to me during my Ph.D. program. I am also very grateful to the Maringá State University (UEM) and the PeixeGen Research Group for the support and opportunity to develop my research and the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq, Brazil; process no. 148102/2018-2) for financial support. This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior* – Brazil (CAPES) – Finance Code 001.

My sincere gratitude to my eternal advisor, Prof. Dr. Fabyano Fonseca e Silva, for the supervision, trust, advice, patience, friendship, and opportunities given to me. His early death was the hardest part of my Ph.D. I preserve in my memory and carry to my life his teachings, joy, desire to help others, and love for the profession. You will never be forgotten!

I thank my advisor, Prof. Dr. Paulo Sávio Lopes, for his understanding, support, and teachings during my Ph.D., especially in the final stage. I have learned a lot from you.

My sincere thanks to my co-advisor, Prof. Dr. Carlos Antonio Lopes de Oliveira, for providing all the necessary structure for carrying out this research. Thank you for receiving me so well at the PeixeGen group. I am grateful for the teachings, suggestions, patience, motivation, friendship, and opportunities. You are not only big in size but also a professional human being with a big heart.

I also thank my co-advisor, Prof. Dr. Leonardo Bonato Felix, for the support and motivation during my Ph.D.

I am very grateful to all Professors of the Department of Animal Science (DZO/UFV) for their teachings during my doctorate. Especially, I thank Profa. Dra. Renata Veroneze for her support, teachings, motivation, and friendship. Also, I thank the Professors of the Department of Statistics for their teachings.

I thank the members and ex-members of the Animal Breeding and Genetics Discussion Group (GDMA/UFV) for the discussions, good conversations, barbecues, and friendship, especially Arícia Reis, Layla Dias, Arielly Garcia, Eula Carrara, Daniele Marques, Leticia Fernanda, Pedro Vital, Talita Santana, Ingrid Garcia, Leandro Dornelas, Matheus Vieira, Túlio Vilar, Natália Galoro, Pedro Henrique, José Teodoro, Delvan Alves, Alessandra Alves, Hugo Silva, Darlene Duarte, and Ivan Carvalho. Also, a special thanks to my undergraduate friend, Profa. Dra. Susana Amaral Teixeira, for encouraging me to study Animal Breeding and Genetics and for all the help during my Ph.D.

I am very grateful to the PeixeGen Research Group (at UEM) for the excellent reception, hard work, and funny moments, especially those who helped me with dataset collection. A big thanks to Eric Campos, Carolina Schlotefeldt, Karla Tsujii, Ítalo Peres, Humberto Todesco, Gisele Silva, Filipe Chagas, Bruno Pires, Simone Siemer, Luiz Fernando, Gabriel Oliveira, Rodrigo Feuerharmel, Giovana Oliveira, Lais Mantovani, Elisangela De Cesaro, Maria Alice, Sérgio Costa, Sabrina Saito, and Nayara Vieira. I also, thank the group's general coordinator, Prof. Dr. Ricardo Pereira Ribeiro, for the support, teachings, and opportunities. It's a privilege to be part of this family.

I thank the technicians of the Aquaculture Station (CODAPAR/UEM), Vitor Moisés ("Vitão"), Cleiton Santos, and José Geraldo, for the support, hard work, and funny moments. I also thank the technicians Mário Everaldo and Antônio ("Teba") for their support and help in dataset collection.

I am very grateful to all Department of Animal Science (DZO/UFV) staff. Thanks, Alda Reis, for the coffee and gossip between classes. I also thank the Animal Science Graduate Program secretary, Fernanda Vieira, for all the clarifications.

I am very grateful to old friends that Animal Science introduced to me, Juliana Rodrigues, Naiara Marcos, Raully Lucas, João Vítor, Mariana Freitas, Camila Miranda, Felipe Martins, and Renata Sampaio, for the funny moments and advice in stressful moments. Thanks, Alícia Zem and Marcos Soares, for the many discussions, support, and friendship. Thank you to the friends of "Lapada F.C.". Also, my sincere thanks to Dr. Gutierrez José de Freitas Assis for all the help to establish an image capture protocol using Microsoft® Kinect.

I would also like to thank friends at LAFAP, especially Maria Tatiana, Wanderson Valente, and Karina Boechat for their support, conversations, and funny moments. Many thanks to my master's degree advisor, Prof. Dr. Jener Alexandre Sampaio Zuanon, for the motivation, opportunities, and help as always.

Thank you, Angelita ("Mercearia Paraná"), for softening the stressful moments, with that cold "BOA".

And last but not least, I would like to thank my parents (Francisco and Elza Helena) and my brothers (Willian and Wanderson) for their support and encouragement. Thank you very much for everything.

Thank you very much, God, for my life and connection with all these amazing people.

BIOGRAPHY

Alex Júnio da Silva Cardoso, son of Francisco Edilson Cardoso and Elza Helena da Silva Cardoso, was born on April 3, 1993, in Viçosa, Minas Gerais, Brazil. In December 2010, he finished high school at *Escola Estadual "Effie Rolfs"*, Viçosa, Minas Gerais, Brazil. In March 2011, he started his undergraduate studies in Animal Science at *Universidade Federal de Viçosa* (UFV). During his Bachelor's degree, he participated in research and extension activities in fish nutrition and fish physiology under Prof. Dr. Jener Alexandre Sampaio Zuanon. In January 2016, he completed his Bachelor's degree.

In March 2016, he started his studies in the Animal Biology Graduate Program at UFV, under Prof. Dr. Jener Alexandre Sampaio Zuanon, and completed it in April 2018.

In August 2018, he started his Ph.D. in the Animal Science Graduate Program at UFV under the supervision of Prof. Dr. Fabyano Fonseca e Silva. His Ph.D. project was conducted at *Universidade Estadual de Maringá* (UEM), under the supervision of Prof. Dr. Carlos Antonio Lopes de Oliveira. During his Ph.D., he participated in the coordination of the Animal Breeding and Genetics Discussion Group (GDMA) at UFV and the research group PeixeGen at UEM. His Ph.D. thesis was approved on April 18, 2022.

ABSTRACT

CARDOSO, Alex Júnio da Silva, D.Sc., Universidade Federal de Viçosa, April, 2022. **Digital image analysis as a tool for phenotyping in Nile tilapia selective breeding**. Adviser: Paulo Sávio Lopes. Co-advisers: Carlos Antonio Lopes de Oliveira and Leonardo Bonato Felix.

Phenotyping is an important step for successful animal selective breeding. Computer vision systems, such as digital image analysis, paired with machine learning (e.g., artificial neural networks, ANN), have the potential to be used in precision aquaculture and genetic improvement programs. Digital image analyses are suitable for determining morphometric traits (e.g., length, height, and width) and traits difficult to measure using traditional techniques, such as body areas, while reducing animal handling. Furthermore, images can provide explanatory variables for posterior prediction of growth, carcass, and fillet traits through ANN models. Therefore, we aimed to (i) develop a fast and straightforward method to measure the length, height, and body areas of Nile tilapia using digital image analysis, (ii) estimate genetic parameters for these traits, and (iii) apply image traits in machine learning algorithms to predict body weight (BW), carcass weight (CW), fillet weight (FW), and fillet yield (FY). The fish used in the study belonged to the 10th and 11th generation of the Nile tilapia breeding program (TILAMAX strain) of the *Universidade Estadual de Maringá*. In the first study, 656 fish (366 days old at harvest, BW of 414 ± 98 g) were photographed and subjected to image analysis to measure the trunk area (TA), head area (HA), caudal fin area (CFA), and fillet area (FA). Heritability estimates (h^2) for BW, TA, HA, CFA, and FA were 0.25, 0.23, 0.26, 0.21, and 0.25, respectively. Genetic correlations between the traits were positive and high, ranging from 0.70 to 0.98. We highlight the genetic correlation between BW and TA (r_G = 0.98) and FA (r_G = 0.97). Given the observed results, it can be concluded that selecting for body areas obtained by digital image analysis can lead to indirect genetic gains in weight and other areas. However, genetic correlations of these body areas with fillet weight and fillet yield were unknown. For the second study, 1,161 fish (427 days old at harvest, BW of 1,093 ± 346 g) were photographed. Body lengths (3 sections), heights (5 sections), TA, HA, FA, and total area (TOT) were measured from the coordinate values (*x* and *y* values in the center of each pixel) of 20 pre-set landmarks on the surface of fish images using the free R software. The proposed method allowed to measure 12 traits in 46 s. The h^2 for lengths and heights were moderate to high, ranging from 0.22 to 0.37. The h^2 values for TA, HA, FA, and TOT were 0.26, 0.35, 0.25, and 0.27, respectively. Positive and moderate to high genetic correlations were observed between morphometric traits and BW (0.66 to 0.98), FW (0.50 to 0.91), and CW (0.77 to 0.98). We highlight the genetic correlation of TA with BW (r_G = 0.98), FW (r_G = 0.91), and CW (r_G = 0.96). The TA/TOT ratio showed a positive and moderate genetic correlation (0.54) with FY. We investigated five supervised machine learning methods for predicting BW, CW, FW, and FY using image traits: multiple linear regression, feed-forward artificial neural network, deep learning, Bayesian regularization for feed-forward neural networks, and random forests. To verify the effectiveness of prediction methods, we used a 10-fold cross-validation procedure with 5 replicates, and the folds were randomly split to provide the training (n = 1045) and validation (n = 116) datasets. Pearson's correlation coefficient (*r*), mean absolute error (MAE), and root mean square error (RMSE) between predicted and observed values were calculated. In general, the Bayesian regularization model showed better performance and accuracy in predicting BW (*r* = 0.99, MAE = 39.54, RMSE = 54.70), CW (*r* = 0.98, MAE = 27.82, RMSE = 40.03), and FW (*r* = 0.96, MAE = 23.26, RMSE = 33.42). For FY prediction, all evaluated models had low performance and accuracy (*r* = 0.29, MAE = 1.55, RMSE = 2.24). The findings demonstrate that digital image analysis is a promising tool for measuring morphometric traits in Nile tilapia, given its non-invasive nature, fast operation, and low cost. Additionally, it was found that body areas can be used as selection criteria, particularly in future studies on body shape changes, with positively correlated responses to FW and positive, albeit lower, correlations with FY. Finally, the Bayesian regularization for the feed-forward neural network method showed the best performance in predicting BW, CW, and FW in Nile tilapia from image traits as predictor variables.

Keywords: Morphometric traits. Computer vision. Genetic parameters. Irregular polygons. Artificial neural networks.

RESUMO

CARDOSO, Alex Júnio da Silva, D.Sc., Universidade Federal de Viçosa, abril de 2022. **Análise de imagens digitais como ferramenta para fenotipagem no melhoramento genético de tilápia do Nilo.** Orientador: Paulo Sávio Lopes. Coorientadores: Carlos Antonio Lopes de Oliveira e Leonardo Bonato Felix.

A fenotipagem é uma etapa importante para o sucesso do melhoramento genético animal. Os sistemas de visão computacional, como a análise de imagens digitais, juntamente com o aprendizado de máquina (por exemplo, redes neurais artificiais, RNA), têm potencial para serem utilizados na aquicultura de precisão e em programas de melhoramento genético. As análises de imagens digitais são adequadas para determinar características morfométricas (por exemplo, comprimento, altura e largura), e também características difíceis de mensurar por meio de técnicas tradicionais, como as áreas corporais, reduzindo o manuseio dos animais. Além disso, as imagens podem fornecer variáveis explicativas para posterior predição de características de crescimento, carcaça e filé, por meio de modelos de RNA. Portanto, objetivamos (i) desenvolver um método rápido e descomplicado para medir o comprimento, a altura e as áreas corporais de tilápias do Nilo utilizando análise de imagens digitais, (ii) estimar parâmetros genéticos para essas características e (iii) aplicar as características obtidas por imagem em algoritmos de aprendizado de máquina para predizer o peso corporal (BW), peso da carcaça (CW), peso do filé (FW) e rendimento do filé (FY). Os peixes utilizados no estudo pertenciam à 10ª e 11ª geração do programa de melhoramento genético de tilápia do Nilo (linhagem TILAMAX) da Universidade Estadual de Maringá. No primeiro estudo, 656 peixes (366 dias de idade na despesca, BW de 414 ± 98 g) foram fotografados e submetidos à análise de imagens para mensuração da área do tronco (TA), área da cabeça (HA), área da nadadeira caudal (CFA) e área do filé (FA). As estimativas de herdabilidade (h^2) para BW, TA, HA, CFA e FA foram 0,25, 0,23, 0,26, 0,21 e 0,25, respectivamente. As correlações genéticas entre as características foram positivas e altas, variando de 0,70 a 0,98. Destacamos a correlação genética entre BW e TA (r_G = 0,98) e FA (r_G = 0,97). Diante dos resultados observados, pode-se concluir que a seleção para áreas corporais obtidas por análise de imagens digitais pode levar a ganhos genéticos indiretos no peso e em outras áreas. No entanto, as correlações genéticas dessas áreas corporais com o peso do filé e o rendimento do filé eram desconhecidas. Para o segundo estudo, 1.161 peixes (427 dias de idade na despesca, BW de 1.093 ± 346 g) foram fotografados. Os comprimentos (3 seções), alturas (5 seções), TA, HA, FA e área total (TOT) foram medidos a partir dos valores das coordenadas (valores x e y no centro de cada pixel) de

20 pontos de referência pré-definidos nas imagens, utilizando o software livre R. O método proposto permitiu medir 12 características em 46 s. As h^2 para os comprimentos e as alturas foram de moderadas a altas, variando de 0,22 a 0,37. Os valores de h^2 para TA, HA, FA e TOT foram 0,26, 0,35, 0,25 e 0,27, respectivamente. Correlações genéticas positivas e moderadas a altas foram observadas entre as características morfométricas e BW (0,66 a 0,98), FW (0,50 a 0,91) e CW (0,77 a 0,98). Destacamos a correlação genética da TA com BW (rG = 0,98), FW (rG = 0,91) e CW (rG = 0,96). A relação TA/TOT apresentou correlação genética positiva e moderada (0,54) com FY. Investigamos cinco métodos supervisionados de aprendizado de máquina para predição do BW, CW, FW e FY utilizando características de imagem: regressão linear múltipla, rede neural artificial *feed-forward*, aprendizado profundo, regularização Bayesiana para redes neurais *feed-forward* e florestas aleatórias. Para verificar a eficácia dos métodos de predição, utilizamos um procedimento de validação cruzada de 10-*folds* com 5 repetições, e os *folds* foram separados aleatoriamente para fornecer os conjuntos de dados de treinamento (n = 1045) e validação (n = 116). O coeficiente de correlação de Pearson (r), erro absoluto médio (MAE) e erro quadrático médio (RMSE) entre os valores previstos e observados foram calculados. Em geral, o modelo de regularização Bayesiana apresentou melhor desempenho e acurácia na predição de BW (r = 0,99, MAE = 39,54, RMSE = 54,70), CW (r = 0,98, MAE = 27,82, RMSE = 40,03) e FW (r = 0,96, MAE = 23,26, RMSE = 33,42). Para a predição de FY, todos os modelos avaliados apresentaram baixo desempenho e acurácia (r = 0,29, MAE = 1,55, RMSE = 2,24). Os resultados observados evidenciam que a análise de imagens digitais é uma ferramenta promissora para medir características morfométricas em tilápias do Nilo, dada a sua natureza não invasiva, rápida e de baixo custo. Além disso, verificou-se que as áreas corporais podem ser utilizadas como critérios de seleção, principalmente em estudos futuros sobre mudanças na forma corporal, e também, apresentam respostas correlacionadas positivas com o FW, bem como correlações positivas, embora menores, com o FY. Por fim, a regularização Bayesiana para o método de rede neural *feed-forward* apresentou o melhor desempenho na predição de BW, CW e FW em tilápias do Nilo utilizando características mensuradas por imagem como variáveis preditoras.

Palavras-chave: Características morfométricas. Visão computacional. Parâmetros genéticos. Polígonos irregulares. Redes neurais artificiais.

SUMMARY

CHAPTER 1:

Literature review

1.1. Selective breeding of Nile tilapia

Aquaculture has been growing steadily since 1950, mainly as a result of increased fish consumption and stagnation of extractive fisheries (FAO, 2020). Global fish production by inland and marine farming reached 54.3 million tonnes in 2018, representing 47.4% of the total aquaculture production (FAO, 2020). Nile tilapia (*Oreochromis niloticus*) ranked third among the most cultivated species in the world, with 4.5 million tonnes produced in 2018 (8.3% of farmed fish production) (FAO, 2020). Brazil is the fourth-largest producer of tilapia in the world, with 534 thousand tonnes produced in 2021, which represents 63.5% of the national fish production (PEIXE BR, 2022).

Nile tilapia, belonging to the family Cichlidae, is native to Africa and distributed in tropical and subtropical regions (Popma and Masser, 1999). This freshwater species is widely cultivated in several parts of the world because it displays interesting productive and industrial characteristics, such as sexual precocity, natural reproduction, capacity to feed on a wide variety of natural foods, accept rations in the larval stage, fast growth, hardiness, high resistance to diseases and parasites, suitability for a wide range of farming systems, absence of Y-shaped intramuscular fishbones, and fillets with good organoleptic traits that are highly appreciated by consumers.

A genetic improvement is a crucial tool for increasing Nile tilapia production in Brazil and worldwide, which, together with adequate animal nutrition, handling, and welfare, contributes to enhancing yield and product quality. There are several approaches to the genetic improvement of aquatic animals, such as hybridization and crossbreeding, chromosomal manipulation, transgenesis, sexual control, and selective breeding (Ponzoni et al., 2011). Selective breeding programs provide continued and permanent genetic gains that are transmitted from generation to generation through the mating of individuals more adapted to the cultivation environment and genetically superior for certain traits. For this, genetic quality indicators (additive genetic values) are predicted from information on individual traits (phenotypes) and kinship between individuals.

The first study investigating selection in fish dates back to 1925, whereby trout (*Salvelinus fontinalis*) were selected for increased survival from furunculosis (Embody and Hayford, 1925; Gjedrem, 2012). In the 1970s, the first efficient genetic improvement program for Atlantic salmon (*Salmo salar*) was implemented in Norway to initially master breeding and cultivation techniques and then of selecting animals (between and within families) for body weight at harvest and age at sexual maturity (Gjedrem, 2012). Between 1986 and 1988, a major genetic improvement program for Nile tilapia was developed by the WorldFish Center in the Philippines under the name GIFT (Genetic Improvement of Farmed Tilapias), which focused primarily on the estimation of phenotypic and genetic parameters for growth (Gjedrem, 2012). To ensure high genetic variability, the GIFT program used four wild strains from Africa (Egypt, Ghana, Kenya, and Senegal) and four commercial strains from Asia (Israel, Singapore, Taiwan, and Thailand) (Gupta and Acosta, 2004; Gjedrem, 2012). In 2005, GIFT tilapia was introduced in Brazil through the *Universidade Estadual de Maringá* (UEM), where the first selective breeding program was established. The base population comprised 600 animals from 30 families (full-sibs) of the eighth generation of selection imported from Malaysia (Santos et al., 2011; Yoshida et al., 2021). Initially, fish were selected for growth rate. Currently, the program is in the 13th generation of selection.

A typical breeding program for Nile tilapia can be divided into three phases (Figure 1). (1) The breeding/selection center is a facility that breeds animals for selection and keeps full- and half-sib families (depending on the mating structure) separately in hapas/tanks until they reach the appropriate size to be tagged with passive integrated transponder tags. (2) The second step consists of cultivation and genetic evaluation. A random sample of individuals from each full-sib family are tagged, and breeding candidates are cultivated in commercial farm environments (different production systems under extreme and/or challenged conditions) or in the breeding/selection center. During this phase, a phenotypic analysis of important traits is performed according to the objectives and selection criteria of the breeding program. Data on test animals and/or breeding candidates are used for genetic evaluation, and animals with high genetic values for a given trait are selected to be parents of the next generation, being returned to the breeding center. (3) The third phase is conducted in multiplier units responsible for disseminating the improved material to producers and the industry.

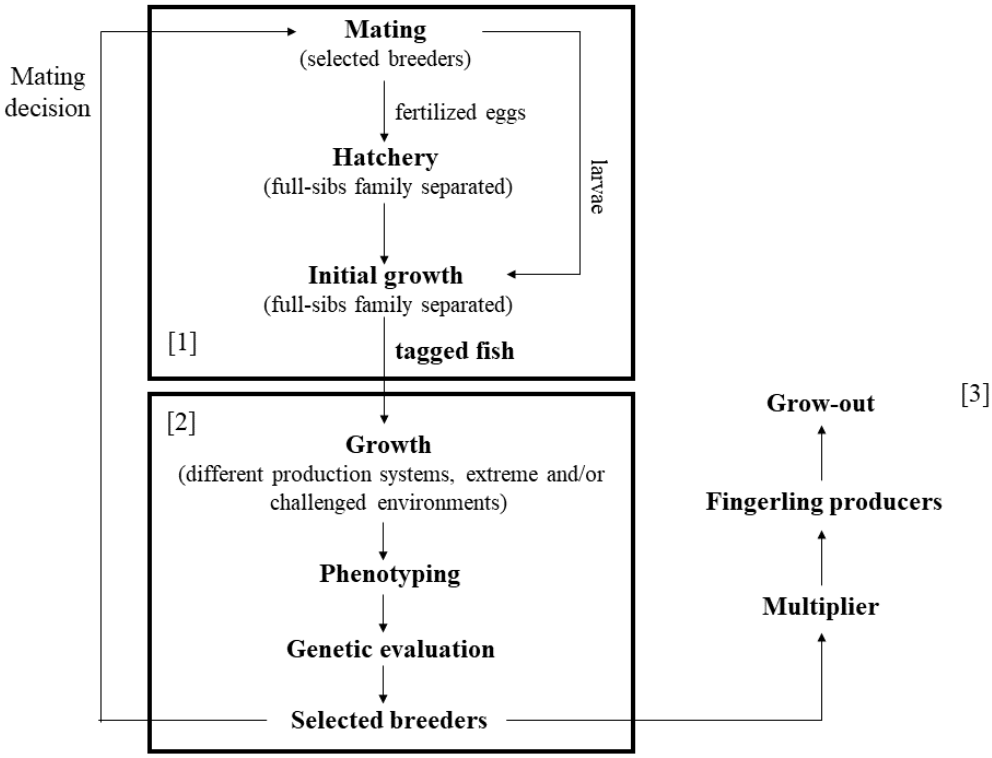

Figure 1. Main elements of a typical breeding program for Nile tilapia (adapted from Gjerde, 2005).

Growth-related traits (e.g., body weight and daily weight gain) are the major traits targeted in tilapia breeding programs (Ponzoni et al., 2005; Thodesen et al., 2011; Oliveira et al., 2016; Yáñez et al., 2020; Yoshida et al., 2021), to produce fish with higher growth rates and lower harvest ages. The heritability (h^2) for Nile tilapia body weight varies greatly, ranging from 0.06 to 0.84 (Ponzoni et al., 2005; Thodesen et al., 2011; Bentsen et al., 2012; Trong et al., 2013; Oliveira et al., 2016; Turra et al., 2016; Araújo et al., 2020). Furthermore, morphometric traits such as length (total and standard), width, and height are routinely recorded and used to select body shape, they are highly correlated with growth traits (Nguyen et al., 2007; Trong et al., 2013). In general, heritability estimates for tilapia morphometric traits are moderate to high, ranging from 0.20 to 0.60 (Rutten et al., 2005; Nguyen et al., 2007; Trong et al., 2013; Reis Neto et al., 2014; Araújo et al., 2020). Of note, in the referred studies, phenotypic traits were measured manually using rulers, requiring intense animal handling.

The initial focus for selective breeding has traditionally been the growth rate. However, some genetic improvement programs have started to consider traits related to fish processing as selection criteria. In the tilapia market, fillet weight and yield (fillet weight/body weight ratio) are economically important traits, given that tilapia are mostly sold as fresh or frozen fillets and, less commonly, as whole fish. It is important to note, however, that fillet yield has low economic value in itself. From an economic perspective, a determinant trait is total fillet production, as a small fish with a high fillet yield has a lower value than a large fish with the same fillet yield (Gjerde et al., 2012).

Fillet yield in Nile tilapia ranges from 33 to 45%, and heritability estimates are low to moderate (0.06 to 0.32) (Rutten et al., 2005; Nguyen et al., 2010; Gjerde et al., 2012; Garcia et al., 2017), suggesting that this trait can be improved by selection. The methods most commonly used for the selection of fillet weight and yield are sib selection or correlation with other body measurements because fillet-related traits are costly to determine and cannot be measured on live breeding candidates by traditional techniques (Rutten et al., 2005; Nguyen et al., 2010; Gjerde et al., 2012). In selection based on sib information, breeding candidates are ranked according to the average performance of slaughtered sibs. However, selection intensity using sib information is less compared to direct and correlated traits selection. The use of indirect criteria highly correlated with fillet weight and/or yield, which can be measured on live candidates, allows exploring the existing variation with higher success than sib selection (Gjerde et al., 2012). Genetic correlations between fillet weight and growth-related traits (e.g., body weight) are high, whereas that between fillet yield and body weight is low (Rutten et al., 2005; Nguyen et al., 2010; Gjerde et al., 2012).

1.2. Phenotyping by digital image analysis

The success of selective breeding is highly dependent on the quality and regularity of recorded phenotypes as well as on the proper use of phenotypic and pedigree information (Ventura et al., 2020). According to Ventura et al. (2020), the development of new technologies (e.g., "omics"), equipment, efficient measurement protocols (e.g., visual computing, use of images and sensors), and new statistical and bioinformatics tools can accelerate genetic progress per unit of time.

Phenotypic data used in breeding programs for Nile tilapia and other aquatic species are conventionally recorded visually or manually, which makes the procedures labor-intensive and, in some cases, inaccurate as a result of human error (Ventura et al., 2020; Fu and Yuna, 2022). Another disadvantage of such methods is that manual measurement of growth-related traits requires the removal of live fish from culturing places and, depending on the traits to be measured, the intensive handling can cause stress and injury (Fu and Yuna, 2022). Automation of phenotyping processes in aquaculture would allow continuous monitoring in the field and the assessment of larger sample sizes, resulting in reduced handling, improved fish welfare, and greater accuracy and repeatability of phenotypic measurements (Ventura et al., 2020; Fu and Yuna, 2022).

Machine vision (or computer vision) is a technique that can be used for the precise measurement of important traits in fish genetic improvement. Machine vision uses image processing and analysis techniques for the development of artificial systems capable of handling visual problems of interest (Fernandes et al., 2020a). Image-based approaches can be used to obtain digital phenotypes in a highly automated manner, avoiding excessive animal handling and manual recording of data (Ventura et al., 2020). In addition to being more economical and faster, machine vision is 10% more accurate than traditional methods (Hao et al., 2015). With the advancement of artificial intelligence through the development of deep learning algorithms, studies in fisheries and aquaculture are related to fish size measurement (e.g., length and height), body mass estimates, behavior monitoring, classification, detection, and counting (Li and Du, 2021).

During the formation of digital images, luminous signals are captured by a sensor (individual sensors arranged in the form of a 2D array, such as charge-coupled devices), coded, and stored in a matrix of numerical values corresponding to the luminous intensity (color) of each pixel (Gonzalez and Woods, 2018; Fernandes et al., 2020a). This array can be a single matrix for binary images (black and white) and grayscale images or a combination of three matrices for color images (intensities for red, green, and blue, corresponding to the RGB color space) (Fernandes et al., 2020a). With this image structure, it is possible to measure the fish length from the proportional relationship between the number of pixels comprising the fish body and the number of pixels of a reference scale with known length (Hsieh et al., 2011). Furthermore, measurements of linear and dimensional morphological traits can be performed using the coordinate system of the image space (Wang et al., 2005). These two approaches also can be used to measure fish body height.

Image analysis techniques have been applied to measure body length and height in several fish species. Navarro et al. (2016) developed an image analysis software (named

IMAFISH_ML) to measure 27 morphometric traits (body lengths, heights, widths, and areas) in gilthead seabream (*Sparus aurata*), meager (*Argyrosomus regius*), and red porgy (*Pagrus pagrus*), using two images of each fish at lateral and dorsal angles. The computer vision algorithms were programmed in MATLAB®. In a study by White et al. (2006), image processing algorithms allowed to classify, based on body shape, seven species of flatfish from the Barents Sea with up to 99.8% reliability and measure body length with a standard deviation of 1.2 mm. Most studies involving the use of digital images to measure morphometric traits focus on fish caught by fishing vessels; very few works use this technique to measure fish for selective breeding.

Digital image analysis also can assist in the collection of traits that are difficult to measure by conventional methods, such as body areas. The real area of a fish body can be determined according to the relationship between the image (number of pixels of a fish region after segmentation) and the real world (reference scale) (Hao et al., 2015). Attributes extracted from images can also be used to predict fish body weight. For instance, Fernandes et al. (2020b) used deep learning methods for image segmentation (background, fish body, and fins) and autonomous measurement of the body area, length, height, and eccentricity in Nile tilapia. The authors found that the body area can be successfully used to predict body ($R^2 = 0.96$) and carcass ($R^2 = 0.95$) weights.

1.3. Future perspectives in phenotyping

In the last decade, several efforts have been made to implement computer vision systems for high-throughput phenotyping in aquaculture. As discussed previously in this chapter, digital image analysis is a non-invasive, low-cost, and fast tool for measuring morphometric traits in several species but has been little applied in fish selective breeding. An additional advantage of the method is that images remain available for later analysis of traits that were not considered at first. There are also, however, some challenges in using computer vision systems in aquaculture, such as (i) the need to remove fish from the culture environment for image acquisition for greater control of image quality (light, contrast, and absence of noise); (ii) images acquired by underwater cameras have low quality because of noise caused by light reflection, water turbidity, and fish movement (curved body and overlapping); and (iii) the models and software developed to apply only to the species under study.

With the rapid development of information, image analysis, and big data technologies, automated phenotyping methods for fish selective breeding programs will soon become a reality. As for future directions, it is proposed that other technologies and imaging sensors should be evaluated in aquaculture, such as 3D imaging sensors already in use for cattle, pig, and poultry analysis. The use of depth sensors based on a structured infrared-light system (e.g., Microsoft Kinect® and Microsoft Azure®) eliminates errors related to variations in animal color and ambient lighting (Kongsro, 2014). Moreover, 3D imaging sensors allow estimating body volume based on depth multiplied by width or area (Kongsro, 2014). Traits determined from 2D or 3D images can be evaluated in the prediction of other traits, such as weights (body and fillet) and yields (fillet and carcass). Although promising, automated phenotyping requires further research for application in aquaculture.

1.4. Objectives

1.4.1. General objective

Use digital image analysis to measure important traits for Nile tilapia breeding programs.

1.4.2. Specific objectives

- Develop a fast and simple method to measure the length, height, and body areas of Nile tilapia by digital image analysis;
- Estimate variance components and genetic parameters for the traits obtained by digital image analysis;
- Apply a machine learning approach to predict body, fillet, and carcass weights and fillet yield of Nile tilapia using traits obtained by digital image analysis.

1.5. References

Bentsen, H.B., Gjerde, B., Nguyen, N.H., Rye, M., Ponzoni, R.W., de Vera, M.S.P., Bolivar, H.L., Velasco, R.R., Danting, J.C., Dionisio, E.E., Longalong, F.M., Reyes, R.A., Abella, T.A., Tayamen, M.M., Eknath, A.E. (2012). Genetic improvement of farmed tilapias: Genetic parameters for body weight at harvest in Nile tilapia (*Oreochromis niloticus*) during five generations of testing in multiple environments. Aquaculture, 338, pp.56-65. https://doi.org/10.1016/j.aquaculture.2012.01.027

Araújo, F.C.T., Oliveira, C.A.L., Campos, E.C., Yoshida, G.M., Lewandowski, V., Todesco, H., Nguyen, N.H., Ribeiro, R.P. (2020). Effects of genotype × environment interaction on the estimation of genetic parameters and gains in Nile tilapia. Journal of Applied Genetics, 61(4), 575-580.https://doi.org/10.1007/s13353-020-00576-2

Embody, G.C., Hayford, C.O. (1925). The advantage of rearing brook trout fingerlings from selected breeders. Transactions of the American Fisheries Society, 55(1), 135-148. https://doi.org/10.1577/1548-8659(1925)55[135:TAORBT]2.0.CO;2

FAO - Food and Agriculture Organization of the United Nations (2020). The state of world fisheries and aquaculture 2020: Sustainability in action. *Food and Agriculture Organization of the United Nations*, Rome.

Fernandes, A.F.A, Dórea, J.R.R., Rosa, G.J.M. (2020a). Image analysis and computer vision applications in animal sciences: an overview. Frontiers in Veterinary Science, 7, 1-10. https://doi.org/10.3389/fvets.2020.551269

Fernandes, A.F.A., Turra, E.M., de Alvarenga, É.R., Passafaro, T.L., Lopes, F.B., Alves, G.F.O., Singh, V., Rosa, G.J.M. (2020b). Deep Learning image segmentation for extraction of fish body measurements and prediction of body weight and carcass traits in Nile tilapia. Computers and Electronics in Agriculture, 170, 105274. https://doi.org/10.1016/j.compag.2020.105274

Fu, G., Yuna, Y. (2022). Phenotyping and phenomics in aquaculture breeding. Aquaculture and Fisheries, 7(2), 140 – 146. https://doi.org/10.1016/j.aaf.2021.07.001

Garcia, A.L.S., de Oliveira, C.A.L., Karim, H.M., Sary, C., Todesco, H., Ribeiro, R.P. (2017). Genetic parameters for growth performance, fillet traits, and fat percentage of male Nile tilapia (*Oreochromis niloticus*). Journal of Applied Genetics, 58(4), 527-533. https://doi.org/10.1007/s13353-017-0413-6

Gjedrem, T. (2012). Genetic improvement for the development of efficient global aquaculture: a personal opinion review. Aquaculture, 344, 12-22. https://doi.org/10.1016/j.aquaculture.2012.03.003

Gjerde, B. (2005). Design of breeding programs. In Selection and Breeding Programs in Aquaculture (pp. 173-195). Springer, Dordrecht.

Gjerde, B., Mengistu, S.B., Ødegård, J., Johansen, H., Altamirano, D.S. (2012). Quantitative genetics of body weight, fillet weight, and fillet yield in Nile tilapia (*Oreochromis niloticus*). Aquaculture, 342, 117-124. https://doi.org/10.1016/j.aquaculture.2012.02.015

Gonzalez, R. C., Woods, R. E. (2018). Digital Image Fundamentals, in: Gonzalez, R. C., Woods, R. E. (Eds.), Digital image processing – fourth edition. Pearson, New York, pp. 47-118.

Gupta, M. V., Acosta, B. O. (2004). From drawing board to dining table: the success story of the GIFT project. NAGA, WorldFish Center Quarterly, 27(3-4), 4-14.

Hao, M., Yu, H., Li, D. (2015). The measurement of fish size by machine vision-a review. In International Conference on Computer and Computing Technologies in Agriculture (pp. 15-32). Springer, Cham.

Hsieh, C.L., Chang, H.Y., Chen, F.H., Liou, J.H., Chang, S.K., Lin, T.T. (2011). A simple and effective digital imaging approach for tuna fish length measurement compatible with fishing operations. Computers and Electronics in Agriculture, 75(1), 44-51. https://doi.org/10.1016/j.compag.2010.09.009

Kongsro, J. (2014). Estimation of pig weight using a Microsoft Kinect prototype imaging system. Computers and Electronics in Agriculture, 109, 32-35. https://doi.org/10.1016/j.compag.2014.08.008

Li, D., Du, L. (2021). Recent advances of deep learning algorithms for aquacultural machine vision systems with emphasis on fish. Artificial Intelligence Review, 1-40. https://doi.org/10.1007/s10462-021-10102-3

Navarro, A., Lee-Montero, I., Santana, D., Henríquez, P., Ferrer, M. A., Morales, A., Soula, M., Badilla, R., Negrín-Báez, D., Zamorano, M.J., Afonso, J. M. (2016). IMAFISH_ML: A fully-automated image analysis software for assessing fish morphometric traits on gilthead seabream (*Sparus aurata* L.), meagre (*Argyrosomus regius*) and red porgy (*Pagrus pagrus*). Computers and Electronics in Agriculture, 121, 66-73. https://doi.org/10.1016/j.compag.2015.11.015

Nguyen, N.H., Khaw, H.L., Ponzoni, R.W., Hamzah, A., Kamaruzzaman, N. (2007). Can sexual dimorphism and body shape be altered in Nile tilapia (*Oreochromis niloticus*) by genetic means?. Aquaculture, 272, S38-S46. https://doi.org/10.1016/j.aquaculture.2007.08.013.

Nguyen, N.H., Ponzoni, R.W., Abu-Bakar, K.R., Hamzah, A., Khaw, H.L., Yee, H.Y. (2010). Correlated response in fillet weight and yield to selection for increased harvest weight in genetically improved farmed tilapia (GIFT strain), *Oreochromis niloticus*. Aquaculture, 305(1-4), 1-5. https://doi.org/10.1016/j.aquaculture.2010.04.007

Oliveira, C.A.L., Ribeiro, R.P., Yoshida, G.M., Kunita, N.M., Rizzato, G.S., Oliveira, S.N., Santos, A.I., Nguyen, N.H. (2016). Correlated changes in body shape after five generations of selection to improve growth rate in a breeding program for Nile tilapia *Oreochromis niloticus* in Brazil. Journal of Applied Genetics, 57(4), 487-493. https://doi.org/10.1007/s13353-016-0338-5

PEIXE BR (2022). Anuário Brasileiro da Piscicultura. *Associação Brasileira da Piscicultura,* São Paulo, Brazil. Available in: https://www.peixebr.com.br/anuario2022, accessed in: March 2, 2022.

Ponzoni, R.W., Hamzah, A., Tan, S., Kamaruzzaman, N. (2005). Genetic parameters and response to selection for live weight in the GIFT strain of Nile tilapia (*Oreochromis niloticus*). Aquaculture, 247(1-4), 203-210. https://doi.org/10.1016/j.aquaculture.2005.02.020

Ponzoni, R.W., Nguyen, N.H., Khaw, H.L., Hamzah, A., Bakar, K.R.A., Yee, H.Y. (2011). Genetic improvement of Nile tilapia (*Oreochromis niloticus*) with special reference to the work conducted by the WorldFish Center with the GIFT strain. Reviews in Aquaculture, 3(1), 27-41. https://doi.org/10.1111/j.1753-5131.2010.01041.x

Popma, T., Masser, M. (1999). Tilapia life history and biology. Southern Regional Aquaculture Center. Publication, (283).

Reis Neto, R.V., Oliveira, C.A.L., Ribeiro, R.P., Freitas, R.T.F., Allaman, I.B., Oliveira, S.N. (2014). Genetic parameters and trends of morphometric traits of GIFT tilapia under selection for weight gain. Scientia Agricola, 71, 259-265. https://doi.org/10.1590/0103-9016-2013-0294

Rutten, M.J., Bovenhuis, H., Komen, H. (2005). Genetic parameters for fillet traits and body measurements in Nile tilapia (*Oreochromis niloticus* L.). Aquaculture, 246(1-4), 125-132. https://doi.org/10.1016/j.aquaculture.2005.01.006

Santos, A.I., Ribeiro, R.P., Vargas, L., Mora, F., Alexandre Filho, L., Fornari, D.C., Oliveira, S.N. (2011). Bayesian genetic parameters for body weight and survival of Nile tilapia farmed in Brazil. Pesquisa Agropecuária Brasileira, 46, 33-43. https://doi.org/10.1590/S0100-204X2011000100005

Thodesen, J., Rye, M., Wang, Y.X., Yang, K.S., Bentsen, H.B., Gjedrem, T. (2011). Genetic improvement of tilapias in China: genetic parameters and selection responses in growth of Nile tilapia (*Oreochromis niloticus*) after six generations of multi-trait selection for growth and fillet yield. Aquaculture, 322, 51-64. https://doi.org/10.1016/j.aquaculture.2011.10.010

Trọng, T.Q., Mulder, H.A., van Arendonk, J.A., Komen, H. (2013). Heritability and genotype by environment interaction estimates for harvest weight, growth rate, and shape of Nile tilapia (*Oreochromis niloticus*) grown in river cage and VAC in Vietnam. Aquaculture, 384, 119-127. https://doi.org/10.1016/j.aquaculture.2012.12.022

Turra, E.M., Toral, F.L.B., de Alvarenga, É.R., Raidan, F.S.S., Fernandes, A.F.A., Alves, G.F.O, Sales, S.C.M., Teixeira, E.A., Manduca, L.G., Brito, T.S., Silva, M.A., da Silva Júnior, A.F., Almeida, L.F.C., Santos, C.R., e Silva, M.A. (2016). Genotype × environment interaction for growth traits of Nile tilapia in biofloc technology, recirculating water and Cage systems. Aquaculture, 460, 98-104. https://doi.org/10.1016/j.aquaculture.2016.04.020

Ventura, R.V., e Silva, F.F., Yáñez, J.M., Brito, L.F. (2020). Opportunities and challenges of phenomics applied to livestock and aquaculture breeding in South America. Animal Frontiers:

The Review Magazine of Animal Agriculture, 10(2), 45 - 52. https://doi.org/10.1093/af/vfaa008

Wang, L., Zhang, Y., Feng, J. (2005). On the Euclidean distance of images. IEEE transactions on pattern analysis and machine intelligence, 27(8), 1334-1339.

White, D.J., Svellingen, C., Strachan, N.J. (2006). Automated measurement of species and length of fish by computer vision. Fisheries Research, 80(2-3), 203-210. https://doi.org/10.1016/j.fishres.2006.04.009

Yáñez, J.M., Joshi, R., Yoshida, G.M. (2020). Genomics to accelerate genetic improvement in tilapia. Animal Genetics, 51(5), 658-674. https://doi.org/10.1111/age.12989

Yoshida, G.M., de Oliveira, C.A.L., Campos, E.C., Todesco, H., Araújo, F.C., Karin, H.M., Zardin, A.M., Bezerra Júnior, J.S., Filho, L.A., Vargas, L., Ribeiro, R.P. (2021). A breeding program for Nile tilapia in Brazil: Results from nine generations of selection to increase the growth rate in cages. Journal of Animal Breeding and Genetics. Early View. https://doi.org/10.1111/jbg.12650

CHAPTER 2:

Estimation of genetic parameters for body areas in Nile tilapia measured by digital image analysis

Running head: Genetic parameters for body areas by digital image

Alex Júnio da Silva Cardoso[1], Carlos Antonio Lopes de Oliveira[2,*], Eric Costa Campos[2], Ricardo Pereira Ribeiro[2], Gutierrez José de Freitas Assis[1], Fabyano Fonseca e Silva[1]

[1] Department of Animal Science, Universidade Federal de Viçosa, Campus Universitário, s/n, Viçosa, Minas Gerais, 36570-900, Brazil.

[2] Department of Animal Science, Universidade Estadual de Maringá, Av. Colombo, 5790, Jd. Universitário, Maringá, Paraná, 87020-900, Brazil.

* Corresponding author: E-mail address: caloliveira@uem.br

Manuscript published in Journal of Animal Breeding and Genetics

Received: 29 December 2020

Accepted: 04 April 2021

DOI: 10.1111/jbg.12551

2.1. Abstract

Digital image analysis is a practical, non-invasive, and relatively low-cost tool that may assist in the evaluation of body traits in Nile tilapia, being particularly useful for assessing difficult-to-measure variables, such as body areas. In this study, we aimed to estimate variance components and genetic parameters for body areas of Nile tilapia obtained by digital images. The data set comprised body weight (BW) records of 1,917 pond-reared fish at 366 days of age. Of this total, 656 animals were photographed and subjected to image analysis of trunk area (TA), head area (HA), caudal fin area (CFA), and fillet area (FA). Heritabilities and genetic correlations were estimated through multiple-trait models based on Bayesian inference. Heritability estimates for BW, TA, HA, CFA and FA were 0.25, 0.23, 0.26, 0.21 and 0.25, respectively. Genetic correlations between the traits were high and positive, ranging from 0.70 to 0.98. We highlight the genetic correlation between BW and TA (r_G = 0.98) and FA (r_G = 0.97). In view of the observed results, it can be concluded that trunk and fillet areas obtained by digital image analysis can lead to indirect genetic gains in weight and other body areas. In addition, the areas studied have potential as a selection criterion and may assist in studies on changes in the body shape in Nile tilapia.

Keywords: genetic correlation, genetic gain, heritability estimate, machine vision

2.2. Introduction

Aquaculture has been growing linearly since 1950 mainly because of the increase in fish consumption by humans (FAO, 2020). The production of farmed fish for food reached 54.3 million tonnes in 2018, which represents 47% of global aquaculture production (FAO, 2020). Nile tilapia (*Oreochromis niloticus*) ranks third among the major fish species produced worldwide, with 4.5 million tonnes produced in 2018 (accounting for 8.3% of farmed fish production) (FAO, 2020). In part, this success is due to breeding programmes that allow effective genetic selection for economically important traits.

Generally, Nile tilapia breeding programmes adopt body weight and/or daily weight gain as the main selection criteria to improve growth rate and reduce age at harvest (Oliveira et al., 2016; Ponzoni et al., 2005; Thodesen et al., 2011). In addition, morphometric traits such as body length, width and height are routinely recorded and used in breeding programmes to select for body shape, as these parameters tend to be highly correlated with growth-related traits (Blonk et al., 2010; Nguyen et al., 2007; Oliveira et al., 2016; Omasaki et al., 2016; Trong et al., 2013). Another morphometric trait with great potential to be exploited in breeding programmes is body area. This trait encompasses the association between growth and body shape measures simultaneously, thus improving the production and appearance of fish according to market requirements. However, body area is difficult to measure by devices conventionally used to obtain phenotypic data in Nile tilapia breeding programmes, as evidenced by the small amount of information on the genetic parameters of this trait.

Digital image analysis is a suitable, non-invasive and practical alternative to obtain body measurements at a relatively low cost. Digital images allow assessing morphometric data of several traits of interest (Hao et al., 2016). Fernandes et al. (2020) demonstrated the effectiveness of a computer vision system in measuring morphometric traits of Nile tilapia, such as body area, length, width and eccentricity; these data were applied to predict body and carcass weight, with good results. Traditional morphometric traits can be evaluated by digital images (Blonk et al., 2010; Hao et al., 2016; Omasaki et al., 2016) and used to infer body shape, as performed in previous geometric morphometry studies (Costa et al., 2010; Prchal et al., 2018). In general, the accuracy of machine-based vision methods is 10% higher than that of traditional methods (Hao et al., 2016). In this study, we aimed to estimate variance components and genetic parameters for body areas of Nile tilapia obtained by digital images.

2.3. Material and methods

2.3.1. Sample population

This study assessed animals from the Nile tilapia breeding programme (Tilamax strain) of the State University of Maringá (UEM), Paraná State, Brazil. The programme was established in 2006 from 30 families of Genetically Improved Farmed Tilapia (GIFT) imported from the WorldFish Center (Malaysia). The fish used were from the 10th generation. Pedigree data were composed by 9,559 animals born between 2015 and 2018, constituting a total of four generations. The pedigree contained 209 sires and 287 dams.

All animal handling procedures were performed in accordance with ethical standards and were approved by the Animal Ethics Committee of UEM (protocol no. 9452160720). Fish reproduction and larvae and fingerling cultivation were performed at the Experimental Aquaculture Station of UEM, located in Floriano, Paraná, Brazil (23°31′08.3″S 52°02′20.5″W). The reproduction period lasted from October to December 2018. The population had been produced by natural mating of 58 sires and 78 dams. Reproductive females were kept in individual hapas (1 m^3) and mated with a male. Once a week the eggs were collected and transferred to incubators equipped with a water recirculation system, water heaters set at 27 °C, and a thermostat. Spawned females were replaced and/or males were transferred to other hapas. In total, 83 families of full- and half-sibs were formed.

After yolk sac absorption, progenies were transferred to individual hapas, suspended at different sites of a same pond. Fish were fed a starter diet containing 42% crude protein twice a day. When the fish reached a mean weight of 20 g, they were individually weighed and tagged with passive integrated transponder tags. After tagging, the fry was transferred to Acqua Sul Piscicultura®, a fish farm in Ilhota, Santa Catarina State, Brazil (26°51′7.5″S 48°51′37.8″W). Performance tests were conducted between March (autumn) and November (spring) 2019 in a 2000 m^2 pond containing 1 fish/m^2. All weight and image data used in the present study were provided by the company.

2.3.2. Image acquisition and measured traits

At the final performance test, 1,917 fish were weighed and 656 were photographed for image analysis (mean of seven fish/family). Photographs were acquired by using a digital camera with 1,920 × 1,080 pixels resolution and an 84.1 × 53.8° field of view (Microsoft®). The camera was fixed at a height of 0.5 m on a support and connected to a computer (Intel® Core i5-7200 U, 2.5 GHz and 8 GB RAM). A blue surface was used as background to improve image contrast. Image identification was the tag number of fish.

Trunk area, head area, caudal fin area and fillet area (depicted in Figure 1) were measured with ImageJ free software (Schneider et al., 2012). Images were analysed and calibrated individually for conversion from pixels to centimeters. The tool “Polygon selection” was used to demarcate the areas of interest and avoid overlapping information.

2.3.3. Statistical analysis

The following general multiple-trait Bayesian model (Sorensen & Gianola, 2002) was fitted to the data (Equation 1):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{c} + \mathbf{Z}\mathbf{a} + \mathbf{e} \tag{1}$$

where $\mathbf{y}$ is the vector of phenotypic observations for all traits, assumed as $\mathbf{y}|\boldsymbol{\beta}, \mathbf{c}, \mathbf{a}, \mathbf{C_0}, \mathbf{G_0}, \mathbf{R_0} \sim \mathrm{N}\ (\mathbf{X}\boldsymbol{\beta} + \mathbf{Wc} + \mathbf{Za}, \mathbf{R_0} \otimes \mathbf{I})$, being $\otimes$ the Kronecker product (direct product) operator; $\mathbf{X}$ is the incidence matrix of systematic effects; $\boldsymbol{\beta}$ is the vector of systematic effects, with prior distribution assumed as $\boldsymbol{\beta} \sim \mathrm{N}\ (\mathbf{0},\ \sum_{\boldsymbol{\beta}} \otimes \mathbf{I})$, being $\sum_{\boldsymbol{\beta}}$ the covariance matrix (in this case, it is given by a diagonal matrix), with large variances (10^{+6}) to represent vague prior knowledge, and $\mathbf{I}$ an identity matrix; $\mathbf{W}$ is the incidence matrix of common family environmental effects; $\mathbf{c}$ is the random vector of common environmental effects related to full-sib families, with prior distribution given by $\mathbf{c}|\mathbf{C_0} \sim \mathrm{N}\ (\mathbf{0}, \mathbf{C_0} \otimes \mathbf{I})$, being $\mathbf{C_0}$ the covariance matrix of common family environmental effects; $\mathbf{Z}$ is the incidence matrix of additive genetic effects; $\mathbf{a}$ is the random vector of additive genetic effects, with prior distribution given by $\mathbf{a}|\mathbf{G_0}, \mathbf{A} \sim \mathrm{N}\ (\mathbf{0}, \mathbf{G_0} \otimes \mathbf{A})$, being $\mathbf{A}$ the additive relationship matrix based on fish genealogical information and $\mathbf{G_0}$ the additive genetic covariance matrix; $\mathbf{e}$ is the random vector of residuals, distributed as $\mathbf{e}|\mathbf{R_0} \sim \mathrm{N}\ (\mathbf{0}, \mathbf{R_0} \otimes \mathbf{I})$, being $\mathbf{R_0}$ the residual covariance matrix. Furthermore, it was assumed that $\mathbf{C_0}$, $\mathbf{G_0}$ and $\mathbf{R_0}$ follow an inverse Wishart prior distribution, WI (v, $\mathbf{V}$), with hyperparameters v = 2 and $\mathbf{V}$ given by a diagonal matrix.

Different models derived from Equation (1) were compared (Table 2). The additive genetic effect ($\mathbf{a}$) was considered significant and included in all models, whereas the common family environmental effect ($\mathbf{c}$) was tested for significance. Additionally, sex, age (linear and quadratic covariate) and tagging weight (linear and quadratic covariate) effects were tested as systematic effects ($\boldsymbol{\beta}$). Goodness of fit was assessed by the Deviance Information Criterion (DIC), proposed by Spiegelhalter et al. (2002). Thus, the lower the DIC value, the better the goodness of fit of the model to experimental data.

Variance components and genetic parameters were estimated using the Gibbs sampling algorithm of GIBBS1F90 software (Misztal et al., 2016). A total of 5 million Markov Chain Monte Carlo (MCMC) samples were generated, with a burn-in period (discarded samples) of

500,000 iterations and a sample interval ("thin") of 200 iterations. An effective sample size of 22,500 iterations was used to obtain the mean posterior distribution (Bayesian estimate) of variance components and genetic parameters. The posterior standard deviation (PSD) and 95% highest posterior density intervals (HPD) were also calculated. Posterior distributions for the heritability of trait i (h_i^2), for the common family environmental effect as a proportion of phenotypic variance for trait i (c_i^2), and for the genetic correlation between traits i and j (r_{ij}) were obtained from Gibbs sampling iterations (k), as follows (Equations 2, 3 and 4):

$$h_i^{2(k)} = \frac{\sigma_{ai}^{2(k)}}{\left(\sigma_{ai}^{2(k)} + \sigma_{ci}^{2(k)} + \sigma_{ei}^{2(k)}\right)} \qquad (2)$$

$$c_i^{2(k)} = \frac{\sigma_{ci}^{2(k)}}{\left(\sigma_{ai}^{2(k)} + \sigma_{ci}^{2(k)} + \sigma_{ei}^{2(k)}\right)} \qquad (3)$$

$$r_{ij}^{(k)} = \frac{\sigma_{aij}^{(k)}}{\sqrt{\sigma_{ai}^{2(k)}\,\sigma_{aj}^{2(k)}}} \qquad (4)$$

where $\sigma_{ai}^{2(k)}$, $\sigma_{ci}^{2(k)}$ and $\sigma_{ei}^{2(k)}$ represent the additive genetic variance, common family environmental variance and residual variance, respectively, of trait i in iteration k and $\sigma_{aij}^{(k)}$ is the additive genetic covariance between traits i and j in iteration k.

Convergence of MCMC chains was assessed for each trait by the Geweke test using POSTGIBBSF90 software (Misztal et al., 2016) and by the convergence criterion of Raftery & Lewis (1992) using the package of convergence diagnosis and output analysis for MCMC (CODA) (Plummer et al., 2006) in R software (R Core Team, 2020).

2.4. Results

2.4.1. Descriptive statistics

The mean body weight at 366 days of age was 414.86 g (Table 1). Fillet area accounted for 66.90% of the trunk area, which does not include head or caudal fin areas (Table 1).

2.4.2. Model comparisons

The chosen model, which had the lowest DIC, considered additive genetic and common family/environment as random effects and sex, weight at tagging (linear covariate) and age (linear covariate) as systematic effects (Table 2). The convergence of MCMC chains was achieved by Geweke and Raftery & Lewis criteria.

2.4.3. Genetic parameters

Estimates of common family environmental variance as a proportion of phenotypic variation (c^2) were low for all traits, ranging from 0.09 to 0.12. However, when considering credibility intervals, we observed a relevant contribution of this effect, thereby justifying its inclusion in the model (Table 3). Heritability estimates were classified as moderate, with values of 0.25 for body weight and 0.21 – 0.26 for body areas measured using digital images (Table 3).

Genetic correlations between weight and areas were high and positive (Table 4), particularly for trunk area (r_G = 0.98) and fillet area (r_G = 0.97). Genetic correlations between areas were also high and positive (Table 4); for instance, the correlation between trunk area and fillet area was 0.98. Phenotypic correlations were quite similar to genetic correlations (Table 4).

2.5. Discussion

2.5.1. Chosen model

In Nile tilapia breeding programmes, full-sib families are kept in a common environment (individual hapa by family) until they reach tagging weight; therefore, it is usually necessary to include shared family environmental effects (**c**) in the model. The observed c^2 values agree with those reported by Oliveira et al. (2016) for the same species. Although the magnitude of c^2 was low, common family environmental effects were considered, as, according to Pante et al. (2002), when the effect is omitted from analysis, the model can overestimate additive genetic variances, thereby reducing selection accuracy. Nguyen et al. (2007) observed a reduction of 17% in selection accuracy for weight at harvest when common family environmental effects were omitted. Common family environmental effects generally decrease as fish age increases (Turra et al., 2012), which explains the low values found in the present study. In addition, full-sib families were stocked in same pond and under same conditions after tagging.

Due to the reproduction period considered in the present study, the fish had different ages at harvest time, as well as different weight at tagging. When we do not consider age and weight at tagging as a covariate in the model, we observe an increase in the DIC value, therefore, these covariables had a significant effect on the genetic parameter’s estimation. It's known that body weight increases with increasing age (Bentsen et al., 2012), which may influence the estimating genetic parameters and the fish genetic merit. The fish in the present study went through a period of low temperatures, typical of the south region of Brazil, thus compromising their growth. At the harvest time, the fish were probably in a compensatory gain period, which explain the significant linear effect for the age covariate. Age is frequently considered a covariate in the estimation of genetic parameters for growth and morphometric traits in Nile tilapia (Bentsen et al., 2012; Oliveira et al., 2016; Reis Neto et al., 2014). Garcia et al. (2017) when considering weight at tagging as covariable, found a better the goodness of model's fit to experimental data, in a study with Nile tilapia.

The use of the multi-trait model allows considering possible correlations between all the traits studied, increasing the precision in genetic parameters estimation (Thompson & Meyer, 1986). Due to the relatively small-samples number, we used Bayesian inference by Markov Chain Monte Carlo (MCMC) sampling methods. The MCMC method provides random samples of the posterior joint and marginal distributions for variance components and genetic parameters, which allows to directly estimate the mean, standard deviation and credibility intervals (Blasco, 2001). However, a convergence analysis must be performed, since the inferences need to be based only on stationary marginal posterior distributions. In the present study, the burn-in period and sample interval ("thin") used were sufficient to reach the equilibrium condition, a fact confirmed by Geweke and Raftery & Lewis criteria.

2.5.2. Genetic parameters

Heritability estimates varies from moderate to high for morphometric traits in Nile tilapia (Fernandes et al., 2015; Nguyen et al., 2007; Reis Neto et al., 2014; Rutten et al., 2005a). In the referred studies, the morphometric traits evaluated (total and standard length, height and width) were measured manually using rulers. To the best of our knowledge, the present study is the first approach to the estimation of heritability in Nile tilapia using body areas measured by digital image analysis. Heritability estimates for trunk area (0.23), head area (0.26), caudal fin area (0.21) and fillet area (0.25) indicate that these traits are suitable for body shape selection. Reis Neto et al. (2014) estimated the heritability of total body area to be 0.30; in their study, however, total body area was calculated as the product of standard length and body depth, measured at the beginning of the dorsal fin. This led to overestimation of the true body area. In general, the use of digital images allows obtaining accurate body areas, with the added benefit of being an easy and low-cost technique (Fernandes et al., 2020). Therefore, digital imagery is an extremely relevant tool that can be included in aquaculture breeding programmes for the adoption of new selection criteria.

The heritability estimates for body weight found here are consistent with those observed in several studies on Nile tilapia (Garcia et al., 2017; Hamzah et al., 2017; Rutten et al., 2005a; Rutten et al., 2005b), although, in other reports, estimates ranged from 0.10 to 0.84 (Hamzah et al., 2017; He et al., 2018; Oliveira et al., 2016; Rutten et al., 2005a; Rutten et al., 2005b; Turra et al., 2016). Various factors contribute to the high heterogeneity of heritability estimates for body weight in Nile tilapia, such as the number of traits considered for genetic parameter estimation, whether the model is based on single or multiple traits, and the genetic and environmental correlations between traits. Multiple-trait models are generally more accurate than single-trait models, as the former incorporate covariances between traits (Mrode, 2014). Other factors, such as common family environmental effects (Pante et al., 2002), characteristics of the fish farming environment (Turra et al., 2016), sex and age at measurement (He et al., 2018; Rutten et al., 2005b), can also contribute to such differences.

Genetic correlations between growth and morphometric traits are typically high and positive (Campos et al., 2020; Fernandes et al., 2015; Reis Neto et al., 2014; Trong et al., 2013). In the present study, the high positive correlations between weight and body areas indicate that it is possible to select fish with high growth rates and larger fillet areas and/or smaller head and/or fin areas. Maximization of edible parts can contribute to reducing environmental impacts of fish production while increasing profitability.

Selection for growth may result in changes to body shape, as reported by Oliveira et al. (2016) for Nile tilapia. In view of the high genetic and phenotypic correlation between body weight and trunk and fillet areas, we suggest the use of these areas (determined by digital image analysis) as selection criteria to change body shape, increase fillet yield and help farmers meet the demands of the consumer market. In regions where Nile tilapia is culturally consumed whole and fresh, body shape traits can be decisive in consumer preference. For Nguyen et al. (2007), round fish may be perceived as having a greater amount of edible meat, and thus more attractive to consumers. In this sense, fish with a larger area of trunk and fillet has greater body weight and larger area of edible meat. However, studies still need to be performed out to verify the effectiveness of using these areas as an indirect measure for selecting the weight and yield of fillet in Nile tilapia.

2.5.3. Digital imagery for body shape studies

Body shape is often considered in genetic studies; however, this trait must be determined through indirect measurements, which can be difficult. Body shape is calculated from body dimensions and their proportions, including as a simple ratio of height to length or length to other body parts and as ellipticity of midsagittal, transverse and frontal planes (Blonk et al., 2010; Nguyen et al., 2007; Oliveira et al., 2016; Omasaki et al., 2016; Trong et al., 2013). Body shape has also been defined as a condition factor (weight/length^3) (Elvingson & Johansson, 1993; Kause et al., 2003; Trong et al., 2013) or classified by subjective visual examination (Kause et al., 2003).

Digital image analysis can be used to measure traits commonly evaluated in fish breeding programs (Blonk et al., 2010; Omasaki et al., 2016) or new traits, such as body areas, contributing to selection for body shape. Analysis of geometric morphometry by digital imagery allows visualization and quantification of differences in body shape between individuals through reference points on the fish body. Costa et al. (2010) assessed the geometric morphometry of European sea bass (*Dicentrarchus labrax*) using digital images to investigate genetic and environmental contributions to body shape and observed a high genetic variation in this trait. In vivo morphological predictors, determined using digital images, contributed to improving the slaughter yields of common carp (*Cyprinus carpio*) (Prchal et al., 2018).

2.6. Conclusion

Moderate heritability estimates were observed for body areas measured using digital images, indicating the potential of these traits as selection criteria for body shape. Selection for trunk and fillet areas can lead to indirect genetic gains in weight and other body areas.

2.7. Acknowledgments

We thank the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) for granting a scholarship to the first author (process no. 148102/2018-2). We are grateful to the PeixeGen Research Group (which specializes in management, breeding and molecular genetics in freshwater fish farming), Acqua Sul Piscicultura®, and all of the Aquaculture Station of the State University of Maringá for their assistance in this study.

2.8. Conflict of interest statement

The authors declare they have no competing interests.

2.9. Data availability statement

The data that support the findings of this study are presented in the paper.

2.10. ORCID

Alex Júnio da Silva Cardoso: https://orcid.org/0000-0002-5841-3882

Carlos Antonio Lopes de Oliveira: https://orcid.org/0000-0003-2504-3261

Eric Costa Campos: https://orcid.org/0000-0002-3503-0889

Ricardo Pereira Ribeiro: https://orcid.org/0000-0001-7752-3692

Gutierrez José de Freitas Assis: https://orcid.org/0000-0002-0156-9376

Fabyano Fonseca e Silva: https://orcid.org/0000-0001-9536-1113

2.11. References

Bentsen, H. B., Gjerde, B., Nguyen, N. H., Rye, M., Ponzoni, R. W., de Vera, M. S. P., Bolivar, H. L., Velasco, R. R., Danting, J. C., Dionisio, E. E., Longalong, F. M., Reyes, R. A., Abella, T. A., Tayamen, M. M., & Eknath, A. E. (2012). Genetic improvement of farmed tilapias: genetic parameters for body weight at harvest in Nile tilapia (*Oreochromis niloticus*) during five generations of testing in multiple environments. *Aquaculture*, 338, 56-65. https://doi.org/10.1016/j.aquaculture.2012.01.027.

Blasco, A. (2001). The Bayesian controversy in animal breeding. *Journal of Animal Science*, 79, 2023-2046. https://doi.org/10.2527/2001.7982023x.

Blonk, R. J. W., Komen, J., Tenghe, A., Kamstra, A., & Van Arendonk, J. A. M. (2010). Heritability of shape in common sole, *Solea solea*, estimated from image analysis data. *Aquaculture*, 307, 6-11. https://doi.org/10.1016/j.aquaculture.2010.06.025.

Campos, E. C., Oliveira, C. A. L., Araújo, F. C. T., Todesco, H., Souza, F. N., Rossi, R. M., Fornari, D. C., & Ribeiro, R. P. (2020). Genetic parameters and response to selection for growth in tambaqui. *animal*, 1- 9. https://doi.org/10.1017/S1751731120000488.

Costa, C., Vandeputte, M., Antonucci, F., Boglione, C., Menesatti, P., Cenadelli, S., Parati, K., Chavanne, H., & Chatain, B. (2010). Genetic and environmental influences on shape variation in the European sea bass (*Dicentrarchus labrax*). *Biological Journal of the Linnean Society*, 101, 427-436. https://doi.org/10.1111/j.1095-8312.2010.01512.x.

Elvingson, P., & Johansson, K. (1993). Genetic and environmental components of variation in body traits of rainbow trout (*Oncorhynchus mykiss*) in relation to age. *Aquaculture*, 118, 191-204. https://doi.org/10.1016/0044-8486(93)90456-9.

FAO (2020). The State of World Fisheries and Aquaculture 2020 - sustainability in action. Food and Agriculture Organization of the United Nations, Rome.

Fernandes, A. F. A., Silva, M. A., de Alvarenga, E. R., Teixeira, E. A., Junior, A. F. S., Alves, G. F. O., Salles, S. C. M., Manduca, L. G., & Turra, E. M. (2015). Morphometric traits as selection criteria for carcass yield and body weight in Nile tilapia (*Oreochromis niloticus* L.) at five ages. *Aquaculture*, 446, 303-309. https://doi.org/10.1016/j.aquaculture.2015.05.009.

Fernandes, A. F. A., Turra, E. M., Alvarenga, É. R., Passafaro, T. L., Lopes, F. B., Alves, G. F. O., Singh, V., & Rosa, G. J. M. (2020). Deep Learning image segmentation for extraction of fish body measurements and prediction of body weight and carcass traits in Nile tilapia. *Computers and Electronics in Agriculture*, 170, 105274. https://doi.org/10.1016/j.compag.2020.105274.

Garcia, A. L. S., Oliveira, C. A. L., Karim, H. M., Sary, C., Todesco, H., & Ribeiro, R. P. (2017). Genetic parameters for growth performance, fillet traits, and fat percentage of male Nile tilapia (*Oreochromis niloticus*). *Journal of Applied Genetics*, 58, 527-533. https://doi.org/10.1007/s13353-017-0413-6.

Hamzah, A., Mekkawy, W., Khaw, H. L., Nguyen, N. H., Yee, H. Y., Abu Bakar, K. R., Nor, S. A. M., & Ponzoni, R.W. (2017). Genetic parameters for survival during the grow-out period in the GIFT strain of Nile tilapia (*Oreochromis niloticus*) and correlated response to selection for harvest weight. *Aquaculture Research*, 48, 47-55. https://doi.org/10.1111/are.12859.

Hao, M., Yu, H., & Li, D. (2016). The Measurement of Fish Size by Machine Vision - A Review, in: Li D., Li Z. (Eds), Computer and Computing Technologies in Agriculture IX. CCTA 2015. IFIP Advances in Information and Communication Technology, vol 479. Springer, Cham. https://doi.org/10.1007/978-3-319-48354-2_2.

He, J., Zhao, Y., Zhao, J., Gao, J., Xu, P., & Yang, R. (2018). Random regression analysis for body weights and main morphological traits in genetically improved farmed tilapia (*Oreochromis niloticus*). *Journal of Applied Genetics*, 59, 99-107. https://doi.org/10.1007/s13353-018-0428-7.

Kause, A., Ritola, O., Paananen, T., Eskelinen, U., & Mäntysaari, E. (2003). Big and beautiful? Quantitative genetic parameters for appearance of large rainbow trout. *Journal of Fish Biology*, 62, 610-622. https://doi.org/10.1046/j.1095-8649.2003.00051.x.

Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., & Vitezica, Z. (2016). Manual for BLUPF90 family of programs. University of Georgia, Athens, GA, USA.

Mrode, R. A. (2014). Best Linear Unbiased Prediction of Breeding Value: Multivariate Animal Models, in R.A. Mrode (Ed.), Linear models for the prediction of animal breeding values, 3rd edition. Cabi, United Kingdom, pp.70 – 86.

Nguyen, N. H., Khaw, H. L., Ponzoni, R. W., Hamzah, A., & Kamaruzzaman, N. (2007). Can sexual dimorphism and body shape be altered in Nile tilapia (*Oreochromis niloticus*) by genetic means?. *Aquaculture*, 272, S38-S46. https://doi.org/10.1016/j.aquaculture.2007.08.013.

Oliveira, C. A. L., Ribeiro, R. P., Yoshida, G. M., Kunita, N. M., Rizzato, G. S., Oliveira, S. N., Santos, A. I., & Nguyen, N. H. (2016). Correlated changes in body shape after five generations of selection to improve growth rate in a breeding program for Nile tilapia *Oreochromis niloticus* in Brazil. *Journal of Applied Genetics*, 57, 487-493. https://doi.org/10.1007/s13353-016-0338-5.

Omasaki, S. K., Charo-Karisa, H., Kahi, A. K., & Komen, H. (2016). Genotype by environment interaction for harvest weight, growth rate and shape between monosex and mixed sex Nile tilapia (*Oreochromis niloticus*). *Aquaculture*, 458, 75-81. https://doi.org/10.1016/j.aquaculture.2016.02.033.

Pante, M. J. R., Gjerde, B., McMillan, I., & Misztal, I. (2002). Estimation of additive and dominance genetic variances for body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. *Aquaculture*, 204, 383-392. https://doi.org/10.1016/S0044-8486(01)00825-0.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. *R news*, 6, 7-11.

Ponzoni, R. W., Hamzah, A., Tan, S., & Kamaruzzaman, N. (2005). Genetic parameters and response to selection for live weight in the GIFT strain of Nile tilapia (*Oreochromis niloticus*). *Aquaculture*, 247, 203-210. https://doi.org/10.1016/j.aquaculture.2005.02.020.

Prchal, M., Bugeon, J., Vandeputte, M., Kause, A., Vergnet, A., Zhao, J., Gela, D., Genestout, L., Bestin, A., Haffray, P., & Kocour, M. (2018). Potential for genetic improvement of the main slaughter yields in common carp with in vivo morphological predictors. *Frontiers in genetics*, 9, 283. https://doi.org/10.3389/fgene.2018.00283.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Raftery, A. E., & Lewis, S. (1992). How many iterations in the Gibbs sampler?, in: Bernardo, J. M., Berger, J., Dawid, A. P., Smith, A. F. M. (Eds.), Bayesian Statistics 4. Oxford: Oxford University Press, pp. 763–773.

Reis Neto, R. V., Oliveira, C. A. L., Ribeiro, R. P., Freitas, R. T. F., Allaman, I. B., & Oliveira, S. N. (2014). Genetic parameters and trends of morphometric traits of GIFT tilapia under selection for weight gain. *Scientia Agricola*, 71, 259-265. https://doi.org/10.1590/0103-9016-2013-0294.

Rutten, M. J., Bovenhuis, H., & Komen, H. (2005a). Genetic parameters for fillet traits and body measurements in Nile tilapia (*Oreochromis niloticus* L.). *Aquaculture*, 246, 125-132. https://doi.org/10.1016/j.aquaculture.2005.01.006.

Rutten, M. J., Komen, H., & Bovenhuis, H. (2005b). Longitudinal genetic analysis of Nile tilapia (*Oreochromis niloticus* L.) body weight using a random regression model. *Aquaculture*, 246, 101-113. https://doi.org/10.1016/j.aquaculture.2004.12.020.

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods*, 9, 671-675. https://doi.org/10.1038/nmeth.2089.

Sorensen, D., & Gianola, D. (2002). Likelihood, Bayesian and MCMC methods in quantitative genetics. New York: Springer-Verlag.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series b (statistical methodology)*, 64, 583–639. https://doi.org/10.1111/1467-9868.00353.

Thodesen, J., Rye, M., Wang, Y. X., Yang, K. S., Bentsen, H. B., & Gjedrem, T. (2011). Genetic improvement of tilapias in China: genetic parameters and selection responses in growth of Nile tilapia (*Oreochromis niloticus*) after six generations of multi-trait selection for growth and fillet yield. *Aquaculture*, 322, 51-64. https://doi.org/10.1016/j.aquaculture.2011.10.010.

Thompson, R., & Meyer, K. (1986). A review of theoretical aspects in the estimation of breeding values for multi-trait selection. *Livestock Production Science*, 15, 299-313. https://doi.org/10.1016/0301-6226(86)90071-0.

Trong, T. Q., Mulder, H. A., van Arendonk, J. A., & Komen, H. (2013). Heritability and genotype by environment interaction estimates for harvest weight, growth rate, and shape of Nile tilapia (*Oreochromis niloticus*) grown in river cage and VAC in Vietnam. *Aquaculture*, 384, 119-127. https://doi.org/10.1016/j.aquaculture.2012.12.022.

Turra, E. M., Oliveira, D. A. A., Valente, B. D., Teixeira, E. A, Prado, S. A, Melo, D. C., Fernandes, A. F. A., Alvarenga, E. R., & e Silva, M. D. A. (2012). Estimation of genetic parameters for body weights of Nile tilapia *Oreochromis niloticus* using random regression models. *Aquaculture*, 354, 31-37. https://doi.org/10.1016/j.aquaculture.2012.04.035.

Turra, E. M., Toral, F. L. B., Alvarenga, É. R., Raidan, F. S. S., Fernandes, A. F. A, Alves, G. F. O., Sales, S. C. M., Teixeira, E. A., Manduca, L. G., Brito, T. S., Silva, M. A., Silva Junior, A. F., Almeida, L. F. C., Santos, C. R., & Silva, M. A. (2016). Genotype × environment interaction for growth traits of Nile tilapia in biofloc technology, recirculating water and Cage systems. *Aquaculture*, 460, 98-104. https://doi.org/10.1016/j.aquaculture.2016.04.020.

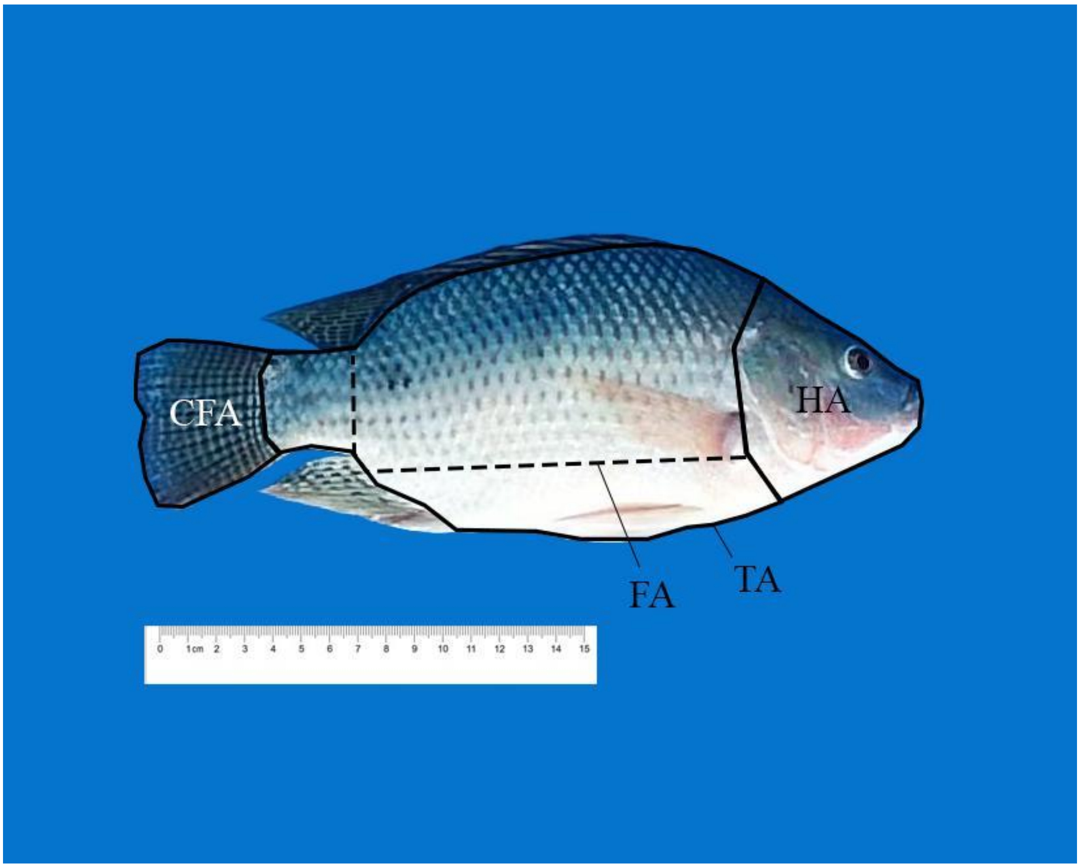

Figure 1. Representation of areas measured by digital image analysis. TA, trunk area; HA, head area; CFA, caudal fin area; FA, fillet area.

Table 1. Descriptive statistics for the studied traits.

Trait	*n*			Mean	Min	Max	SD
	Male	Female	Total				
Age (days)	1,020	897	1,917	366.37	337.00	393.00	17.86
Weight at tagging (g)	1,020	897	1,917	21.09	5.00	81.00	9.45
Body weight (g)	1,020	897	1,917	414.86	111.00	753.00	98.86
Trunk area (cm^2)	338	318	656	113.34	49.16	171.52	19.84
Head area (cm^2)	338	318	656	30.92	15.41	44.61	4.30
Caudal fin area (cm^2)	338	318	656	23.07	8.81	38.26	4.86
Fillet area (cm^2)	338	318	656	75.82	32.58	118.24	13.24

Abbreviations: *n*, number of individuals; Min, minimum; Max, maximum; SD, standard deviation.

Table 2. Deviance information criterion (DIC) values for the compared models.

Models	a	c	sex	Age	Age^2	WT	WT^2	DIC
M1	✓	✓	✓	✓	-	✓	-	46,062.564[a]
M2	✓	-	✓	✓	-	✓	-	46,367.800
M3	✓	✓	-	✓	-	✓	-	47,555.724
M4	✓	✓	✓	-	-	✓	-	46,508.702
M5	✓	✓	✓	✓	-	-	-	46,709.194
M6	✓	✓	✓	✓	✓	✓	-	46,779.814
M7	✓	✓	✓	✓	-	✓	✓	46,724.729
M8	✓	✓	✓	✓	✓	✓	✓	46,269.679

Abbreviations: a, additive genetic effect; c, common family environmental effect; Age, linear effect of age; Age^2, quadratic effect of age; WT, linear effect of tagging weight; WT^2, quadratic effect of tagging weight; ✓, significant (included in the model); -, effect not included.

[a] Best model according to DIC.

Table 3. Posterior mean estimate, posterior standard deviation (superscript in parentheses), and 95% credibility intervals (below the estimate, in parentheses) of variance components and genetic parameters for Nile tilapia weight and body areas measured by digital image analysis.

Trait	σ_a^2	σ_c^2	σ_e^2	c^2	h^2
BW	1,851.00 $^{(697.71)}$	766.18 $^{(295.82)}$	4,835.30 $^{(400.04)}$	0.10 $^{(0.04)}$	0.25 $^{(0.09)}$
	(568.80, 3188.00)	(262.30, 1369.00)	(4,026.00, 5,572.00)	(0.04, 0.18)	(0.09, 0.41)
TA	71.38 $^{(27.03)}$	34.60 $^{(12.92)}$	201.83 $^{(16.53)}$	0.11 $^{(0.04)}$	0.23 $^{(0.08)}$
	(23.46, 125.00)	(11.81, 60.44)	(169.50, 233.70)	(0.04, 0.19)	(0.08, 0.38)
HA	4.37 $^{(1.79)}$	1.96 $^{(0.83)}$	10.17 $^{(1.12)}$	0.12 $^{(0.05)}$	0.26 $^{(0.10)}$
	(1.10, 7.81)	(0.50, 3.58)	(8.01, 12.37)	(0.04, 0.21)	(0.08, 0.45)
CFA	3.80 $^{(1.59)}$	1.69 $^{(0.73)}$	12.78 $^{(1.12)}$	0.09 $^{(0.04)}$	0.21 $^{(0.08)}$
	(1.10, 7.01)	(0.43, 3.12)	(10.48, 14.91)	(0.03, 0.16)	(0.06, 0.36)
FA	34.71 $^{(12.52)}$	16.12 $^{(5.94)}$	85.41 $^{(7.62)}$	0.12 $^{(0.04)}$	0.25 $^{(0.09)}$
	(13.26, 60.16)	(5.52, 27.90)	(70.43, 99.84)	(0.04, 0.19)	(0.10, 0.42)

Abbreviations: σ_a^2, additive genetic variance; σ_c^2, common family environmental variance; σ_e^2, residual variance; c^2, common family environmental effect as a proportion of phenotypic variance; h^2, heritability; BW, body weight; TA, trunk area; HA, head area; CFA, caudal fin area; FA, fillet area.

Table 4. Posterior mean estimate, posterior standard deviation (superscript in parentheses), and 95% credibility intervals (below the estimate, in parentheses) for genetic (below the diagonal) and phenotypic (above the diagonal) correlations between Nile tilapia weight and body areas measured by digital image analysis.

Trait	BW	TA	HA	CFA	FA
BW	…	0.94 $^{(0.005)}$ (0.93, 0.95)	0.67 $^{(0.03)}$ (0.62, 0.72)	0.60 $^{(0.03)}$ (0.54, 0.66)	0.92 $^{(0.007)}$ (0.91, 0.93)
TA	0.98 $^{(0.01)}$ (0.96, 0.99)	…	0.69 $^{(0.03)}$ (0.63, 0.73)	0.67 $^{(0.03)}$ (0.61, 0.71)	0.98 $^{(0.002)}$ (0.97, 0.98)
HA	0.74 $^{(0.15)}$ (0.43, 0.97)	0.71 $^{(0.16)}$ (0.39, 0.96)	…	0.52 $^{(0.04)}$ (0.45, 0.59)	0.69 $^{(0.03)}$ (0.64, 0.74)
CFA	0.81 $^{(0.13)}$ (0.55, 0.99)	0.88 $^{(0.09)}$ (0.72, 0.99)	0.72 $^{(0.18)}$ (0.37, 0.98)	…	0.67 $^{(0.03)}$ (0.62, 0.72)
FA	0.97 $^{(0.02)}$ (0.94, 0.99)	0.98 $^{(0.01)}$ (0.96, 0.99)	0.70 $^{(0.16)}$ (0.38, 0.95)	0.85 $^{(0.10)}$ (0.66, 0.98)	…

Abbreviations: BW, body weight; TA, trunk area; HA, head area; CFA, caudal fin area; FA, fillet area.

CHAPTER 3:

Morphometric traits measured by digital image analysis, genetic parameter estimates, and application of a machine learning approach to predict growth, carcass, and fillet traits in Nile tilapia

Alex Júnio da Silva Cardoso[a,*], Carlos Antonio Lopes de Oliveira[b], Eric Costa Campos[b], Carolina Schlotefeldt[b], Karla Miky Tsujii[b], Humberto Todesco[b], Ricardo Pereira Ribeiro[b], Paulo Sávio Lopes[a], Fabyano Fonseca e Silva[a,†]

[a] Department of Animal Science, Universidade Federal de Viçosa, Campus Universitário, s/n, Viçosa, Minas Gerais, 36570-900, Brazil.

[b] Department of Animal Science, Universidade Estadual de Maringá, Av. Colombo, 5790, Jd. Universitário, Maringá, Paraná, 87020-900, Brazil.

[†]In memoriam.

* Corresponding author: E-mail address: alexjuniocardoso@hotmail.com

The Manuscript was written based on the guide for authors of *Aquaculture*.

3.1. Highlights

- A fast method was proposed to measure Nile tilapia morphometric traits using images
- The method allowed to measure 12 morphometric traits in 46 s
- Selection for body areas may lead to indirect genetic gains in fillet weight
- Image traits were used as predictor variables in supervised machine learning (ML)
- ML with Bayesian regularization showed better predictive ability for weight traits

3.2. Abstract

In this study, we aimed to (i) develop a fast and straightforward method to measure the length, height, and body areas of Nile tilapia by digital image analysis, (ii) estimate genetic parameters for these traits, and (iii) apply image traits in machine learning algorithms to predict body weight (BW), carcass weight (CW), fillet weight (FW), and fillet yield (FY). In total, 1,161 fish (427 days old at harvest, BW of 1,093 ± 346 g) were photographed. Body lengths (3 sections), heights (5 sections), trunk area (TA), head area (HA), fillet area (FA), and total area (TOT) were measured from the coordinate values (x and y values in the center of each pixel) of 20 pre-set landmarks on the surface of fish images using the free R software. The proposed method allowed to measure 12 traits in 46 s. Heritability estimates (h^2) for lengths and heights were moderate to high, ranging from 0.22 to 0.37. The h^2 values for TA, HA, FA, and TOT were 0.26, 0.35, 0.25, and 0.27, respectively. Positive and moderate to high genetic correlations were observed between morphometric traits and BW (0.66 to 0.98), FW (0.50 to 0.91), and CW (0.77 to 0.98). We highlight the genetic correlation of TA with BW (r_G = 0.98), FW (r_G = 0.91), and CW (r_G = 0.96). The TA/TOT ratio showed a positive and moderate genetic correlation (0.54) with FY. We investigated five supervised machine learning methods for predicting BW, CW, FW, and FY using image traits: multiple linear regression, feed-forward artificial neural network, deep learning, Bayesian regularization for feed-forward neural networks, and random forests. The Bayesian regularization model generally showed better performance and accuracy in predicting BW, CW, and FW. For FY prediction, all evaluated models had low performance and accuracy. The findings demonstrate that digital image analysis is a promising tool for measuring morphometric traits in Nile tilapia, given its non-invasive nature, fast operation, and low cost. Additionally, it was found that body areas can be used as selection criteria, particularly in future studies on body shape changes, and also, because they show positively correlated responses to FW and positive, albeit lower, correlations with FY. Finally, the Bayesian

regularization for the feed-forward neural network method showed the best performance in predicting BW, CW, and FW in Nile tilapia from image traits as predictor variables.

Keywords: computer vision, irregular polygons, heritability, genetic correlations, artificial neural networks

3.3. Introduction

Phenotyping is an important step for successful selective breeding. In Nile tilapia breeding programs, the growth rate has been the main selection objective, with body weight and morphometric traits (e.g., length, height, and width) as selection criteria (Ponzoni et al., 2005; Rutten et al., 2005; Nguyen et al., 2007; Thodesen et al., 2011; Reis Neto et al., 2014; Fernandes et al., 2015; Oliveira et al., 2016; Yoshida et al., 2021). However, the tilapia industry is interested in improving other economically important traits, such as body shape, fillet weight, and fillet yield (fillet weight/body weight ratio).

Measurement of fillet-related traits is unfeasible in live breeding candidates. Alternatively, it is possible to select fish based on sib information or through genetically correlated body traits. In general, genetic correlations between fillet weight and growth-related traits (e.g., body weight, length, height, and width) are high and positive, contrary to what is observed for fillet yield (Rutten et al., 2005; Nguyen et al., 2010; Turra et al., 2012; Gjerde et al., 2012; Garcia et al., 2017). Another important consideration is that body measurements influence the animal's overall body shape, which can impact consumer preference. For instance, longer and slimmer tilapia are less attractive to consumers, given the perception of the reduced amount of edible flesh (Nguyen et al., 2007).

In Nile tilapia selective breeding, phenotypic traits are measured manually. Such procedures are labor-intensive and, in some cases, inaccurate because of human error (Ventura et al., 2020). Computer vision systems, such as digital image analysis, combined with machine learning (e.g., artificial neural networks, ANN), have the potential to be used in precision aquaculture and genetic improvement programs (Fernandes et al., 2020a). The accuracy of machine vision is 10% higher than traditional methods, besides being more economical and faster (Hao et al., 2015). Digital image analyses are suitable to determine morphometric traits that are difficult to measure through traditional techniques, such as body areas, with the added benefit of reducing animal handling and eliminating the need for manual recording of data. Furthermore, images can provide explanatory variables for posterior prediction of growth, carcass, and fillet traits through ANN models. In a previous study, we observed that Nile tilapia body areas measured by image analysis have moderate heritability and are highly genetically correlated with body weight (Cardoso et al., 2021). However, genetic correlations of these body areas with fillet weight and yield are unknown.

Given these observations, we aimed to (i) develop a fast and straightforward method to measure the length, height, and body areas of Nile tilapia using digital image analysis, (ii) estimate genetic parameters for these traits, and (iii) apply image traits in machine learning algorithms to predict body weight, carcass weight, fillet weight, and fillet yield.

3.4. Material and methods

3.4.1. Ethics statement

All animal handling and slaughter procedures were approved by the Animal Use Ethics Committee of *Universidade Estadual de Maringá* (UEM, protocol no. 9452160720) and the Ethics Commission on the Use of Farm Animals of *Universidade Federal de Viçosa* (UFV, protocol no. 049/2020), in accordance by the norms of the National Council for the Control of Animal Experiments (CONCEA) and Brazilian legislation.

3.4.2. Population and growth test

The fish used in the study belonged to the 11th generation of the Nile tilapia breeding program (TILAMAX strain) of UEM. The test population was obtained by mating 48 sires and 74 dams from November to December 2019. Fish reproduction and larvae and fingerling cultivation were performed at the Experimental Aquaculture Station of UEM (23°31′08.3″S 52°02′20.5″W), Floriano, Paraná, Brazil, as described by Cardoso et al. (2021).

For the growth test, the fish were randomly stocked in four freshwater cages with an individual volume of 6 m^3 and a final stocking density of 75 kg/m^3, located in the Corvo River (22°39'24.9"S 52°46'51.5"W), Diamante do Norte, Paraná, Brazil. The growth test lasted about 300 days (from April 2020 to February 2021), during which period the fish were fed twice a day with commercial feed, as described by Oliveira et al. (2016).

3.4.3. Image acquisition

At the end of the growth test, a total of 1,161 fish were photographed with a high-resolution digital camera (Microsoft® Kinect v2.0, RGB color model, 1920 × 1080 pixels, 84.1° × 53.8° field of view) at a fixed height of 0.65 m from a blue surface used to support the animal and improve image contrast. The camera was connected to a computer (Intel® Core i5-7200 U, 2.5 GHz, 8 GB RAM), and images were captured using an in-house code implemented in the MATLAB software (2018). Before image acquisition, the fish were anesthetized with 100 mg of eugenol (Biodinâmica®, Paraná, Brazil) dissolved in 1 L of water (Simões et al., 2010) for safe handling of the animals.

All photographed fish were weighed and slaughtered to remove the fillet. The skin was manually removed and filleting was performed on both sides in sequence by a single trained professional. After filleting, the fillet weight was recorded and the fillet yield (FY) was calculated according to the equation FY (%) = (Fillet weight/Total weight) × 100.

3.4.4. Image analysis

We proposed a digital image analysis method to measure the lengths, heights, and body areas of Nile tilapia using R software (R Core Team, 2022). The imager package (Barthelmé and Tschumperlé, 2019) was used to open images in the software and the locator function (R Core Team, 2022) to obtain the coordinates (*x* and *y* in the center of each pixel) of 20 pre-set landmarks on the surface of the fish image (Figure 1).

Body heights and lengths (Figure 2) were obtained using the Euclidean distance (Wang et al., 2005) between two landmarks in a two-dimensional plane, as described in Eq. (1):

$$d_{\mathrm{AB}} = \sqrt{(x_{\mathrm{B}} - x_{\mathrm{A}})^2 + (y_{\mathrm{B}} - y_{\mathrm{A}})^2} \times \mathrm{CF} \qquad (1)$$

where d_{AB} is the distance between landmarks A and B; $(x_{\mathrm{A}}, y_{\mathrm{A}})$ and $(x_{\mathrm{B}}, y_{\mathrm{B}})$ are the coordinate values of landmarks A and B, respectively; and CF is the correction factor from pixels to centimeters, obtained by the formula $\mathrm{CF} = \frac{M}{d_{\mathrm{r}}}$, where M is the maximum measurement value of the ruler used in the image (in cm) and d_{r} is the Euclidean distance between landmarks 1 and 2 on the ruler (in pixels).

Head area (HA), trunk area (TA), fillet area (FA), and total area (TOT) (Figure 2) were calculated using the Gauss area formula (shoelace formula), as described by Lee and Lim (2017) (Eq. 2):

$$\mathrm{Area} = \frac{1}{2} \left|\sum_{i=1}^{n} (x_i \times y_{i+1}) - (x_{i+1} \times y_i)\right| \times \mathrm{CF}^2 \qquad (2)$$

for an n-sided polygon considering the union of the first source (x_1, y_1) to an n-landmark (x_n, y_n) counterclockwise with return to the source landmark for each body area. The ratios HA/TOT, TA/TOT, and FA/TOT were also calculated.

3.4.5. Estimation of genetic parameters

A preliminary analysis was performed to assess goodness of fit of the model for all studied traits. Two linear mixed models with different random effects were tested for each trait using the restricted maximum likelihood method (REML). Model 1 included only the random additive genetic effect, and model 2 included the random additive genetic effect and the common environmental random effect related to full-sib families. Fixed effects were the same for the two models. The fixed effects considered here were sex, cage, and age at harvest (linear and quadratic covariates). Models were compared using the Akaike information criterion (AIC) (Akaike, 1973), whereby the lower the AIC value, the higher the goodness of fit of the model to experimental data. For all the traits studied, the best-fitting model was that which did not include the family common environment random effect (Table S1, Supplementary Material). The significance of fixed effects was evaluated using the Wald test for the best-fitting model. Subsequent analyses included only significant fixed effects ($p < 0.05$) (Table S1, Supplementary Material). All analyses were performed using the nlme package (Pinheiro et al., 2021) of R software (R Core Team, 2022).

Estimates of variance components and heritability were obtained using a single-trait animal model, described in matrix form as shown in Eq. (3):

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Za} + \mathbf{e} \qquad (3)$$

where $\mathbf{y}$ is the vector of phenotypic observations; $\boldsymbol{\beta}$ is the vector of fixed effects (sex - except for FY, cage, and the linear covariate age); $\mathbf{a}$ is the random vector of additive genetic effects, distributed as $\mathbf{a} \sim N(0, \mathbf{A}\sigma_\mathrm{a}^2)$, being $\mathbf{A}$ the additive relationship matrix based on fish genealogical information and σ_a^2 the additive genetic variance; $\mathbf{e}$ is the random vector of residuals, distributed as $\mathbf{e} \sim N(0, \mathbf{I}\sigma_\mathrm{e}^2)$, being $\mathbf{I}$ an identity matrix and σ_e^2 the residual variance; and $\mathbf{X}$ and $\mathbf{Z}$ are incidence matrices related to $\boldsymbol{\beta}$ and $\mathbf{a}$, respectively. Heritability (h^2) was calculated as $h^2 = {\sigma_\mathrm{a}^2}/{(\sigma_\mathrm{a}^2 + \sigma_\mathrm{e}^2)}$.

Genetic correlation estimates were obtained using a two-trait animal model, with the same matrix form described previously but with $\mathbf{a} \sim N(0, \mathbf{A} \otimes \mathbf{G})$ and $\mathbf{e} \sim N(0, \mathbf{I} \otimes \mathbf{R})$, where $\mathbf{G}$ and $\mathbf{R}$ are the additive genetic and residual (co)variance matrices, respectively. $\mathbf{A}$ and $\mathbf{I}$ were previously defined, and $\otimes$ is the Kronecker product (direct product) operator. Genetic correlation (r_{ij}) was calculated as $r_{ij} = \sigma_{aij} \Big/ \sqrt{\sigma_{ai}^2 \times \sigma_{aj}^2}$, where σ_{aij} is the additive genetic covariance between traits *i* and *j*.

The AIREMLF90 software (Misztal et al., 2014) was used in these analyses. The genealogical dataset consisted of 25,388 animals born between 2008 and 2019, comprising 11 generations.

3.4.6. Machine learning approach

Supervised machine learning techniques were used to predict body weight (BW), carcass weight (CW), fillet weight (FW), and FY, with sex and traits measured by digital image analysis (lengths, heights, and body areas) used as predictor variables.

Five machine learning models were constructed using the free R software. (1) In the multiple linear regression model, traits extracted from images were used as independent variables (inputs) and traits to be predicted were used as dependent variables (responses). (2) In the ANN model, a feed-forward neural network composed of an input layer (13 nodes), a single hidden layer (two neurons), and an output layer was developed, with skip-layer connections and switch for linear output units. The ANN model was implemented using the nnet package (Venables and Ripley, 2002). (3) A deep learning (DL) model was developed as a feed-forward neural network composed of two hidden layers (200 neurons each), with rectified linear unit activation function, automatic loss function, and 0.005 learning rate. The DL model was implemented using the h2o package (LeDell et al., 2022). (4) A Bayesian regularization for the feed-forward neural network (BRNN) model was developed using a feed-forward neural network with Nguyen and Widrow's (1990) algorithm to assign initial weights and the Gauss-Newton algorithm (Foresee and Hagan, 1997) to perform the optimization. The

BRNN model was implemented using the brnn package (Rodriguez and Gianola, 2020). (5) A random forest (RF) model was created using an ensemble learning algorithm, where several decision trees obtained from training data samples were trained to predict the desired traits. The RF model was implemented using the h2o package (LeDell et al., 2022).

To verify the effectiveness of prediction methods, we used a 10-fold cross-validation procedure with 5 replicates, totaling 50 runs per model and predicted trait. Cross-validation folds were randomly split to provide the training (n = 1045) and validation (n = 116) datasets. Pearson's correlation coefficient (r), mean absolute error (MAE), and root mean square error (RMSE) between predicted and observed values were calculated. The Kruskal–Wallis test followed by Dunn's multiple comparison test with Bonferroni correction was performed to compare r, MAE, and RMSE values among models. The level of significance was set at $p < 0.05$.

3.5. Results

3.5.1. Descriptive statistics

In total, 624 males and 537 females (1,161 individuals), 427 days old at harvest, were photographed, and 12 traits were measured by image analysis (Table 1). The mean FY was 33.24% and showed a low coefficient of variation (7.14%) (Table 1). The coefficients of variation for body areas (21.41–22.14%) were higher than those for other traits (9.95–12.48%) measured by image analysis (Table 1). The image analysis method proposed in the present study allowed to measure body lengths, heights, and areas simultaneously in 46.46 ± 3.94 s per image.

3.5.2. Heritability and genetic correlations

Heritability estimates for body lengths, heights, and areas measured by digital image analysis were moderate, ranging from 0.22 to 0.37 (Table 2). We observed first-time heritability estimates for FA/TOT, TA/TOT, and HA/TOT ratios of 0.36, 0.31, and 0.31, respectively (Table 2). Moderate heritability estimates were also found for body weight at harvest (BW; 0.26), carcass weight (CW; 0.31), fillet weight (FW; 0.25), and fillet yield (FY; 0.28) (Table 2).

Positive and moderate to high genetic correlations were observed among morphometric traits and body weight at harvest (0.66 to 0.98), fillet weight (0.50 to 0.91), and carcass weight (0.77 to 0.98) (Figure 3). Among morphometric traits, the genetic correlation ranged from 0.53 to 0.99 (Figure 3). Genetic correlations among fillet yield and other traits were low or null (Figure 3). The TA/TOT ratio showed a positive and moderate genetic correlation (0.54) with FY (Figure 3). The opposite was observed for HA/TOT, which showed a negative and moderate genetic correlation (-0.54) with FY (Figure 3).

3.5.3. Machine learning for prediction of growth, carcass, and fillet traits

For BW prediction from traits measured by image analysis, the BRNN model resulted in a higher correlation between observed and predicted values (r = 0.987) and lower prediction errors (MAE = 39.545, RMSE = 54.700) than the other models (Table 3). For CW and FW, the BRNN model showed higher absolute values for correlation and lower values for MAE and RMSE; however, it did not differ statistically from linear regression or the ANN model (Table 3). In general, the BRNN model showed better performance and accuracy in predicting BW, CW, and FW. Regarding FY prediction, all evaluated models resulted in low correlation values, even though MAE and RMSE were also low (Table 3).

3.6. Discussion

3.6.1. Phenotyping

In this study, we proposed an image analysis method based on the R software that allowed to measure 12 morphometric traits of Nile tilapia (lengths, heights, and body areas) in 46 s. Although the technique uses manual steps for image acquisition and definition of reference points, traits are measured through a coordinate system (*x*- and *y*-axes), which allows extrapolation to other fish species. Furthermore, the method can be applied to images obtained by conventional sensors, with little interference from image background, fish color, or ambient lighting.

In fisheries and aquaculture, computer vision technologies have been used to measure size, estimate body mass, monitor animal welfare and behavior, detect sex, count individuals, and identify species (Zion, 2012). However, there are few reports of applications in fish selective breeding. Genetic improvement programs require accurate measurement of traits in a large number of animals and in a short time. Currently, phenotypic datasets used in Nile tilapia breeding programs are manually recorded. The procedures are labor-intensive and, in some cases, can provide inaccurate results biased by human error (Ventura et al., 2020; Fu and Yuna, 2022). Digital image analysis can provide objective, fast, non-invasive, and accurate measurements of morphometric traits in several fish species (Hao et al., 2015; Navarro et al., 2016; Fernandes et al., 2020b), with a positive impact on animal welfare. An additional advantage is that standardized measurement methods such as image analysis can decrease environmental variance, improving the accuracy of genetic parameter estimates (Navarro et al., 2016).

Automated phenotyping techniques using image analysis have been developed for some fish species, including Nile tilapia. Navarro et al. (2016) developed an image analysis software (IMAFISH_ML) to measure 27 morphometric traits (lengths, heights, widths, and areas) in gilthead seabream (*Sparus aurata*), meager (*Argyrosomus regius*), and red porgy (*Pagrus pagrus*) using two images of each fish at different angles (lateral and dorsal). Fernandes et al. (2020b) used DL methods for image segmentation (image background, fish body, and fins) and autonomous measurement of the body area, length, height, and eccentricity in Nile tilapia. There are however some limitations. These techniques require that fish be removed from the culture environment for image acquisition, the algorithms are applied only to the studied species, and operators must have advanced knowledge in programming languages and/or specialized software for image analysis.

3.6.2. Genetic parameters

Heritability estimates for lengths and heights measured by digital image analysis were moderate to high (0.22–0.37) and are similar to those found in studies using manual methods to measure the same traits in tilapia (Rutten et al., 2005; Nguyen et al., 2007; Nguyen et al., 2010; Reis Neto et al., 2014; Fernandes et al., 2015). By contrast, Elalfy et al. (2021) observed a 24% increase in the heritability of body length in gilthead seabream (*S. aurata*) measured using IMAFISH software (developed by Navarro et al., 2016). Therefore, non-invasive technologies for measuring traits can contribute to increasing the accuracy of genetic parameter estimates.

Heritability estimates for body areas (0.26–0.35) obtained by image analysis were consistent with those observed in a previous study with tilapia grown in pond systems (Cardoso et al., 2021). We found first-time heritability estimates of 0.36, 0.31, and 0.31 for FA/TOT, TA/TOT, and HA/TOT ratios, respectively, which indicates the potential of these new traits as selection criteria for targeted changes in Nile tilapia body shape.

The mean FY found in the current study was 33.24%, within the range of 33% to 45% reported for Nile tilapia in previous studies (Rutten et al., 2005; Nguyen et al., 2010; Gjerde et al., 2012; Garcia et al., 2017). Heritability estimates for FY and fillet weight (FW) were 0.28 and 0.25, respectively, suggesting that selection to improve these traits can be effective. However, selection for fillet traits is not an easy task because they are expensive to measure and cannot be recorded by traditional methods in live breeding candidates (Rutten et al., 2005; Nguyen et al., 2010; Garcia et al., 2017). For this reason, the selection is usually performed based on sib information or correlated body measurements. In general, genetic correlations between FW and BW are positive and high, whereas those between FY and BW are weakly positive or negative (Nguyen et al., 2010; Garcia et al., 2017; Gjerde et al., 2012).

Genetic correlations between morphometric traits and FW in the current study were positive and high, as reported in other studies where these traits were measured manually in Nile tilapia (Rutten et al., 2005; Nguyen et al., 2010). To the best of our knowledge, our study is the first to report high genetic correlations between body areas measured by image analysis and FW (0.62 to 0.91). The high genetic correlation for the trunk area (0.91) is particularly noteworthy. The findings indicate those body areas measured by image analysis constitute good selection criteria to indirectly improve FW in Nile tilapia, precluding the need to slaughter the animals for analysis, thereby increasing selection intensity.

Improving FY through selection for correlated responses is challenging, given the low correlations between FY and other body measurements. Contrary to what we observed in our study, Rutten et al. (2005) estimated high to moderate genetic correlations between FY and standard length (0.62), head length (0.47), and width (0.98) in Nile tilapia. However, it must be noted that, in the referred study, tilapia had a mean weight of 787.7 g and were grown in a closed recirculation system, fillets were cut using a filleting machine, and the skin was not removed from fillets, resulting in a coefficient of variation of 15.5% for FY. Therefore, the discrepancy between estimates of genetic correlations between the present study and that of

Rutten et al. (2005) is likely due to differences in cultivation conditions, fish age, filleting method, and statistical approaches.

When we simulated a ratio trait as in FY, the TA/TOT ratio was positively correlated with FY (0.54), whereas HA/TOT showed the opposite behavior (−0.54), indicating a restricted correlated response that was, nevertheless, superior to those observed for the other morphometric traits. Determination of body volume by 3D imaging sensors, such as depth sensors based on a structured infrared-light system (Kongsro, 2014; Fernandes et al., 2020a), might allow establishing ratios more related to FY, possibly revealing stronger genetic correlations.

3.6.3. Machine learning approach

This research investigated the possibility of using morphometric traits measured by digital image analysis to predict BW, CW, FW, and FY in Nile tilapia through algorithms based on non-linear behavior, such as machine learning. All 12 measured traits (body lengths, heights, and areas) were applied as input variables in prediction models. Sex was tested as a categorical variable because, in Nile tilapia, growth patterns differ between sexes (sexual size dimorphism), with males being larger than females. When sex was included in the models, there was an improvement in predicted results.

In general, the BRNN model had better performance and accuracy in the prediction of BW, CW, and FW. Linear regression models are the most commonly used in aquaculture. Rutten et al. (2004) found a high correlation (0.98) between observed and predicted values of FW in Nile tilapia when BW, length, height, width, and corrected length were considered in a linear regression model. By contrast, the correlation between observed and predicted FY values was low (0.38). In the present study, by applying the BRNN model, we also observed a high correlation (0.97) between observed and predicted values for FW, with MAE and RSME values of 23.25 and 33.42, respectively. Regarding FY prediction, although MAE and RSME values were low, correlations were also low for all evaluated models. Fernandes et al. (2020b) found that the best predictive model for BW and CW in Nile tilapia included body area and its square as covariates in linear regression, resulting in MAE and RSME values of 57.77 and 77.52 for BW and 36.07 and 49.98 for CW, respectively. In the present study, MAE and RSME values for the BRNN model were lower than those reported by Fernandes et al. (2020b), indicating lower prediction errors for BW and CW.

Artificial neural networks are structured by connections between input, middle ("hidden"), and output layers as well as by the type of processing performed at each neuron (activation function), providing non-linear relationships between predictors and responses (Gianola et al., 2011). However, neural networks are prone to overfitting the training dataset, preventing good predictions in dataset validation, when new data are presented to the network (Morota et al., 2018; Sinecen, 2019). In general, models of high complexity tend to suffer from overfitting unless big datasets are available (Morota et al., 2018). The BRNN model prevents overfitting by restricting the size of the network connection (e.g., via shrinkage) through a process known as regularization, which uses Bayesian methods (Burden and Winkler, 2008; Gianola et al., 2011). Therefore, the BRNN method may have better fit quality and predictive capacity than other methods, as observed in the current study for BW, CW, and FW prediction in Nile tilapia.

Machine learning methods can be used in tilapia breeding programs, mainly for determining traits that are difficult to measure (e.g., fillet weight), which increases the selection intensity for these traits. Given the promising results observed for FW prediction, FY can be estimated indirectly through the predicted FW with the BW (obtained by weighing scale or predicted). For that, evaluations in a larger image dataset are necessary to increase the accuracy and decrease the prediction error. However, the challenges of automating the phenotyping process in aquaculture lie in capturing images in the culture environment and storing and analyzing large image datasets. Surpassing these limitations will require collaboration among transdisciplinary fields with complementary backgrounds, such as computer science, animal science, engineering, mathematics, statistics, and the fish industry.

3.7. Conclusions

Digital image analysis is a relevant tool for measuring morphometric traits in Nile tilapia. Additionally, body areas can be used as selection criteria, particularly in future studies of body shape changes, and also, because they show favorable correlated responses with FW and favorable, albeit lower, correlations with FY. Finally, the BRNN method best predicts BW in Nile tilapia using image traits as predictor variables. For CW and FW prediction, the BRNN, linear regression, and ANN models are indicated.

3.8. Acknowledgments

We thank the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) for granting a scholarship to the first author (process no. 148102/2018-2). We are grateful to the PeixeGen Research Group, which specializes in management, breeding, and molecular genetics in freshwater fish farming, and all members of the Aquaculture Station of the State University of Maringá for their assistance in this study.

3.9. References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle, in: Parzen, E., Tanabe, K., Kitagawa, G. (Eds.), Selected Papers of Hirotugu Akaike. Springer, New York, pp. 199-213.

Barthelmé, S., Tschumperlé, D., 2019. imager: an R package for image processing based on CImg. J. Open Source Softw. 4, 1012. https://doi.org/10.21105/joss.01012

Burden F., Winkler D., 2008. Bayesian Regularization of Neural Networks, in: Livingstone, D.J. (Eds.), Artificial Neural Networks. Humana Press, New Jersey, pp.23-42. https://doi.org/10.1007/978-1-60327-101-1_3

Cardoso, A.J.S., Oliveira, C.A.L., Campos, E.C., Ribeiro, R.P., Assis, G.J.F., Silva, F.F., 2021. Estimation of genetic parameters for body areas in Nile tilapia measured by digital image analysis. J. Anim. Breed. Genet. 138, 731–738. https://doi.org/10.1111/jbg.12551

Elalfy, I., Shin, H.S., Negrín-Báez, D., Navarro, A., Zamorano, M.J., Manchado, M., Afonso, J.M., 2021. Genetic parameters for quality traits by non-invasive methods and their G x E interactions in ocean cages and estuaries on gilthead seabream (*Sparus aurata*). Aquaculture, 537, 736462. https://doi.org/10.1016/j.aquaculture.2021.736462

Fernandes, A.F.A., Silva, M.A., de Alvarenga, E.R., Teixeira, E.A., Junior, A.F.S., Alves, G.F.O., Salles, S.C.M., Manduca, L.G., Turra, E.M., 2015. Morphometric traits as selection criteria for carcass yield and body weight in Nile tilapia (*Oreochromis niloticus* L.) at five ages. Aquaculture, 446, 303-309. https://doi.org/10.1016/j.aquaculture.2015.05.009.

Fernandes, A.F.A, Dórea, J.R.R., Rosa, G.J.M., 2020a. Image analysis and computer vision applications in animal sciences: an overview. Front. Vet. Sci. 7, 1-10. https://doi.org/10.3389/fvets.2020.551269

Fernandes, A.F.A., Turra, E.M., de Alvarenga, É.R., Passafaro, T.L., Lopes, F.B., Alves, G.F.O., Singh, V., Rosa, G.J.M., 2020b. Deep Learning image segmentation for extraction of fish body measurements and prediction of body weight and carcass traits in Nile tilapia. Comput. Electron. Agric. 170, 105274. https://doi.org/10.1016/j.compag.2020.105274

Foresee, F. D., Hagan, M. T., 1997. Gauss-Newton approximation to Bayesian learning. In Proceedings of international conference on neural networks (ICNN'97), 3, 1930-1935. https://doi.org/10.1109/ICNN.1997.614194

Fu, G., Yuna, Y., 2022. Phenotyping and phenomics in aquaculture breeding. Aquac. Fish. 7, 140-146. https://doi.org/10.1016/j.aaf.2021.07.001

Garcia, A.L.S., Oliveira, C.A.L., Karim, H.M., Sary, C., Todesco, H., Ribeiro, R.P., 2017. Genetic parameters for growth performance, fillet traits, and fat percentage of male Nile tilapia (*Oreochromis niloticus*). J. Appl. Genet. 58, 527-533. https://doi.org/10.1007/s13353-017-0413-6

Gianola, D., Okut, H., Weigel, K.A., Rosa, G.J., 2011. Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. BMC Genet. 12, 1-14. https://doi.org/10.1186/1471-2156-12-87

Gjerde, B., Mengistu, S.B., Ødegård, J., Johansen, H., Altamirano, D.S., 2012. Quantitative genetics of body weight, fillet weight and fillet yield in Nile tilapia (*Oreochromis niloticus*). Aquaculture, 342, 117-124. https://doi.org/10.1016/j.aquaculture.2012.02.015

Hao, M., Yu, H., Li, D., 2015. The measurement of fish size by machine vision-a review. In International Conference on Computer and Computing Technologies in Agriculture. Springer, Cham, pp. 15-32.

Kongsro, J., 2014. Estimation of pig weight using a Microsoft Kinect prototype imaging system. Comput. Electron. Agric. 109, 32-35. https://doi.org/10.1016/j.compag.2014.08.008

LeDell, E., Gill, N., Aiello, S., Fu, A., Candel, A., Click, C., Kraljevic, T., Nykodym, T., Aboyoun, P., Kurka, M., Malohlava, M., 2022. h2o: R Interface for the 'H2O' Scalable Machine Learning Platform. R package version 3.36.0.2. https://CRAN.R-project.org/package=h2o (accessed 24 January 2022).

Lee, Y., Lim, W., 2017. Shoelace Formula: Connecting the Area of a Polygon and the Vector Cross Product. Math. Teach. 110, 631–636. https://doi.org/10.5951/mathteacher.110.8.0631

Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., Vitezica, Z., 2014. Manual for BLUPF90 family of programs. University of Georgia, Athens, GA, USA.

Morota, G., Ventura, R.V., Silva, F.F., Koyama, M., Fernando, S.C., 2018. Big data analytics and precision animal agriculture symposium: Machine learning and data mining advance predictive big data analysis in precision animal agriculture. J. Anim. Sci. 96, 1540-1550. https://doi.org/10.1093/jas/sky014

Navarro, A., Lee-Montero, I., Santana, D., Henríquez, P., Ferrer, M. A., Morales, A., Soula, M., Badilla, R., Negrín-Báez, D., Zamorano, M.J., Afonso, J. M., 2016. IMAFISH_ML: A fully-automated image analysis software for assessing fish morphometric traits on gilthead seabream (*Sparus aurata* L.), meagre (*Argyrosomus regius*) and red porgy (*Pagrus pagrus*). Comput. Electron. Agric. 121, 66-73. https://doi.org/10.1016/j.compag.2015.11.015

Nguyen, D., Widrow, B., 1990. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. Proc. Int. Jt. Conf. Neural Netw. 3, 21-26.

https://doi.org/10.1109/IJCNN.1990.137819

Nguyen, N.H., Khaw, H.L., Ponzoni, R.W., Hamzah, A., Kamaruzzaman, N., 2007. Can sexual dimorphism and body shape be altered in Nile tilapia (*Oreochromis niloticus*) by genetic means?. Aquaculture, 272, S38-S46. https://doi.org/10.1016/j.aquaculture.2007.08.013.

Nguyen, N.H., Ponzoni, R.W., Abu-Bakar, K.R., Hamzah, A., Khaw, H.L., Yee, H. Y., 2010. Correlated response in fillet weight and yield to selection for increased harvest weight in genetically improved farmed tilapia (GIFT strain), *Oreochromis niloticus.* Aquaculture, 305, 1-5. https://doi.org/10.1016/j.aquaculture.2010.04.007

Oliveira, C.A.L., Ribeiro, R.P., Yoshida, G.M., Kunita, N.M., Rizzato, G.S., de Oliveira, S.N., dos Santos, A.I., Nguyen, N.H., 2016. Correlated changes in body shape after five generations of selection to improve growth rate in a breeding program for Nile tilapia *Oreochromis niloticus* in Brazil. J. Appl. Genet. 57, 487–493. https://doi.org/10.1007/s13353-016-0338-5

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team, 2021. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-152. https://CRAN.R-project.org/package=nlme (accessed 24 January 2022).

Ponzoni, R.W., Hamzah, A., Tan, S., Kamaruzzaman, N., 2005. Genetic parameters and response to selection for live weight in the GIFT strain of Nile tilapia (*Oreochromis niloticus*). Aquaculture, 247(1-4), 203-210. https://doi.org/10.1016/j.aquaculture.2005.02.020

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reis Neto, R.V., Oliveira, C.A.L., Ribeiro, R.P., Freitas, R.T.F., Allaman, I.B., Oliveira, S.N., 2014. Genetic parameters and trends of morphometric traits of GIFT tilapia under selection for weight gain. Sci. Agric. 71, 259-265. https://doi.org/10.1590/0103-9016-2013-0294.

Rodriguez, P. P., Gianola, D., 2020. brnn: Bayesian Regularization for Feed-Forward Neural Networks. R package version 0.8. https://CRAN.R-project.org/package=brnn (accessed 24 January 2022).

Rutten, M. J., Bovenhuis, H., Komen, H., 2005. Genetic parameters for fillet traits and body measurements in Nile tilapia (*Oreochromis niloticus* L.). Aquaculture, 246, 125-132. https://doi.org/10.1016/j.aquaculture.2005.01.006

Rutten, M.J., Bovenhuis, H., Komen, H., 2004. Modeling fillet traits based on body measurements in three Nile tilapia strains (*Oreochromis niloticus* L.). Aquaculture, 231, 113-122. https://doi.org/10.1016/j.aquaculture.2003.11.002

Simões, L.N., Paiva, G., de Carvalho Gomes, L., 2010. Óleo de cravo como anestésico em adultos de tilápia-do-Nilo. Pesqui. Agropecu. Bras. 45, 1472–1477. https://doi.org/10.1590/S0100-204X2010001200019

Sinecen, M., 2019. Comparison of genomic best linear unbiased prediction and Bayesian regularization neural networks for genomic selection. IEEE Access, 7, 79199-79210. https://doi.org/ 10.1109/ACCESS.2019.2922006

Thodesen, J., Rye, M., Wang, Y.X., Yang, K.S., Bentsen, H.B., Gjedrem, T., 2011. Genetic improvement of tilapias in China: genetic parameters and selection responses in growth of Nile tilapia (*Oreochromis niloticus*) after six generations of multi-trait selection for growth and fillet yield. Aquaculture, 322, 51-64. https://doi.org/10.1016/j.aquaculture.2011.10.010

Turra, E.M., Oliveira, D.A.A., Valente, B.D., Teixeira, E.A., Prado, S.A., Alvarenga, E. R., Melo, D.C., Felipe, V.P.S., Fernandes, A.F.A., Silva, M.A., 2012. Longitudinal genetic analyses of fillet traits in Nile tilapia *Oreochromis niloticus*. Aquaculture, 356, 381-390. https://doi.org/10.1016/j.aquaculture.2012.04.039

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, Fourth edition. Springer, New York.

Ventura, R.V., Silva, F.F., Yáñez, J.M., Brito, L.F., 2020. Opportunities and challenges of phenomics applied to livestock and aquaculture breeding in South America. Anim. Front. 10, 45-52. https://doi.org/10.1093/af/vfaa008

Wang, L., Zhang, Y., Feng, J., 2005. On the Euclidean distance of images. IEEE Trans. Pattern Anal. Mach. Intell. 27, 1334–1339. https://doi.org/10.1109/TPAMI.2005.165

Yoshida, G.M., de Oliveira, C.A.L., Campos, E.C., Todesco, H., Araújo, F.C., Karin, H.M., Zardin, A.M., Bezerra Júnior, J.S., Filho, L.A., Vargas, L., Ribeiro, R.P., 2021. A breeding program for Nile tilapia in Brazil: Results from nine generations of selection to increase the growth rate in cages. J. Anim. Breed. Genet. Early View. https://doi.org/10.1111/jbg.12650

Zion, B., 2012. The use of computer vision technologies in aquaculture – a review. Comput. Electron. Agric. 88, 125-132. https://doi.org/10.1016/j.compag.2012.07.010

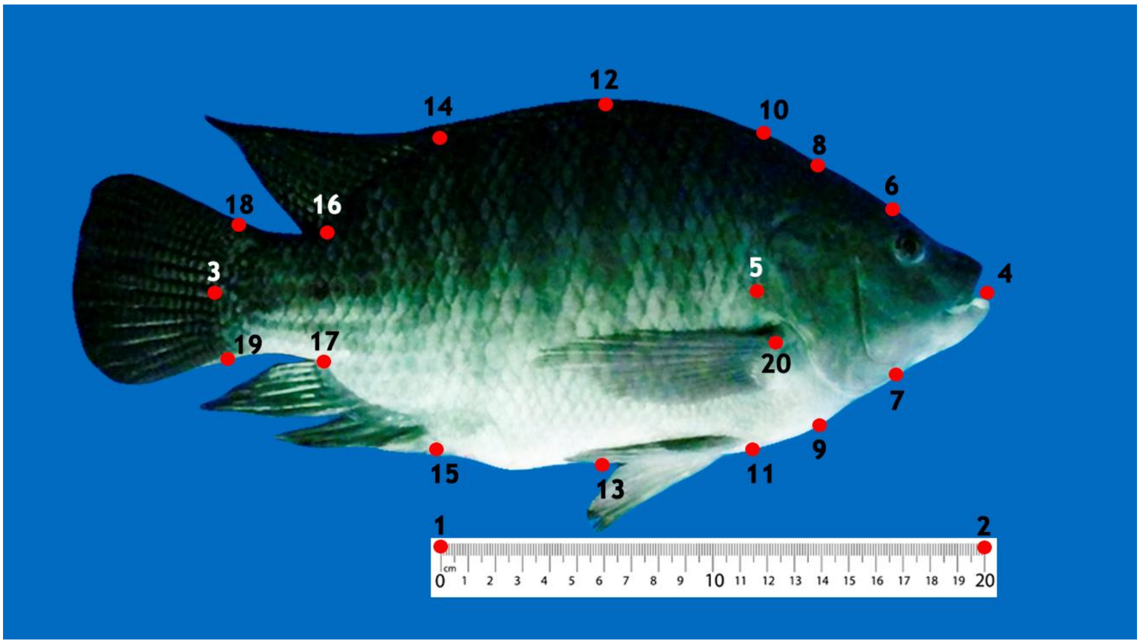

Figure 1. Landmarks placed on each Nile tilapia image for the measurement of morphometric traits. (1) Initial position of the ruler (0 cm), (2) final position of the ruler (20 cm), (3) intersection between trunk and caudal fin (center), (4) head extremity, (5) larger extremity of the operculum, (6) upper dorsal lateral region of the eye, (7) perpendicular to landmark 6, (8) tangent to landmark 5, (9) perpendicular to landmark 8, (10) beginning of dorsal fin, (11) perpendicular to landmark 10, (12) higher trunk height, (13) perpendicular to landmark 12, (14 and 15) landmarks perpendicular to the beginning of the anal fin, (16) dorsal beginning of the caudal peduncle, (17) ventral beginning of caudal peduncle and end of the anal fin, (18) dorsal intersection of the trunk and caudal fin, (19) ventral intersection of the trunk and caudal fin, and (20) beginning of pectoral fin.

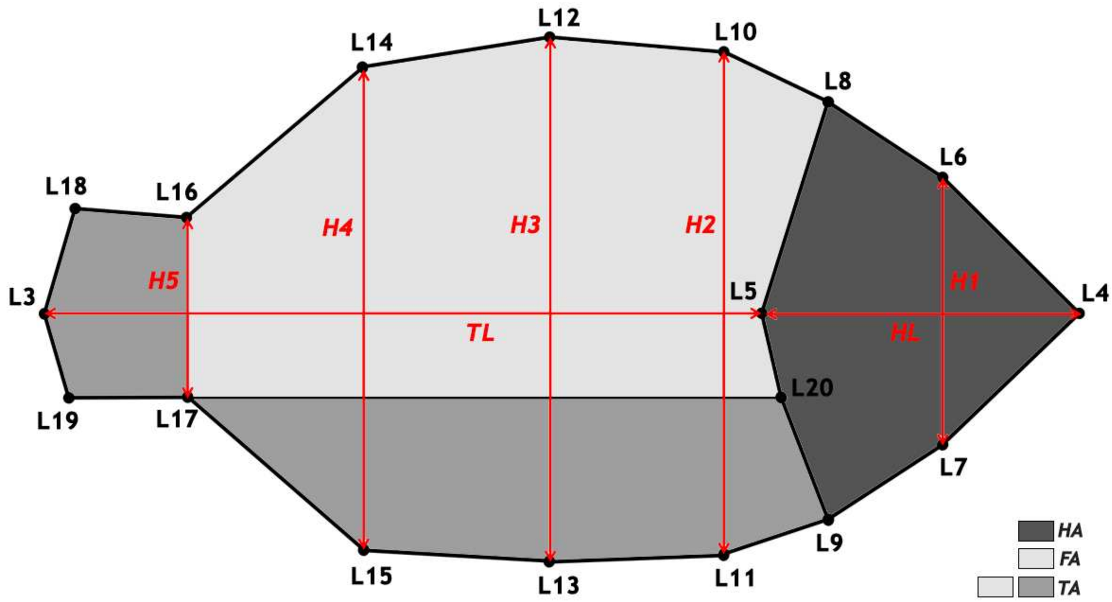

Figure 2. Representation of trunk length ($TL = d_{L3,L5}$), head length ($HL = d_{L5,L4}$), standard length ($SL = TL + HL$), height 1 ($H1 = d_{L6,L7}$), height 2 ($H2 = d_{L10,L11}$), height 3 ($H3 = d_{L12,L13}$), height 4 ($H4 = d_{L14,L15}$), height 5 ($H5 = d_{L16,L17}$), head area ($HA = Area_{L4,L6,L8,L5,L20,L9,L7,L4}$), trunk area ($TA = Area_{L8,L10,L12,L14,\ldots,L11,L9,L20,L5,L8}$), fillet area ($FA = Area_{L8,L10,L12,L14,L16,L17,L20,L5,L8}$), and total area ($TOT = TA + HA$), as measured by digital image analysis.

Table 1. Descriptive statistics for the studied traits.

Trait*	Mean	Min	Max	SD	CV (%)
Age (days)	427.57	412.00	445.00	10.18	2.38
BW (g)	1,093.83	320.00	2,290.00	346.97	31.72
TL (cm)	21.85	6.06	27.80	2.32	10.63
HL (cm)	9.42	6.37	16.68	1.04	11.09
SL (cm)	31.26	20.48	38.88	3.11	9.95
H1 (cm)	6.49	3.29	9.44	0.75	11.52
H2 (cm)	12.04	6.08	16.54	1.44	11.97
H3 (cm)	12.89	6.95	17.34	1.59	12.34
H4 (cm)	11.14	6.38	15.20	1.39	12.48
H5 (cm)	4.67	3.09	6.49	0.55	11.86
TA (cm^2)	229.49	29.67	379.35	50.80	22.14
HA (cm^2)	51.50	6.81	97.42	11.03	21.41
FA (cm^2)	153.54	19.34	254.23	33.47	21.80
TOT (cm^2)	280.99	36.48	461.68	60.77	21.63
FW (g)	364.02	103.86	792.78	119.90	32.94
CW (g)	638.91	188.14	1,351.94	203.48	31.85
FY (%)	33.24	15.60	49.24	2.37	7.14
FA/TOT (%)	54.63	49.04	59.35	1.51	2.77
TA/TOT (%)	81.59	76.54	85.46	1.54	1.89
HA/TOT (%)	18.41	14.54	23.46	1.54	8.38

*Total sample of 1,161 individuals (624 males and 537 females).

Min, minimum; Max, maximum; SD, standard deviation; CV, coefficient of variation; BW, body weight; TL, trunk length; HL, head length; SL, standard length; H1, H2, H3, H4, and H5 represent heights 1, 2, 3, 4, and 5, respectively; TA, trunk area; HA, head area; FA, fillet area; TOT, total area; FW, fillet weight; CW, carcass weight; FY, fillet yield.

Table 2. Means and standard errors (superscript in parentheses) of variance components and genetic parameters for the studied traits.

Trait	σ_a^2	σ_e^2	h^2
BW	$14{,}810.00^{(3{,}826.3)}$	$41{,}817.00^{(2{,}961.70)}$	$0.26^{(0.06)}$
TL	$0.67^{(0.19)}$	$2.41^{(0.16)}$	$0.22^{(0.06)}$
HL	$0.24^{(0.06)}$	$0.41^{(0.04)}$	$0.37^{(0.07)}$
SL	$1.35^{(0.35)}$	$3.75^{(0.27)}$	$0.26^{(0.06)}$
H1	$0.09^{(0.02)}$	$0.24^{(0.02)}$	$0.28^{(0.06)}$
H2	$0.27^{(0.07)}$	$0.80^{(0.06)}$	$0.25^{(0.06)}$
H3	$0.40^{(0.10)}$	$0.92^{(0.07)}$	$0.30^{(0.07)}$
H4	$0.29^{(0.07)}$	$0.70^{(0.05)}$	$0.30^{(0.07)}$
H5	$0.05^{(0.01)}$	$0.12^{(0.01)}$	$0.29^{(0.06)}$
TA	$326.30^{(85.23)}$	$940.73^{(66.31)}$	$0.26^{(0.06)}$
HA	$21.97^{(5.15)}$	$40.64^{(3.51)}$	$0.35^{(0.07)}$
FA	$140.79^{(37.21)}$	$422.02^{(29.32)}$	$0.25^{(0.06)}$
TOT	$487.31^{(124.78)}$	$1303.50^{(94.72)}$	$0.27^{(0.06)}$
FW	$1{,}850.00^{(485.80)}$	$5{,}544.80^{(383.68)}$	$0.25^{(0.06)}$
CW	$6{,}070.60^{(1{,}486.30)}$	$13{,}571.00^{(1{,}067.70)}$	$0.31^{(0.07)}$
FY	$1.51^{(0.40)}$	$3.92^{(0.30)}$	$0.28^{(0.07)}$
FA/TOT	$0.72^{(0.17)}$	$1.30^{(0.11)}$	$0.36^{(0.07)}$
TA/TOT	$0.56^{(0.13)}$	$1.26^{(0.10)}$	$0.31^{(0.07)}$
HA/TOT	$0.56^{(0.13)}$	$1.26^{(0.10)}$	$0.31^{(0.06)}$

σ_a^2, additive genetic variance; σ_e^2, residual variance; h^2, heritability; BW, body weight; TL, trunk length; HL, head length; SL, standard length; H1, H2, H3, H4, and H5 represent heights 1, 2, 3, 4, and 5, respectively; TA, trunk area; HA, head area; FA, fillet area; TOT, total area; FW, fillet weight; CW, carcass weight; FY, fillet yield.

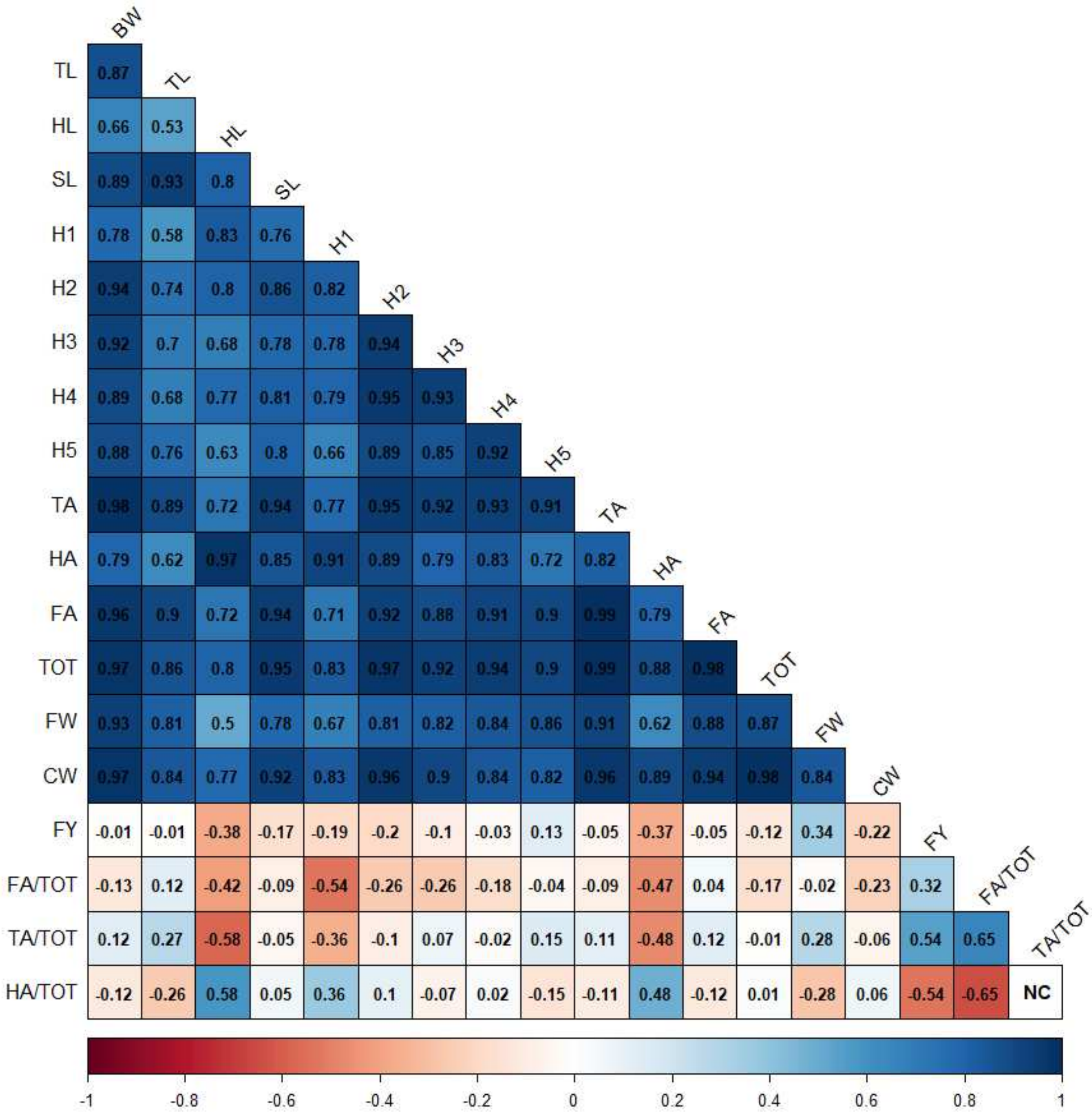

Figure 3. Genetic correlations among the studied traits. BW, body weight; TL, trunk length; HL, head length; SL, standard length; H1, H2, H3, H4, and H5 represent heights 1, 2, 3, 4, and 5, respectively; TA, trunk area; HA, head area; FA, fillet area; TOT, total area; FW, fillet weight; CW, carcass weight; FY, fillet yield; NC, not converged.

Table 3. Means and standard deviations for Pearson's correlation coefficient (*r*), mean absolute error (MAE), and root mean square error (RMSE) of machine learning models using image traits as predictor variables.

Predicted trait	**Model**	**Validation dataset**		
		r	**MAE**	**RMSE**
BW	LR	0.982 ± 0.016 b	43.233 ± 4.161 b	63.420 ± 24.667 b
	ANN	0.982 ± 0.017 b	43.076 ± 4.152 b	63.663 ± 26.310 b
	DL	0.986 ± 0.003 b	52.782 ± 10.704 a	67.285 ± 12.419 a
	BRNN	**0.987 ± 0.006 a**	**39.545 ± 3.423 c**	**54.700 ± 11.254 c**
	RF	0.986 ± 0.003 b	43.764 ± 3.746 b	58.584 ± 5.479 ab
CW	LR	0.975 ± 0.014 ab	29.391 ± 2.806 bc	44.077 ± 12.102 ab
	ANN	0.975 ± 0.015 ab	29.324 ± 2.864 bc	44.292 ± 13.011 ab
	DL	0.979 ± 0.006 ab	33.633 ± 5.817 a	45.797 ± 7.460 a
	BRNN	**0.980 ± 0.009 a**	**27.820 ± 2.850 c**	**40.033 ± 7.652 b**
	RF	0.977 ± 0.006 b	30.015 ± 2.725 b	43.264 ± 6.410 ab
FW	LR	0.953 ± 0.025 a	24.158 ± 2.336 bc	35.828 ± 8.867 ab
	ANN	0.951 ± 0.029 a	24.096 ± 2.490 bc	36.107 ± 9.726 ab
	DL	0.957 ± 0.015 a	27.761 ± 4.439 a	37.834 ± 6.584 a
	BRNN	**0.959 ± 0.017 a**	**23.255 ± 2.240 c**	**33.416 ± 5.912 b**
	RF	0.955 ± 0.016 a	24.899 ± 2.326 b	35.158 ± 5.288 ab
FY	LR	0.281 ± 0.094 a	1.555 ± 0.156 b	2.254 ± 0.396 a
	ANN	0.280 ± 0.095 a	1.557 ± 0.155 b	2.265 ± 0.390 a
	DL	0.273 ± 0.096 a	1.635 ± 0.211 ab	2.350 ± 0.406 a
	BRNN	0.294 ± 0.094 a	1.545 ± 0.155 b	2.243 ± 0.393 a
	RF	0.164 ± 0.112 b	1.652 ± 0.161 a	2.369 ± 0.391 a

LR, linear regression; ANN, artificial neural networks; DL, deep learning; BRNN, Bayesian regularization for feed-forward neural networks; RF, random forests; BW, body weight; CW, carcass weight; FW, fillet weight; FY, fillet yield.

For each predicted trait, means followed by the same letters in the same column do not differ at the 5% significance level by Dunn's multiple comparison test with Bonferroni correction.

3.10. Supplementary material

Development of the Shoelace formula showed by Lee and Lim (2017) to calculate the area of triangle with vertices A (x_1, y_1), B (x_2, y_2), and C (x_3, y_3):

$$area\ (ABC) = \frac{1}{2}\left|\det\begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} + \det\begin{bmatrix} x_2 & x_3 \\ y_2 & y_3 \end{bmatrix} + \det\begin{bmatrix} x_3 & x_1 \\ y_3 & y_1 \end{bmatrix}\right|$$

$$area\ (ABC) = \frac{1}{2}\ |x_1y_2 - x_2y_1 + x_2y_3 - x_3y_2 + x_3y_1 - x_1y_3|$$

$$area\ (ABC) = \frac{1}{2}\ |(x_1y_2 + x_2y_3 + x_3y_1) - (x_2y_1 + x_3y_2 + x_1y_3)|$$

Extension of the Shoelace formula to an *n*-sided polygon:

$$area\ (A_1A_2 \ldots A_n) = \frac{1}{2}\left|\sum_{i=1}^{n} x_iy_{i+1} - x_{i+1}y_i\right|$$

where A_n is source vertex.

Table S1. Preliminary analysis performed to assess the significance of fixed and random effects for all studied traits.

	Random effects		Fixed effects[2]			
	AIC[1]		*p*-value			
Trait	M1	M2	Sex	Age	Age^2	Cage
BW	**15,970.13**	15,972.13	<0.001	<0.001	0.841	<0.001
TL	**4,626.23**	4,628.23	<0.001	<0.001	0.809	<0.001
HL	**2,848.07**	2,850.07	<0.001	<0.001	0.325	<0.001
SL	**5,208.18**	5,210.18	<0.001	<0.001	0.588	<0.001
H1	**2,051.07**	2,053.07	<0.001	<0.001	0.065	<0.001
H2	**3,423.06**	3,425.06	<0.001	<0.001	0.657	<0.001
H3	**3,655.37**	3,657.37	<0.001	<0.001	0.948	<0.001
H4	**3,321.16**	3,323.16	<0.001	<0.001	0.251	<0.001
H5	**1,317.65**	1,319.65	<0.001	<0.001	0.205	<0.001
TA	**11,581.53**	11,583.53	<0.001	<0.001	0.766	<0.001
HA	**8,113.31**	8,115.31	<0.001	<0.001	0.604	<0.001
FA	**10,644.50**	10,646.50	<0.001	<0.001	0.914	<0.001
TOT	**11,980.12**	11,982.12	<0.001	<0.001	0.878	<0.001
FW	**13,622.29**	13,674.80	<0.001	<0.001	0.607	<0.001
CW	**14,718.30**	14,720.30	<0.001	<0.001	0.987	<0.001
FY	**5,275.92**	5,277.92	0.713	0.016	0.257	<0.001
FA/TOT	**4,156.33**	4,158.33	<0.001	0.028	0.682	<0.001
TA/TOT	**4,050.71**	4,052.71	<0.001	0.002	0.096	<0.001
HA/TOT	**4,050.71**	4,052.71	<0.001	0.002	0.096	<0.001

[1]Akaike information criterion.

[2]Significance of fixed effects considering the best model (AIC values in bold).

BW, body weight; TL, trunk length; HL, head length; SL, standard length; H1, H2, H3, H4, and H5 represent heights 1, 2, 3, 4, and 5, respectively; TA, trunk area; HA, head area; FA, fillet area; TOT, total area; FW, fillet weight; FY, fillet yield; CW, carcass weight.

CHAPTER 4:

Final considerations

Digital image analysis is a promising tool for phenotyping in Nile tilapia selective breeding, given its non-invasive nature, fast operation, low cost, and reduced animal handling. In this study, we proposed a simple and fast method of image analysis (images in RGB color model - Red, Green, and Blue) to measure morphometric traits of Nile tilapia, especially those difficult to measure through traditional techniques, such as body areas. Additionally, moderate heritability estimates were observed for body areas, indicating the potential of these traits as selection criteria for targeted changes in Nile tilapia body shape. Also, the selection for body areas may lead to indirect genetic gains in body weight and fillet weight.

Digital images also can provide explanatory variables for posterior prediction of growth, carcass, and fillet traits through machine learning models. In this context, we evaluated five supervised machine learning models to predict body weight, carcass weight, fillet weight, and fillet yield, using as predictor variables the traits measured by digital image analysis (lengths, heights, and body areas). In general, the Bayesian regularization for the feed-forward neural network method showed the best performance in predicting body weight, carcass weight, and fillet weight. The findings from this thesis will contribute to the advance of precision phenotyping in Nile tilapia breeding programs.

ATTACHMENTS

Attachment 1. Project approved by the Animal Use Ethics Committee at *Universidade Estadual de Maringá*.

UEM Universidade Estadual de Maringá

Comissão de Ética no Uso de Animais da Universidade Estadual de Maringá

CERTIFICADO

Certificamos que a proposta intitulada "Fenotipagem por imagens digitais via sensor Kinect® aplicada ao melhoramento genético de tilápia do Nilo", protocolada sob o CEUA nº 9452160720 (ID 002749), sob a responsabilidade de **Carlos Antonio Lopes de Oliveira** *e equipe; Alex Júnio da Silva Cardoso; Fabyano Fonseca e Silva* - que envolve a produção, manutenção e/ou utilização de animais pertencentes ao filo Chordata, subfilo Vertebrata (exceto o homem), para fins de pesquisa científica ou ensino - está de acordo com os preceitos da Lei 11.794 de 8 de outubro de 2008, com o Decreto 6.899 de 15 de julho de 2009, bem como com as normas editadas pelo Conselho Nacional de Controle da Experimentação Animal (CONCEA), e foi **aprovada** pela Comissão de Ética no Uso de Animais da Universidade Estadual de Maringá (CEUA/UEM) na reunião de 03/09/2020.

We certify that the proposal "Phenotyping by digital images via Kinect® sensor applied to genetic improvement of Nile tilapia", utilizing 1200 Fishes (males and females), protocol number CEUA 9452160720 (ID 002749), under the responsibility of **Carlos Antonio Lopes de Oliveira** *and team; Alex Júnio da Silva Cardoso; Fabyano Fonseca e Silva* - which involves the production, maintenance and/or use of animals belonging to the phylum Chordata, subphylum Vertebrata (except human beings), for scientific research purposes or teaching - is in accordance with Law 11.794 of October 8, 2008, Decree 6899 of July 15, 2009, as well as with the rules issued by the National Council for Control of Animal Experimentation (CONCEA), and was **approved** by the Ethic Committee on Animal Use of the State University of Maringá (CEUA/UEM) in the meeting of 09/03/2020.

Finalidade da Proposta: Pesquisa

Vigência da Proposta: de 11/2020 a 10/2021 Área: Dzo-Zootecnia

Origem:	Setor de Piscicultura de Floriano						
Espécie:	Peixes	sexo:	Machos e Fêmeas	idade:	300 a 350 dias	N:	1200
Linhagem:	Tilápia (Oreochromis niloticus)			Peso:	800 a 1400 g		

Local do experimento: Unidade Demonstrativa de produção em tanques rede, da Universidade Estadual de Maringá, Campus de Diamante do Norte-PR;

Maringá, 03 de setembro de 2020

Tatiana

Profa. Dra. Tatiana Carlesso dos Santos
Coordenadora da CEUA/UEM
Universidade Estadual de Maringá

Profa. Dra. Erika Seki Kioshima Cótica
Coordenadora Adjunta da CEUA/UEM
Universidade Estadual de Maringá

Attachment 2. Project approved by the Ethics Commission on the Use of Farm Animals at *Universidade Federal de Viçosa*.

UNIVERSIDADE FEDERAL DE VIÇOSA
COMISSÃO DE ÉTICA NO USO DE ANIMAIS DE PRODUÇÃO
CEUAP/UFV

Campus Universitário – Viçosa, MG – 36570-900 – Telefone:(31) 3899.3275 – e-mail: ceuap@ufv.br – site: www.ceuap.ufv.br

Viçosa, 20 de Nov. de 2020

CERTIFICADO

Certificamos que o projeto intitulado **"Fenotipagem por imagens digitais via sensor Kinect® aplicada ao melhoramento genético de tilápia do Nilo"**, protocolo **nº 049/2020**, sob a responsabilidade de **Fabyano Fonseca e Silva -** que envolve a produção, manutenção e/ou utilização de animais pertencentes ao filo chordata, subfilo vertebrata (exceto o homem), para fins de pesquisa científica (ou ensino) - encontra-se de acordo com os preceitos da lei nº 11.794, de 8 de outubro de 2008, do decreto nº 6.899, de 15 de julho de 2009, e com as normas editadas pelo Conselho Nacional de Controle da Experimentação Animal (CONCEA), e foi apreciado pela Comissão de Ética no Uso de Animais de Produção da Universidade Federal de Viçosa (CEUAP-UFV) em reunião de **03 de Nov. de 2020.**

Finalidade: **(x)Pesquisa** **( )Ensino** Vigência do Projeto: de **20 de Nov. de 2020** a **01 de out. de 2021**

Espécie/linhagem: **Tilápia do Nilo *(Oreochromis niloticus)*** Nº de animais: **1200**

Peso: **800 a 1400 g** Idade: **300 a 350 dias** Sexo: **Macho/Fêmea** Origem: **Estação de Piscicultura da Universidade Estadual de Maringá, Paraná, Brasil. Cnpj/CPF: 79.151.312/0001-56 Endereço: Estr. Bravin - Distrito Industrial 2, Maringá - PR Responsável: Carlos Antonio Lopes de Oliveira**

CERTIFICATE

We certify that the project entitled **"Phenotyping by digital images via Kinect® sensor applied to genetic improvement of Nile tilapia",** protocol **nº 049/2020**, under the responsibility of **Fabyano Fonseca e Silva** - which involves the production, maintenance and/or use of animals belonging to the phylum chordata, subphylum vertebrata (except man), for scientific research purposes (or education) - is in accordance with the law nº. 11.794, of October 8, 2008, Decree nº. 6899 of July 15, 2009, and the rules issued by the Brazilian National Council for Animal Experimentation Control (CONCEA), and was approved by the Ethics Commission on the use of farm animals of Universidade Federal de Viçosa (CEUAP-UFV) in its meeting on **Nov. 03th, of 2020**.

Finality: **(x)Research** **( )Education**

Duration of the Project: from **Nov. 20th, of 2020** to **Oct. 1st, of 2021**.

Species / strain: **Nile Tilapia *(Oreochromis niloticus)*** Nº of animals: **1200**

Weight: **800 a 1400 g** Age: **300 to 350 days** Sex: **Male/Female** Source: **Estação de Piscicultura da Universidade Estadual de Maringá, Paraná, Brasil. Cnpj/CPF: 79.151.312/0001-56 Endereço: Estr. Bravin - Distrito Industrial 2, Maringá - PR Responsável: Carlos Antonio Lopes de Oliveira**

Luciana Navajas Rennó

Luciana Navajas Rennó
Coordenadora da CEUAP/UFV