Aquaculture 560 (2022) 738522

Contents lists available at ScienceDirect

Aquaculture

journal homepage: www.elsevier.com/locate/aquaculture

Quantitative trait loci controlling swimming performance and their effect on growth in Nile tilapia (*Oreochromis niloticus*)

Xiaofei Yu [a,*], Samuel Bekele Mengistu [a,b], Han A. Mulder [a], Arjan P. Palstra [a], John A. H. Benzie [c,d], Trong Quoc Trinh [c], Martien A.M. Groenen [a], Hans Komen [a], Hendrik-Jan Megens [a]

[a] Animal Breeding and Genomics, Wageningen University & Research, Wageningen, the Netherlands
[b] School of Animal and Range Sciences, College of Agriculture, Hawassa University, Hawassa, Ethiopia
[c] WorldFish Centre, Jalan Batu Maung, Bayan Lepas, Penang, Malaysia
[d] School of Biological Earth and Environmental Sciences, University College Cork, Cork, Ireland

ARTICLE INFO

Keywords:
Nile tilapia
Hypoxia
Swimming performance
Growth
QTL

ABSTRACT

Critical swimming speed (U_{crit}) is an important measurement of swimming performance and a good indicator for cardio-respiratory health. It offers a new opportunity to select fish with better fitness. However, the genomic architecture of swimming performance at whole genome level is not clear in Nile tilapia. For this study, swimming performance was measured in 1500 fish from the Genetic Improvement of Farmed Tilapia strain in their early life, which were subsequently grown in a non-aerated pond (nocturnal hypoxia) until harvest. Our results showed that the heritability for U_{crit} was 0.31 ± 0.04. Genetic correlations between U_{crit} and harvest weight (−0.13 ± 0.13) and between U_{crit} and daily growth coefficient (DGC) (−0.26 ± 0.13) were slightly negative. Nine single nucleotide polymorphisms (SNPs) were found to be suggestively associated with U_{crit}, of which five are located in a region between 12.18 and 19.89 Mb on linkage group (LG)14, while two SNPs are located between 18.85 Mb to 18.94 Mb on LG13. The remaining two SNPs are located on LG19 and LG12, respectively. Candidate genes in high linkage disequilibrium (LD) with these SNPs were identified, including *hip1*, *hectd1*, *elna*, *smyd1b*, *rrp12* and *pprc1*. This suggests possible involvement of neuronal growth, muscle activity, cardiovascular development and angiogenesis, and oxygen/hypoxia regulation. Three of these nine SNPs were significantly associated with both harvest weight and DGC, and SNP genotypes that associated with lowest mean U_{crit} were associated with highest mean harvest weight and DGC. In conclusion, we found a clear pleiotropic effect of some SNPs that affect both growth and swimming performance in a hypoxic environment, while other SNPs had only effect on swimming performance, but not on growth. Although fast swimming fish are assumed to show slower growth, such as lower DGC and harvest weight, candidate genetic markers identified in this study provide an opportunity to select fish with good cardio-respiratory health and growth.

1. Introduction

Tilapia is the second most important farmed fish species for aquaculture production. It is farmed in >120 countries across the world in a wide range of culture environments. Global tilapia aquaculture production grew 11% annually and increased from 383,654 metric tons in 1990 to almost 7,000,000 metric tons in 2020 according to FAO Fisheries and Aquaculture statistics (El-Sayed, 2019) and the worldwide tilapia market was US$ 7.9 billion in 2020 (from IMARC, https://www.imarcgroup.com/tilapia-market). Many selective breeding programs have been established (Tayamen, 2004; Ponzoni et al., 2011; Thodesen et al., 2011). One of the most important breeding programs is called the "Genetic Improvement of Farmed Tilapia" (GIFT) implemented by WorldFish in Malaysia, which has played an important role in boosting tilapia production in many countries in Asia and the Pacific region (Ponzoni et al., 2011; Bentsen et al., 2017). However, a yield gap is often observed in environments where fish are farmed without aeration. Low levels of dissolved oxygen, can adversely affect growth, feed conversion ratio and survival in Nile tilapia (Mengistu et al., 2020b). For most smallholder farmers, aeration of fishponds is not available or too expensive. In non-aerated ponds extreme hypoxia (below the critical level of 3 mg/L) can frequently occur, especially during the night when

* Corresponding author.
E-mail address: xiaofei.yu@wur.nl (X. Yu).

https://doi.org/10.1016/j.aquaculture.2022.738522
Received 13 January 2022; Received in revised form 9 June 2022; Accepted 12 June 2022
Available online 18 June 2022

algae become net oxygen consumers (Mengistu et al., 2020a; Yu et al., 2021). Therefore, it is crucial to select tilapia that grow better and healthier under conditions where dissolved oxygen is limited (such as hypoxia) by breeding companies and organizations. Smallholder tilapia farmers can then use these for more efficient fingerling production.

Swimming performance of fishes has been widely studied over half a century (Kieffer, 2010). Swimming performance is an important feature that correlates with fitness, survival and metabolism in aquacultural species (Palstra and Planas, 2012). It also plays a crucial role in several other aspects, such as migration, habitat selection, predator-prey interaction, and reproduction (Cano-Barbacil et al., 2020). Prolonged swimming performance in fish can be measured in a critical swimming challenge test, by which critical swimming speed (U_{crit}) can be assessed. During the test, swimming speeds are incrementally increased at prescribed intervals until fish become fatigued and stop swimming. The moment of fatigue determines the U_{crit} which is primarily aerobically driven and at which maximum oxygen uptake occurs (as smoothened average over this last swimming period but with significantly higher peaks during burst-and-glide swimming (Brett, 1964; Plaut, 2001; Kieffer, 2010; Palstra and Planas, 2012; Zhang et al., 2019). In general, U_{crit} is used as indicator to evaluate aerobic swimming performance and physical fitness in fish, similar to using a treadmill for human or rodents. For marine fish species such as Atlantic salmon (*Salmo sala*) and Gilthead Seabream (*Sparus aurata*), U_{crit} was used as a predictor for target traits such as growth performance in relation to feed intake and fillet percentage (Palstra et al., 2020). For freshwater species, Tudorache et al. (2008) performed an integrated research on swimming capacity and energy use in seven European freshwater fish species, showing that U_{crit} and oxygen consumption both were positively correlated to migration capacity. In tilapia, U_{crit} has been shown to be highly correlated with maximum metabolic rate (McKenzie et al., 2003). It therefore is reasonable to state that U_{crit} is strongly positively correlated to fitness. A previous study from our group (Mengistu et al., 2021) based on pedigree information, showed the existence of additive genetic variance for U_{crit} in Nile tilapia, and a favorable genetic correlation between U_{crit} and body weight, standard length, height and surface area at swimming test.

Selective breeding of farmed animals for economically important traits has high potential to increase aquaculture production. Traditional breeding strategies are to select families or individuals with excellent traits as parents to set up brood stocks for improvement, which were reported to be time-consuming and labor intensive, especially relatively slow or unstable when selecting economic traits that are determined by multiply segregating loci (Guo et al., 2022). Single nucleotide polymorphisms (SNPs) are DNA sequence variants within the genome. They can be used to map quantitative trait loci (QTLs) and genes affecting traits of interest. The analysis of QTLs for marker-assisted selection is more effective to accelerate genetic gain and reduces economic cost of bringing progeny to maturity compared to traditional breeding strategies (Ma et al., 2021). However, it is still unknown which quantitative trait loci (QTLs) and genes are associated with U_{crit} in tilapia. It is known that athletic performance is heritable in several species, such as human (Guth and Roth, 2013), horse (Schröder et al., 2011), and dog (Kim et al., 2018), and that variations in candidate genes involved in specific metabolic pathways play an important role in this trait. For instance, Ben-Zaken et al. (2017) suggested that the mutations C-1245 T (rs35767) in insulin-like growth factor and Lys(K)-153Arg(R) in myostatin are strongly associated with skeletal muscle phenotypes in human, which are beneficial for endurance and short-distance running. Over the past decade, several studies have been conducted to identify genes that play an important role in determining the swimming performance in skeletal and cardiac muscles of fishes. In Atlantic salmon, Robinson et al. (2017) detected putative SNPs associated with aerobic exercise and swimming performance between wild and domesticated stocks. Analysis of those SNPs showed that they mapped to genes involved in energetic processes, coding for contractile filaments in the muscle and controlling cell proliferation. Raffini et al. (2020) identified genes involved in swimming behaviour, physiology and oxygen intake differently expressed in the gill, by comparing divergent body shapes of two lake cichlid species. However, their contribution to genetic variation of swimming performance and their influence on growth performance in tilapia is not clear.

Hence, we decided to further explore the genomic architecture including SNPs and QTLs associated with swimming performance using the experimental data described in Mengistu et al. (2021). We also identify the effect of candidate SNPs and QTLs for Ucrit on growth performance. All fish from this study were genotyped using the Axiom® SNP array, which contains 65 K SNP markers dispersed over the Nile tilapia reference genome (Peñaloza et al., 2020). The main objectives were: (1) to estimate the genetic correlations between U_{crit} with growth traits at early life stage, and with growth traits at later life stage using a genomic relationship matrix under hypoxia the same condition as used by most smallholder tilapia farmers, (2) to identify QTLs associated with swimming performance and (3) to estimate the effect of significant QTLs for U_{crit} on the growth traits under hypoxia. Overall, this knowledge will help in prioritizing SNPs and QTLs to select for better growth and healthier fish in tilapia breeding programs.

2. Material and methods

2.1. Ethics statement

Phenotypic measurements and sampling of the GIFT strain were conducted as part of the GIFT selective breeding program managed by WorldFish at the Aquaculture Extension Centre. All fish in the GIFT breeding population are managed in accordance with the Guiding Principles of the Animal Care, Welfare and Ethics Policy of WorldFish.

2.2. Experimental design and traits collection

Nile tilapia were part of the GIFT selective breeding program. Sixty families were produced using 31 males and 58 females. The mating ratios were designed for one male to at least two females. The successful mating were: 12 males each mated with one female (resulting in 12 full sib families), 12 males each mated with two females (resulting in 12 half sib groups equivalent to 24 full sib families), 4 males each mated with 3 females (four half sib groups equivalent to 12 full sib families) and 3 males each mated with 4 females (three half sib groups equivalent to 12 full sib families), which consist a total of 60 full sib families. Each full sib family was reared separately in a hapa (fine mesh net enclosure) set up in an earthen pond. A nursing hapa has a dimension of $1.0 \times 1.0 \times 1.0$ m, each was stocked with 120 fry. Thirty to 35 fingerlings from each family were selected, anesthetized using clove oil and then individually tagged using PIT (Passive Integrated Transponder) tags using intraperitoneal injection the position of the PIT tag was in abdominal cavity (peritoneum). The mean body size of fish at swim testing was 10.8 g and the size of the PIT tag is about 95 mg. The swimming test was performed three weeks after PIT tagging. From each family, 25 fish in a range from 5 to 10 cm standard length were selected for the swimming test using a ruler with a centimetre scale. Body weight (BWtest in g) and photographs were made one day before the swimming test. Standard length (SLtest) and height (Htest) at swimming test of the fish were obtained from each fish photographs using image analysis as described previously by Mengistu et al. (2020a). Surface area (SAtest) of Nile tilapia was calculated as:

$$\mathrm{SAtest} = \frac{1}{4}\pi^{*}\mathrm{SLtest}^{*}\mathrm{Htest} \quad (1)$$

Fish feeding was stopped 24 h before the beginning of the swimming test. The fish were acclimatized for one hour in the swimming flume without flow. The critical swimming test was executed using a Brett type rectangular oval shape raceway with a swimming compartment on one side and a propellor for inducing flow on the other side as described by

Mengistu et al. (2021). In short, the swim flume measured 230 cm in length and 90 cm in width with a water depth of 40 cm, and was equipped with a Minn Kota Terrova 80 lbs. propeller. After every 30 mins, the velocity was increased to the next speed interval until the fish fatigued. At each setting, the average water flow velocity was recorded using a FP111 Global Water Flow Probe (Mengistu et al., 2021, supplementary Table 1 for mean water velocities and standard deviations at each propeller speed setting). The mean water temperature in the tank was 28.3 ± 0.6 °C during the swim test. The swimming test takes maximally 4.5 h with 9 propeller speed level. A fish fatigue is defined when it touched the back fence and could not be stimulated to continue swimming. Each fatigued fish was taken out immediately, and fatigue time was recorded. Fatigue time was used to calculate the absolute critical swimming speed (U_{crit}) as below (Brett, 1964):

$$U_{crit} = U_{-1} + \left(\frac{t}{\Delta t}\right)\Delta U \quad (2)$$

where U_{-1} is the highest velocity maintained for the full-time period, t is the time to fatigue at final velocity level in minutes, Δt is the time each velocity level is maintained at (= 30 min), ΔU is velocity increment in cm/s.

The caudal fin clip tissue samples were cut from each individual fish using a 3 mm diameter hole punch before stocking in the pond. Fin clips were preserved in 95% ethanol and stored at −20 °C until DNA extraction. The experimental design and data collection is published by Mengistu et al. (2021). After the swimming test, fish were stocked in a non-aerated pond for grow-out and harvested after 145 or 146 days to keep similar hypoxic conditions as in most smallholder tilapia farming systems. The dissolved oxygen level varied from 0.91 mg/L to 6.21 mg/L measured every 2 h for 24 h using Eco-Sense® DO200A reported in our previous study (Yu et al., 2021). The stocking and harvest weight (Harw) were recorded and daily growth coefficient (DGC) was calculated as below (Bureau et al., 2000).

$$DGC = \left[\frac{\sqrt[3]{harvest\ weight} - \sqrt[3]{stocking\ weight}}{time\ in\ days}\right] \times 100 \quad (3)$$

2.3. SNP genotyping and quality control

A SNP array is a powerful high-throughput genotyping tool to characterize genome-wide single-nucleotide polymorphisms (SNPs). The DNA was extracted and genotyped by Identigen (Dublin, Ireland) using an Axiom® SNP array, which contains 65 K SNP markers dispersed over the Nile tilapia reference genome (Peñaloza et al., 2020). The raw data from SNP array genotyping was imported to the Axiom analysis Suite version 4.0.3 software for genotype calling and quality control. Data was filtered to meet a dish quality control (DQC) >0.82 and call rate for samples (CR) >0.93, respectively. Next, a second quality control step was applied based on per SNP call rate and minor allele frequency (MAF) statistics using PLINK v1.90 (Purcell et al., 2007). A genotype call rate threshold (>90%) was set for SNP filtering. SNPs with MAF higher than 5% were retained. A total of 1388 fish and 51,438 SNPs were used for subsequent analyses.

2.4. Descriptive statistics and genetic parameters estimation

Basic statistics of phenotype data was analysed in R (4.0.2). The difference of traits between male and female were compared by unpaired two-samples *t*-test. We built the genomic relationship matrix (GRM) with 51,438 SNPs using calc_grm program with vanraden2 option (Calus and Vandenplas, 2013). Variance components and heritabilities for traits including U_{crit}, SLtest, Htest, SAtest, BWtest, Harw and DGC were estimated using univariate models based on residual maximum likelihood method using ASReml version 4.1.0 (Gilmour et al., 2015). The following model was applied:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_1a} + \mathbf{Z_2c} + \mathbf{e} \quad (4)$$

where **y** is a vector with observations for one trait, being U_{crit}, SLtest, Htest, SAtest, BWtest, Harw and DGC, and **b** is a vector with fixed effects. Test day and sex are significant factors for SLtest, Htest, SAtest and BWtest; hence, only test day was fitted as class variable for U_{crit}, while age at harvest as a covariate and sex as a class variable for Harw; weight at stocking as a covariate and sex as a class variable for DGC. **a** is a vector of the additive genetic effects of individuals and was assumed to be distributed as $N(\mathbf{0}, \mathbf{G}\sigma_a^2)$, with **G** the genomic relationship matrix and σ_a^2 is the additive genetic variance; **c** is a vector of the common environmental effects, and was assumed to be distributed as N $(\mathbf{0}, \mathbf{I}\sigma_c^2)$, **e** is a vector with the random residual and is assumed to be distributed as $N(\mathbf{0}, \mathbf{I}\sigma_e^2)$, with **I** the identity matrix and σ_e^2 is the residual variance; **X**, $\mathbf{Z_1}$ and $\mathbf{Z_2}$ are design matrices assigning trait values to the fixed effects, additive genetic effects and common environmental effects. Heritability (h^2) was estimated as the ratio of additive genetic variance to the phenotypic variance. Phenotypic and genetic correlations between U_{crit} and traits such as SLtest, Htest, SAtest, BWtest, Harw and DGC were estimated based on a bivariate linear model. The log-likelihood for the bivariate model did not converge when the common environmental effect was fitted as random effect, therefore the common environmental effect was excluded in all bivariate models. The fixed effects were the same as in the univariate models for U_{crit}, SLtest, Htest, SAtest, BWtest, Harw and DGC traits.

2.5. Association analysis for swimming performance

As a Gaussian distribution is assumed for phenotypes in an association test of quantitative trait (Goh and Yap, 2009), the normality of traits including U_{crit}, Harw and DGC was tested with a Shapiro-wilk test. Because it is impossible to completely follow an absolute normal distribution in most cases, we normalized them by the square root method (McDonald, 2009). To test significance of factors in the experiment, a linear model was conducted using Stepwise Algorithm (Neerchal et al., 2014). The fix factors were the same as in the univariate models. Only test day was fitted as class variable for U_{crit}; age at harvest as a covariate and sex as a class variable for Harw, while stocking weight as a covariate and sex as a class variable for DGC. Once the most appropriate linear model had been fitted, residuals were extracted for the subsequent association analysis (Gondro et al., 2013).

Population structure and kinships can be confounding factors in genome-wide association studies (Hoffman, 2013). The top five principal components were added as covariates and included in subsequent GWAS models as fixed effect in this association model, since we observed a slight family structure. All SNPs that passed the quality control were used to generate the genomic relationship matrix.

Animal phenotypes and genotypes were jointly analysed to identify genomic regions associated with U_{crit}, Harw and DGC. The following model was applied:

$$\mathbf{y} = \mathbf{W\alpha} + \mathbf{Z\mu} + \mathbf{x\beta} + \boldsymbol{\varepsilon} \quad (5)$$

where **y** is the vector of each trait from (U_{crit}, Harw and DGC); **α** is a vector of associated fixed effects, **μ** is the vector of additive genetic effects, **W** and **Z** are the corresponding design matrices, **x** being the vector of SNP genotypes and **β** their associated effects, ε is vector with the residual effects. Analyses were conducted using the GEMMA software (Zhou and Stephens, 2012). As the Bonferroni method is overly conservative, SimpleM method which based on the effective number of independent tests (Gao et al., 2010), was used to calculate the suggestive (5.15*E*-05) and genome-wide significance (2.57E-06) thresholds. The empirical *p*-values are based on the Wald tests. Manhattan and quantile-quantile (Q-Q) plots were generated through the "qqman" package (https://cran.r-project.org/web/packages/qqman/). The inflation factor λ was calculated to indicate the influence of population structure in

the association analyses.

2.6. Candidate genes in QTLs

Candidate regions associated with swimming performances were characterized within a 200 kb window size (100 kb upstream and downstream) flanking the candidate SNPs. LD and haplotype blocks were analysed with LDBlockShow (Dong et al., 2020), while $r^2 > 0.8$ as cut-off for blocks. Candidate genes were defined as genes located in the same haplotype blocks as the candidate SNPs or nearby candidate SNPs if no block is present. Afterwards, candidate genes were functionally annotated based on the latest tilapia reference genome (O_niloticus_UMD_NMBU) downloaded from NCBI Genome database (https://www.ncbi.nlm.nih.gov/assembly/GCF_001858045.2). To better understand the function of tilapia genes, we performed a BLAST against zebra fish (*Danio rerio*) proteins based on the genome (GCF_000002035.6_GRCz11), using a threshold of E-value <1e-6.

2.7. The effect of candidate SNPs in swimming performance on growth

In order to investigate the influence on growth of candidate SNPs associated with U_{crit}, we estimated the genetic association between SNPs and growth traits (DGC, Harw) based on generalized linear model in SNPassoc (González et al., 2007). The best fitting genetic model was evaluated based on Akaike Information Criterion (AIC) score. The significant threshold was defined as adjusted *P*-value (FDR) < 0.05.

3. Results

3.1. Descriptive statistics

A total of 1500 fish (60 families and 25 fish per family) were tested for swimming performance. After the swimming test, fish were stocked and allowed to grow out in non-aerated pond, except 260 fish that were dead before stocking. Comparing the U_{crit} of the dead fish to those surviving, showed no significant difference (*P* value = 0.29). Descriptive statistics of U_{crit}, SLtest, Htest, BWtest, SAtest, Harw and DGC for further analyses are presented in Table 1. U_{crit} was not significant different between males and females. However, all other traits presented a significantly higher mean value in males compared to that in females.

3.2. SNP distribution, allele frequency and family structure

In total 55,119 SNPs were exported from quality control in the Axiom Analysis Suite software (v4.03) and further investigated for their distribution, minor allele frequency and family structure. Those SNPs were distributed across the whole genome (Fig. 1a). The highest number of SNPs (4457) was seen for LG7, while the lowest number of SNPs (249) was found on an unplaced contig. The SNP distribution was mostly consistent with the physical length of the linkage groups according to the *Oreochromis niloticus* genome assembly (O_niloticus_UMD_NMBU). The majority of the SNPs belonged to the common (MAF ≥ 0.3) and intermediate (0.3 > MAF ≥ 0.1) groups, which consisted of 19,757 and 21,198 SNPs respectively (Supplementary Fig. 1).

The PCA based on 51,438 informative SNPs showed genetic variation amongst those 60 GIFT families. The first, second and third components explained 15.6%, 13.2% and 11.5% genotype variation, respectively (Fig. 1b). However, it seems that there are several clusters that represent families.

Table 1
Summary statistics of traits for all swimming tested fish.

Trait	Sex	No.	Mean	SE	t-value	Effect of sex (P value)
	male	680	69.05	0.21	−0.12	NS (0.91)
Ucrit (cm/s)	female	702	69.09	0.20		
	male	683	7.24	0.02	4.20	**S (2.84E-05)**
SLtest (cm)	female	705	7.11	0.02		
	male	683	2.74	0.01	5.01	**S (6.15E-07)**
Htest (cm)	female	705	2.67	0.01		
	male	683	11.07	0.10	4.56	**S (5.56E-06)**
BWtest (g)	female	705	10.45	0.09		
	male	683	15.67	0.10	4.82	**S (1.63E-06)**
SAtest (cm^2)	female	705	15.02	0.09		
	male	537	457.1	3.82	16.4	**S (< 2.2e-16)**
Harw (g)	female	589	379.5	2.78		
DGC ($g^{1/3}$ /d)	male	537	3.18	0.01	9.62	**S (< 2.2e-16)**
	female	589	3.00	0.01		

S: significant; NS: non-significant; standard length (SLtest), height (Htest), surface area (SAtest); body weight (BWtest); harvest weight (Harw); daily growth coefficient (DGC) (*P* < 0.05 was set as significance threshold).

3.3. Phenotypic and genetic parameter estimation with different traits at swimming test

Variance components and heritabilities from univariate models for different traits including U_{crit}, SLtest, Htest, SAtest and BWtest are presented in Table 2. The h^2 for U_{crit} was 0.31 with standard error 0.043. The variance explained by common environment was not significant ranging from 1.20E-06 to 3.10E-03. The h^2 for the remaining four traits at the swimming test (SLtest, Htest, SAtest and BWtest) were relatively close, ranging from 0.23 to 0.27. The h^2 estimates for harvest weight and DGC were moderate with 0.29 ± 0.047 and 0.30 ± 0.046, respectively. Phenotypic and genetic correlations are shown in Table 3. The genetic correlation between U_{crit} and other traits at the time of swimming test ranged from 0.35 to 0.43, indicating that fish with better swimming capacity usually show a longer standard length, larger body height, weight and body area compared to fish with poorer swimming capacity. The estimated r_g between U_{crit} and harvest weight (−0.13 ± 0.13) and U_{crit} and DGC (−0.26 ± 0.13) were slightly negative but not different from zero, suggesting that U_{crit} might have a negative genetic correlation to body weight at harvest and growth until harvest.

3.4. Genome-wide association study for swimming performance and growth

The linear mixed model was implemented to identify QTLs associated with critical swimming speed and growth. The *P* values of corrected thresholds for 5% genome-wide significant levels and suggestive association were 2.57E-06 and 5.15E-05, respectively. The genome-wide association results are shown in Fig. 2, while the inflation factor lambda (λ) was estimated to be 0.97, suggesting there is little population stratification present in the association results. In total, nine SNPs located on LG12, LG13, LG14 and LG19 exceeded the genome-wide suggestive threshold for U_{crit} (Fig. 2a), while none exceeded the genome-wide significant threshold. Five out of 9 SNPs were located between 12.18 and 19.89 Mb on LG14. Two SNPs were located between 18.85 and 18.94 Mb on LG13, while the remaining two SNPs were located on LG19:3659540 and LG12:24066436 (Supplementary Table 1). Five of the suggestive SNPs showed the minor allele (ranging from 0.074 to 0.485) to be associated with improved swimming performance and indicated potential for maker-assisted selection for this trait. For harvest weight (Fig. 2b), only four SNPs on LG19 exceeded the genome-wide suggestive threshold. However, there is a peak located on LG16 between 22.55 Mb to 38.36 Mb for DGC (Fig. 2c), which include in total 7 SNPs over the suggestive threshold. Notably, there were three overlapping peaks above the suggestive line from LG7, LG16 and LG21 for both harvest weight and DGC, although a peak on LG11 did not attain any statistical significance.

3.5. Candidate genes in the QTL regions

Candidate genes were defined as genes located in the same haplotype

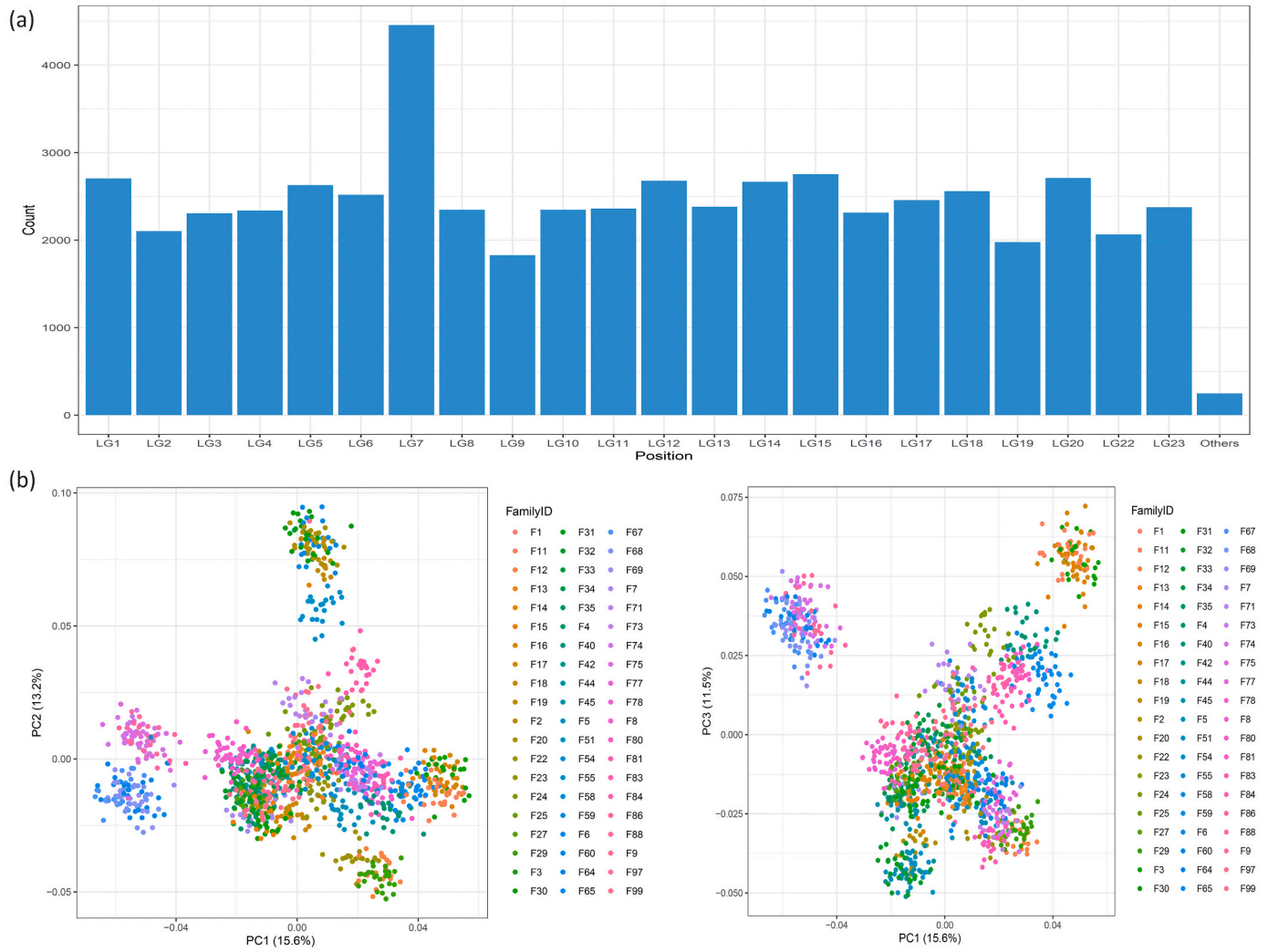

Fig. 1. SNP distribution generated from SNP array (a), PC plot on SNPs (b), each dot represents one individual.

Table 2
Genetic variance components and heritability for traits at swimming test in GIFT population.

Traits	σ_a^2	σ_e^2	σ_c^2	σ_p^2	h^2
U_{crit}	5.15	11.72	1.20E-06	16.88	0.31 (0.043)
SLtest	0.05	0.18	1.70E-04	0.23	0.24 (0.042)
Htest	0.01	0.03	2.40E-05	0.04	0.26 (0.042)
SAtest	1.06	3.47	3.05E-03	4.53	0.23 (0.042)
BWtest	1.23	3.38	3.10E-03	4.61	0.27 (0.043)
Harw	1515.89	3652.26	2.23	5170.40	0.29 (0.047)
DGC	0.02	0.05	0.35E-09	0.08	0.30 (0.046)

U_{crit} when only test day was fitted as class variable in the model; SLtest, Htest, SAtest and BWtest were estimated when test day and sex were included in the model as class variables; age at harvest as a covariate and sex as a class variable for Harw; weight at stocking as covariate and sex as a class variable for DGC.

Table 3
Genetic and phenotypic correlations between U_{crit} and other traits at swimming test, harvest weight, daily growth coefficient.

Traits	r_g	r_p
BWtest	0.43 ± 0.10	0.36 ± 0.03
SLtest	0.42 ± 0.11	0.35 ± 0.03
Htest	0.35 ± 0.11	0.30 ± 0.03
SAtest	0.39 ± 0.11	0.33 ± 0.03
Harw	−0.13 ± 0.13	0.06 ± 0.04
DGC	−0.26 ± 0.13	−0.01 ± 0.04

In a bivariate model, the fixed effect were the same as in the univariate models above except the common environmental effect was excluded in all models.

blocks as the candidate SNPs or nearby candidate SNPs if no block is present (as showed in Supplementary Fig. 2**)**. A summary of all candidate genes and their functions is shown in Table 4. The SNPs with the highest significance for U_{crit} were located in an intron of the *hip1* (huntingtin interacting protein 1) gene on LG14. *Hip1* is mainly involved in regulating the central nervous system and body length. Other candidate genes on LG14 were *limk1a*, *elna*, *lsamp*, *aipl1* and *dner*. Candidate genes on LG13 were *pprc1* and *rrp12*, while two other candidate genes located on LG19 and LG12 were *hectd1* and *smyd1*, respectively. The functional annotation for these candidate genes suggests the involvement of several relevant biological processes, including neuronal growth, muscle activity, cardiovascular development and oxygen/hypoxia regulation.

3.6. *The effect of suggestive SNPs from U_{crit} on growth*

To investigate the effect of candidate SNPs on growth, the significance threshold was defined with FDR multiple testing value <0.05. The summary for the effect of the nine suggestive SNPs with DGC is shown in Supplementary Table 2. One significant SNP with DGC (AX-317442766) followed a dominant genetic model, which means the genotype

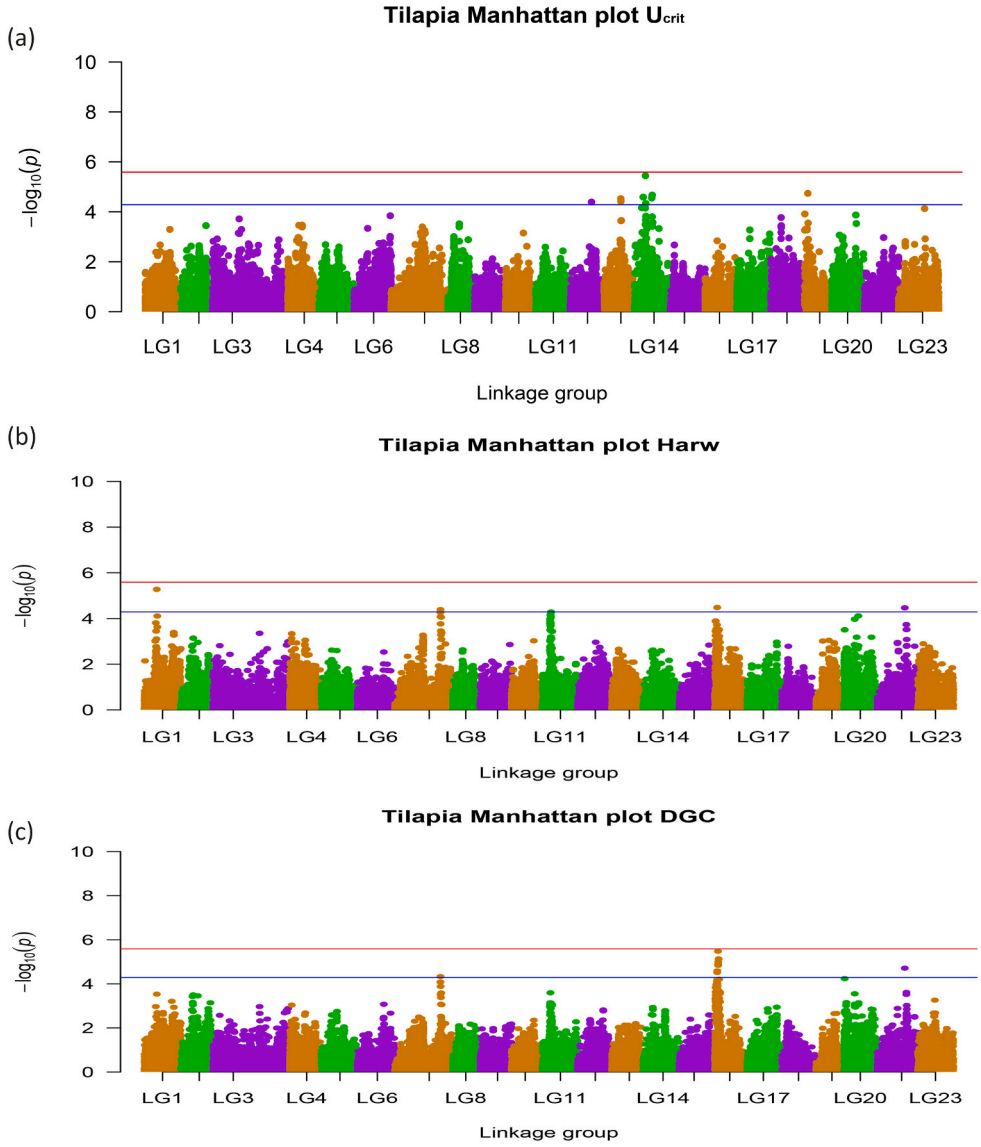

Fig. 2. Manhattan plot for U_{crit} (a), Harw (b) and DGC (c). The orange and blue horizontal line represent the genome-wide significance (2.57E-06) and suggestive significance threshold value (5.15E-05) respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

homozygous for the major allele (TT) had significantly higher DGC (3.12) compared to the mean DGC (3.07) of CT or CC genotype. Two SNPs (AX-317595114 and AX-317418538) were significantly associated with DGC following an over-dominant genetic model. For SNP AX-317595114, the heterozygous genotype (TC) had a significantly higher DGC (3.11) compared to the mean DGC of CC and TT (3.06) genotype. However, for SNP AX-317418538, the heterozygous genotype (AG) had a significantly lower DGC (3.04) compared to the mean DGC of GG and AA (3.10). From the summary for the effect of the nine SNPs with harvest weight (Supplementary Table 3**)**, we found that three SNPs (AX-317442766, AX-317595114 and 317,418,538) were significant, the same for DGC. We also investigated the effect size of the significant SNPs, to understand the interplay amongst U_{crit}, harvest weight and DGC**.** We observed that the TT genotype of SNP AX-317442766 had a lower mean U_{crit} than the genotypes CT and CC but had the highest mean harvest weight and DGC (as shown in Fig. 3a). The same is observed for AX-317418538 (Fig. 3c), in which the GG genotype had a lower mean U_{crit} than genotype AA and AG but had the highest mean harvest weight and DGC. For SNP AX-317595114 (Fig. 3b), the TT genotype had a highest mean U_{crit}, but also a lowest harvest weight and DGC compared to genotype TC and CC. Overall, three of these 9 SNPs were significantly associated with both harvest weight and DGC, and SNP genotypes that presented a low mean U_{crit} had a high mean harvest weight and DGC.

4. Discussion

Following the earlier study of swimming performance in tilapia (Mengistu et al., 2021), we re-examined the genetic correlations between U_{crit} with body size traits at the swimming test and with growth traits after a growing-out period in a non-aerated pond using a genomic relationship matrix. To better understand the genomic architecture of swimming performance in Nile tilapia, we identified QTLs associated with critical swimming speed, and estimated the effect of candidate SNPs for U_{crit} on the growth traits.

4.1. Heritability and genetic correlation

Heritabilities for all traits investigated in this study was moderate to high. Our results confirm those from the earlier study by Mengistu et al. (2021) and provide further evidence and support that swimming

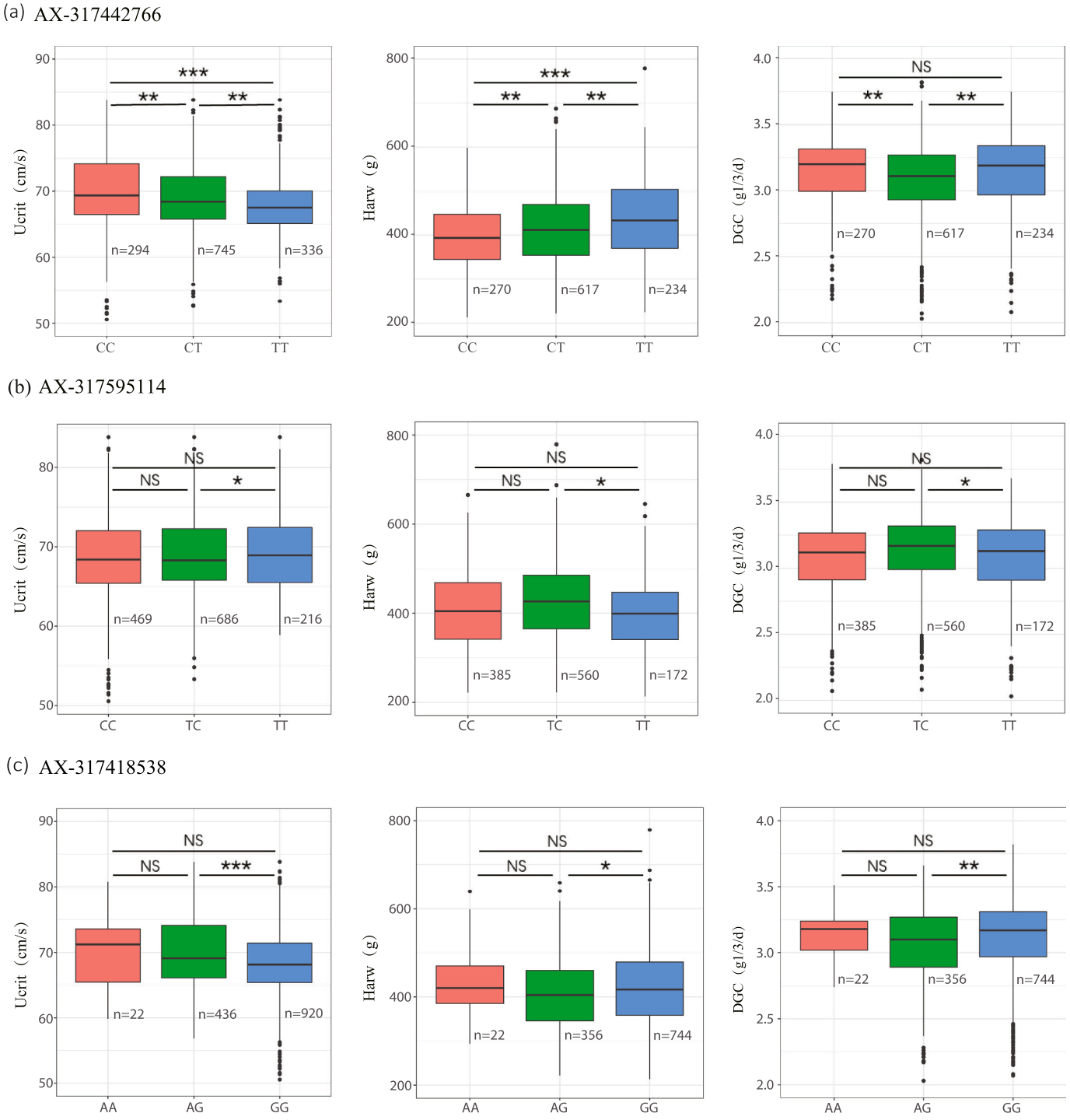

Fig. 3. The effect estimation of candidate SNPs on growth with AX-317442766 (a), AX-317418538 (b) and AX-31739145 (c).

performance is heritable and potentially applicable in genomic selection. The standard error of the heritability estimates was smaller using the genomic relationship matrix method compared to a pedigree-based relationship matrix. The genetic variance estimated for U_{crit} based on genomic relationship matrix is 5.15 in this study while 6.71 in previous study estimated by pedigree relationships by Mengistu et al. (2021). This confirms the idea that variance components are more accurately estimated when using a genomic relationship matrix, which not only utilizes the relationships of an individual with all other individuals in the analysis, but also captures the Mendelian sampling variance (Veerkamp et al., 2011). Our heritability estimate of U_{crit} is within the range with that in other species e.g. 0.24 in Guppy (*Poecilia reticulata*) (Nicoletto, 1995), 0.24 in Atlantic salmon (*Salmo salar*) (Hurley and Schom, 1984), 0.41 in threespine stickleback (*Gasterosteus aculeatus*) (Garenc et al., 1998) and 0.55 in European sea bass (*Dicentrarchus labrax*) (Vandeputte et al., 2016).

Similar to the heritability, the genetic correlations based on the genomic relationship matrix were also more precise compared to those based on pedigree information. The positive phenotypic genetic correlations between U_{crit} and SLtest, Htest, SAtest, BWtest indicate that larger fish swam faster in absolute terms during the swimming test. The observed genetic correlations between U_{crit} and Harw, and U_{crit} and

Table 4
Candidate genes from suggestive SNPs and their biological functions.

LG	rs	pos	Candidate genes	Functional annotation
14	AX-317442766	12,189,511	*hip1*	central nervous system, body length
19	AX-317595114	3,659,540	*hectd1*	hypoxia regulator
14	AX-317032241	19,888,564	LOC112842187 (ncRNA) LOC100693040 (*limk1a*) LOC102081909 (ncRNA) LOC100694912 (*elna*)	cardiovascular development
14	AX-317417214	9,648,027	*lsamp*	neuronal growth
14	AX-317424375	19,382,001	*aipl1* trarg1 LOC109204983 (ncRNA) *gosr1*	photoreceptor
13	AX-317391485	18,854,572	*pprc1*	exercise-related muscle activity
13	AX-317415431	18,943,978	*rrp12*	oxygen adaptation
12	AX-323027008	24,066,436	LOC100708974 (*smyd1b*)	skeletal and cardiac muscles activity
14	AX-317418538	12,496,362	*dner*	neuronal and glial differentiation

Gene names in brackets were annotated based on a BALST against *Danio rerio*.

DGC, suggests an interaction between growth and swimming performance. A negative genetic correlation indicates that Nile tilapia with higher U_{crit} have lower growth and harvest weight in non-aerated ponds. In juvenile common carp (*Cyprinus carpio*), when acclimatized to the lower temperature, U_{crit} was negatively correlated with feeding rate and growth rate, suggesting a trade-off between growth and exercise. But this trade-off in juvenile common carp (*Cyprinus carpio*) can disappear when acclimatized to higher temperature (Pang et al., 2016). On the other hand, a negative correlation between growth and swimming performance could occur in species that mature in their first year, like tilapia. Growth rate becomes positively correlated with U_{crit} when fish mature within a few months after hatching, this pattern has been found in European sea bass (*Dicentrarchus labrax*) (Cominassi et al., 2019), Atlantic Salmon (*Salmo salar*) and Gilthead Seabream (*Sparus aurata*) (Palstra et al., 2020). In addition, selection for high growth rate can lead to the development of cardiac abnormalities, which has been observed in several farmed species, such as broiler chicken (Olkowski, 2007) and rainbow trout (*Oncorhynchus mykiss*) (Brijs et al., 2020). The GIFT strain has been selected for harvest weight over fifteen generations so far. The negative correlations between U_{crit} and growth traits in the study, seems to suggest that a low U_{crit} could lead to cardio-respiratory health problems in the future. These results indicate that selection for better growth in combination with selection for fish with good cardio-respiratory health is feasible in a breeding program, which will aid in providing fingerlings resilient to (periodic) hypoxic culture conditions within smallholder tilapia farming.

4.2. Genome-wide association study

Our results show several QTLs involved in the swimming performance in Nile tilapia. It is well-known that the larger the effect of the QTL, the higher the chance of finding a significant association for a given size of study, i.e. higher power. Our analyses showed limited numbers of clear associations, and even where candidate regions were found, the effect generally seemed limited in size. However, our study was only based on 1500 samples. Increasing the number would increase the power and stronger associations would be expected to emerge. While some medium to strong associations were found, the results also indicated that U_{crit} is a complex trait regulated by many genes with small effect.

Overall, 16 genes in high linkage disequilibrium (LD) with the identified SNPs, were found. The 65 K SNP array used in this study covers <2% of all SNPs presented in the GIFT population (Cádiz et al., 2020). It is more likely candidate SNPs are in strong LD with the casual SNPs rather than causal SNPs. The SNP with the highest significance (LG14:12189511) was linked to the gene *hip1*, which codes for huntingtin-interacting protein. The gene has been suggested to confer narrow body shape in zebrafish (Komoike et al., 2010), whereas *HIP1* in human is presumed to affect cognitive and central nervous function (Metzler et al., 2003; Ramocki et al., 2010). Two other genes (*Lsamp* and *dner*) are also potentially involved in neuronal growth and differentiation (Pimenta et al., 1995; Hsieh et al., 2013). The candidate gene *hectd1* codes for HECT domain E3 ubiquitin ligase 1. A zebrafish (*Danio rerio*) orthologue of the human and mouse *HECTD1* gene is annotated in the ZFIN database. The expression level of *HECTD1* decreased under hypoxia in human cells (Erler, 2014; Wang et al., 2020). The response to oxygen levels in human cells suggests *hectd1* in tilapia may be an interesting candidate gene under similar challenging circumstances. Similarly, the *rrp12* gene has orthologues found in both human and mouse, which have been associated with adaptation to low oxygen. The human orthologue of *RRP12* was found to be associated with living at high altitude in Asian and Amerindian populations (Foll et al., 2014). Further candidate genes identified in the present study include *elna*, for which variant (elnasa12235 c.264 T > A, p.Tyr88*) has been shown to reduced blood flow and to induce heart function abnormalities in zebrafish (*Danio rerio*) (Zorrilla, 2018). Furthermore, zebrafish (*Danio rerio*) and human orthologues of two genes (*smyd1b* and *pprc1*) are potentially involved in cardiac and exercise-related muscle activity (Li et al., 2013; Uguccioni et al., 2010).

Previous studies have shown that skeletal and cardiac muscle tissues are significantly influenced by swimming-induced exercise in several species, such as Atlantic salmon (*Salmo salar*) (Castro et al., 2013) and zebrafish (*Danio rerio*) (Rovira, 2017). Sustained swimming could increase transcriptional activity in white muscle during growth and development in rainbow trout (*Oncorhynchus mykiss*) (Palstra et al., 2013). We further confirmed that a few candidate genes were also involved in skeletal and cardiac muscle in Nile tilapia.

In summary, while the majority of genes involved in the trait probably cannot be fully identified by a study of the scope and size of the current one, we were able to identify a number of good candidates. It was found that pathways including central nervous system and neuron development, oxygen adaptation and hypoxia regulation, cardiac and exercise-related muscle activity are potentially involved in regulation of swimming performance. Overall, these 9 SNPs involved in muscle activity, cardiovascular development and angiogenesis, and oxygen/hypoxia regulation pathways, should be prioritized for marker-assisted selection in breeding program.

4.3. Pleiotropic SNPs with swimming performance and growth

We also investigated the impact of candidate SNPs associated with U_{crit} on growth. As the function of genes becomes better known, it is often revealed that they are involved in multiple pathways, and therefore have multiple effects on the phenotype (pleiotropy). Pleiotropy is widespread in animal breeding. When one trait is under selection, the mean of other traits also changes over generations. This response to selection could be reflected by the genetic correlation between traits (Gratten and Visscher, 2016). Knowing the genetic basis of traits and their pleiotropic effects through candidate gene and molecular pathway analysis can reveal how traits may be related at the functional genomic level.

Our results show only three of the SNPs are associated with harvest

weight and DGC. These three SNPs are linked to three genes (*hip1*, *hectd1*, *dner)*. Gene *hip1* regulates body length, therefore it is not surprising that *hip1* is also associated with growth. The gene *hectd1* is assumed to be involved in adaptation to hypoxia (Wang et al., 2020) and our previous studies have shown that hypoxia can suppress growth in Nile tilapia (Mengistu et al., 2020a; Yu et al., 2021). The candidate genes involved in cardiovascular development and exercise-related muscle activity suggested no relationship with growth. Nevertheless, they potentially can be applied as biomarkers to select fish with good cardio-respiratory health and good growth. Genes involved in neuronal growth and differentiation pathway can influence growth while others do not. After all, growth traits are also polygenic with potential involvement of thousands of QTLs (Zhang et al., 2021).

5. Conclusions

Moderate heritability of swimming performance was found based on genomic relationship matrix estimation. Large fish generally swim faster than smaller fish, whereas fish with better swimming performance showed slower growth later in life. Nine suggestive SNPs between genotype and swimming performance were identified. The identified QTLs indicate that swimming performance is a complex trait with pleiotropic effects on growth. Our results reveal a clear pleiotropic effect of some SNPs associated with swimming performance on growth traits including harvest weight and DGC, while other SNPs had only effect on swimming performance, but not on growth. Using these SNPs in selection has the potential to select fish with good cardio-respiratory health and good growth. Overall, our analyses shed a first light on the genomic mechanisms of swimming performance and growth in tilapia.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aquaculture.2022.738522.

Funding

This work was funded by Koepon Foundation, International Fund for Agricultural Development (IFAD) Grant number: 2000001001, the CGIAR Research Program on Fish Agri-Food Systems (FISH) led by WorldFish. The FISH program is supported by contributors to the CGIAR Trust Fund. The authors highly thank the WorldFish breeding team and the Department of Fisheries Malaysia for their technical support during the experimental phase. The authors also thank to the financial support from China Scholarship Council (CSC, File No. 201803250069), which has sponsored Xiaofei Yu's PhD study at Wageningen University and Research. The funding bodies played no role in design of the study and data collection, analysis, and result interpretation in the manuscript.

Author contributions

X.Y: Writing – original draft, Methodology, Formal analysis, Investigation, Methodology, Visualization. S.M: Writing – review & editing, Conceptualization, data curation, Formal analysis. H.M: Writing – review & editing, Conceptualization, Methodology, Investigation. A.P: Writing – review & editing, Conceptualization, Methodology, Investigation. J.B: Writing – review & editing, Conceptualization, Resources, Funding acquisition. T.T: Writing - Review & Editing, Data Curation, Methodology, Investigation. M.A.M.G: Writing – review & editing, Supervision, Investigation, Resources. H.K: Writing – review & editing, Supervision, Conceptualization, Investigation, Funding acquisition. H.J. M: Writing – original draft, Supervision, Methodology, Investigation.

Data availability

The genotype and phenotype data for swim-tested fish in this study are stored in the Harvard Dataverse repository (https://dataverse.harvard.edu/) with accession number JHQHS1. The authors declare that all data supporting the findings are available within this article and its supplementary files.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Bentsen, H.B., Gjerde, B., Eknath, A.E., de Vera, M.S.P., Velasco, R.R., Danting, J.C., Abella, T.A., 2017. Genetic improvement of farmed tilapias: response to five generations of selection for increased body weight at harvest in *Oreochromis niloticus* and the further impact of the project. Aquaculture 468, 206–217.

Ben-Zaken, S., Meckel, Y., Nemet, D., Eliakim, A., 2017. The combined frequency of IGF and myostatin polymorphism among track & field athletes and swimmers. Growth Hormon. IGF Res. 32, 29–32. https://doi.org/10.1016/j.ghir.2016.12.002.

Brett, J.R., 1964. The respiratory metabolism and swimming performance of young sockeye Salmon. J. Fish. Res. Board Can. 21, 1183–1226. https://doi.org/10.1139/f64-103.

Brijs, J., Hjelmstedt, P., Berg, C., Johansen, I.B., Sundh, H., Roques, J.A.C., Gräns, A., 2020. Prevalence and severity of cardiac abnormalities and arteriosclerosis in farmed rainbow trout (*Oncorhynchus mykiss*). Aquaculture 526, 735417. https://doi.org/10.1016/j.aquaculture.2020.735417.

Bureau, D.P., Azevedo, P.A., Tapia-Salazar, M., Cuzon, G., 2000. Pattern and cost of growth and nutrient deposition in fish and shrimp: potential implications and applications. In: Avances en Nutrición Acuícola V. Memorias del V Simposium Internacional de Nutrición Acuícola, 19, pp. 111–140.

Cádiz, M.I., López, M.E., Díaz-Domínguez, D., Cáceres, G., Yoshida, G.M., Gomez-Uchida, D., Yáñez, J.M., 2020. Whole genome re-sequencing reveals recent signatures of selection in three strains of farmed Nile tilapia (*Oreochromis niloticus*). Sci. Rep. 10, 1–14.

Calus, M.P.L., Vandenplas, J., 2013. Calc_grm—A Program to Compute Pedigree, Genomic, and Combined Relationship Matrices. Animal Breeding and Genomics Centre, Wageningen UR Livestock Research.

Cano-Barbacil, C., Radinger, J., Argudo, M., Rubio-Gracia, F., Vila-Gispert, A., García-Berthou, E., 2020. Key factors explaining critical swimming speed in freshwater fish: a review and statistical analysis for Iberian species. Sci. Rep. 10, 1–12. https://doi.org/10.1038/s41598-020-75974-x.

Castro, V., Grisdale-Helland, B., Helland, S.J., Torgersen, J., Kristensen, T., Claireaux, G., Takle, H., 2013. Cardiac molecular-acclimation mechanisms in response to swimming-induced exercise in Atlantic salmon. PLoS One 8, e55056.

Cominassi, L., Moyano, M., Claireaux, G., Howald, S., Mark, F.C., Zambonino-Infante, J. L., Peck, M.A., 2019. Combined effects of ocean acidification and temperature on larval and juvenile growth, development and swimming performance of European sea bass (*Dicentrarchus labrax*). PLoS One 14, 1–22. https://doi.org/10.1371/journal.pone.0221283.

Dong, S.-S., He, W.-M., Ji, J.-J., Zhang, C., Guo, Y., Yang, T.-L., 2020. LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. Brief. Bioinform. 00, 1–6. https://doi.org/10.1093/bib/bbaa227.

El-Sayed, A.-F.M., 2019. Tilapia Culture. Academic Press.

Erler, J.T., 2014. Identification of Hypoxia-Regulated Proteins Using MALDI-Mass Spectrometry Imaging Combined with Quantitative Proteomics.

Foll, M., Gaggiotti, O.E., Daub, J.T., Vatsiou, A., Excoffier, L., 2014. Widespread signals of convergent adaptation to high altitude in Asia and America. Am. J. Hum. Genet. 95, 394–407. https://doi.org/10.1016/j.ajhg.2014.09.002.

Gao, X., Becker, L.C., Becker, D.M., Starmer, J.D., Province, M.A., 2010. Avoiding the high Bonferroni penalty in genome-wide association studies. Genet. Epidemiol. 34, 100–105.

Garenc, C., Silversides, F.G., Guderley, H., 1998. Burst swimming and its enzymatic correlates in the threespine stickleback (*Gasterosteus aculeatus*): full-sib heritabilities. Can. J. Zool. 76, 680–688.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, Sj, Thompson, R., 2015. ASReml User Guide Release 4.1 Structural Specification. Hemel Hempstead: VSN International Ltd.

Goh, L., Yap, V.B., 2009. Effects of normalization on quantitative traits in association test. BMC Bioinform. 10 https://doi.org/10.1186/1471-2105-10-415.

Gondro, C., Van Der Werf, J.H., Hayes, B., 2013. Genome-Wide Association Studies and Genomic Prediction. Humana Press.

González, J.R., Armengol, L., Solé, X., Guinó, E., Mercader, J.M., Estivill, X., Moreno, V., 2007. SNPassoc: an R package to perform whole genome association studies. Bioinformatics 23, 644–645. https://doi.org/10.1093/bioinformatics/btm025.

Gratten, J., Visscher, P.M., 2016. Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. Genome Med. 8, 8–10. https://doi.org/10.1186/s13073-016-0332-x.

Guo, J., Wang, A., Mao, S., Xu, X., Li, J., Shen, Y., 2022. Construction of high-density genetic linkage map and QTL mapping for growth performance in black carp (*Mylopharyngodon piceus*). Aquaculture 549, 737799. https://doi.org/10.1016/j.aquaculture.2021.737799.

Guth, L.M., Roth, S.M., 2013. Genetic influence on athletic performance. Curr. Opin. Pediatr. 25, 653.

Hoffman, G.E., 2013. Correcting for population structure and kinship using the linear mixed model: theory and extensions. PLoS One 8. https://doi.org/10.1371/journal.pone.0075707.

Hsieh, F.Y., Ma, T.L., Shih, H.Y., Lin, S.J., Huang, C.W., Wang, H.Y., Cheng, Y.C., 2013. Dner inhibits neural progenitor proliferation and induces neuronal and glial differentiation in zebrafish. Dev. Biol. 375, 1–12. https://doi.org/10.1016/j.ydbio.2013.01.007.

Hurley, S.M., Schom, C.B., 1984. Genetic control of swimming stamina in Atlantic salmon (*Salmo salar*). Can. J. Genet. Cytol. 26, 57–61.

Kieffer, James D., 2010. Perspective - exercise in fish: 50+ years and going strong. In: Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology, 156, pp. 163–168. https://doi.org/10.1016/j.cbpa.2010.02.009.

Kim, J., Williams, F.J., Dreger, D.L., Plassais, J., Davis, B.W., Parker, H.G., Ostrander, E. A., 2018. Genetic selection of athletic success in sport-hunting dogs. Proc. Natl. Acad. Sci. U. S. A. 115, E7212–E7221. https://doi.org/10.1073/pnas.1800455115.

Komoike, Y., Fujii, K., Nishimura, A., Hiraki, Y., Hayashidani, M., Shimojima, K., Yamamoto, T., 2010. Zebrafish gene knockdowns imply roles for human YWHAG in infantile spasms and cardiomegaly. Genesis 48, 233–243. https://doi.org/10.1002/dvg.20607.

Li, H., Zhong, Y., Wang, Z., Gao, J., Xu, J., Chu, W., Du, S.J., 2013. Smyd1b is required for skeletal and cardiac muscle function in zebrafish. Mol. Biol. Cell 24, 3511–3521.

Ma, A., Huang, Z., Wang, X., An Xu, Y., Guo, X., 2021. Identification of quantitative trait loci associated with upper temperature tolerance in turbot, *Scophthalmus maximus*. Sci. Rep. 11, 1–12. https://doi.org/10.1038/s41598-021-01062-3.

McDonald, J.H., 2009. Handbook of Biological Statistics. Sparky House Publishing, Baltimore, MD.

McKenzie, D.J., Martinez, R., Morales, A., Acosta, J., Morales, R., Taylor, E.W., Estrada, M.P., 2003. Effects of growth hormone transgenesis on metabolic rate, exercise performance and hypoxia tolerance in tilapia hybrids. J. Fish Biol. 63, 398–409. https://doi.org/10.1046/j.1095-8649.2003.00162.x.

Mengistu, Samuel B., Mulder, H.A., Benzie, J.A.H., Khaw, H.L., Megens, H.J., Trinh, T.Q., Komen, H., 2020a. Genotype by environment interaction between aerated and non-aerated ponds and the impact of aeration on genetic parameters in Nile tilapia (*Oreochromis niloticus*). Aquaculture 529, 735704. https://doi.org/10.1016/j.aquaculture.2020.735704.

Mengistu, Samuel Bekele, Mulder, H.A., Benzie, J.A.H., Komen, H., 2020b. A systematic literature review of the major factors causing yield gap by affecting growth, feed conversion ratio and survival in Nile tilapia (*Oreochromis niloticus*). Rev. Aquac. 12, 524–541. https://doi.org/10.1111/raq.12331.

Mengistu, S.B., Palstra, A.P., Mulder, H.A., Benzie, J.A.H., Trinh, T.Q., Roozeboom, C., Komen, H., 2021. Heritable variation in swimming performance in Nile tilapia (*Oreochromis niloticus*) and negative genetic correlations with growth and harvest weight. Sci. Rep. 11, 1–11. https://doi.org/10.1038/s41598-021-90418-w.

Metzler, M., Li, B., Gan, L., Georgiou, J., Gutekunst, C.A., Wang, Y.T., Hayden, M.R., 2003. Disruption of the endocytic protein HIP1 results in neurological deficits and decreased AMPA receptor trafficking. EMBO J. 22, 3254–3266. https://doi.org/10.1093/emboj/cdg334.

Neerchal, N.K., Morel, J.G., Huang, X., Moluh, A., 2014. A stepwise algorithm for generalized linear mixed models. In: SAS Global Forum, pp. 1822–2014.

Nicoletto, P.F., 1995. Offspring quality and female choice in the guppy, *Poecilia reticulata*. Anim. Behav. 49, 377–387.

Olkowski, A.A., 2007. Pathophysiology of heart failure in broiler chickens: structural, biochemical, and molecular characteristics. Poult. Sci. 86, 999–1005. https://doi.org/10.1093/ps/86.5.999.

Palstra, A.P., Planas, J.V., 2012. Swimming Physiology of Fish: Towards Using Exercise to Farm a Fit Fish in Sustainable Aquaculture. Springer Science & Business Media. https://doi.org/10.1007/978-3-642-31049-2.

Palstra, A.P., Beltran, S., Burgerhout, E., Brittijn, S.A., Magnoni, L.J., Henkel, C.V., Planas, J.V., 2013. Deep RNA sequencing of the skeletal muscle transcriptome in swimming fish. PLoS One 8, e53171.

Palstra, A.P., Kals, J., Böhm, T., Bastiaansen, J.W.M., Komen, H., 2020. Swimming performance and oxygen consumption as non-lethal indicators of production traits in Atlantic Salmon and Gilthead Seabream. Front. Physiol. 11, 1–14. https://doi.org/10.3389/fphys.2020.00759.

Pang, X., Fu, S.J., Zhang, Y.G., 2016. Acclimation temperature alters the relationship between growth and swimming performance among juvenile common carp (*Cyprinus carpio*). Comp. Biochem. Physiol. -Part A: Mol. Integr. Physiol. 199, 111–119. https://doi.org/10.1016/j.cbpa.2016.06.011.

Peñaloza, C., Robledo, D., Barría, A., Trinh, T.Q., Mahmuddin, M., Wiener, P., Houston, R.D., 2020. Development and validation of an open access SNP array for Nile Tilapia (oreochromis niloticus). G3: Genes Genomes Genetics 10, 2777–2785. https://doi.org/10.1534/g3.120.401343.

Pimenta, A.F., Zhukareva, V., Barbe, M.F., Reinoso, B.S., Grimley, C., Henzel, W., Levitt, P., 1995. The limbic system-associated membrane protein is an Ig superfamily member that mediates selective neuronal growth and axon targeting. Neuron 15, 287–297. https://doi.org/10.1016/0896-6273(95)90034-9.

Plaut, I., 2001. Critical swimming speed: its ecological relevance. Comp. Biochem. Physiol. A Mol. Integr. Physiol. 131, 41–50.

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 world fish center with the GIFT strain. Rev. Aquac. 3, 27–41. https://doi.org/10.1111/j.1753-5131.2010.01041.x.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Daly, M. J., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.

Raffini, F., Schneider, R.F., Franchini, P., Kautt, A.F., Meyer, A., 2020. Diving into divergence: differentiation in swimming performances, physiology and gene expression between locally-adapted sympatric cichlid fishes. Mol. Ecol. 29, 1219–1234. https://doi.org/10.1111/mec.15304.

Ramocki, M.B., Bartnik, M., Szafranski, P., Kołodziejska, K.E., Xia, Z., Bravo, J., Stankiewicz, P., 2010. Recurrent distal 7q11.23 deletion including HIP1 and YWHAG identified in patients with intellectual disabilities, epilepsy, and neurobehavioral problems. Am. J. Hum. Genet. 87, 857–865. https://doi.org/10.1016/j.ajhg.2010.10.019.

Robinson, N.A., Timmerhaus, G., Baranski, M., Andersen, Ø., Takle, H., Krasnov, A., 2017. Training the salmon's genes: influence of aerobic exercise, swimming performance and selection on gene expression in Atlantic salmon. BMC Genomics 18, 1–19. https://doi.org/10.1186/s12864-017-4361-7.

Rovira, M., 2017. Skeletal Muscle and Cardiac Adaptations to Swimming-induced Exercise in Adult Zebrafish. PhD thesis.

Schröder, W., Klostermann, A., Distl, O., 2011. Candidate genes for physical performance in the horse. Vet. J. 190, 39–48. https://doi.org/10.1016/j.tvjl.2010.09.029.

Tayamen, M.M., 2004. Nationwide dissemination of GET-EXCEL tilapia in the Philippines. In: Proceedings of the Sixth International Symposium on Tilapia in Aquaculture.

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.

Tudorache, C., Viaene, P., Blust, R., Vereecken, H., De Boeck, G., 2008. A comparison of swimming capacity and energy use in seven European freshwater fish species. Ecol. Freshw. Fish 17, 284–291.

Uguccioni, G., D'souza, D., Hood, D.A., 2010. Regulation of PPAR coactivator-1 function and expression in muscle: effect of exercise. PPAR Res. 2010.

Vandeputte, M., Porte, J.D., Auperin, B., Dupont-Nivet, M., Vergnet, A., Valotaire, C., Chatain, B., 2016. Quantitative genetic variation for post-stress cortisol and swimming performance in growth-selected and control populations of European sea bass (*Dicentrarchus labrax*). Aquaculture 455, 1–7. https://doi.org/10.1016/j.aquaculture.2016.01.003.

Veerkamp, R.F., Mulder, H.A., Thompson, R., Calus, M.P.L., 2011. Genomic and pedigree-based genetic parameters for scarcely recorded traits when some animals are genotyped. J. Dairy Sci. 94, 4189–4197. https://doi.org/10.3168/jds.2011-4223.

Wang, X., De Geyter, C., Jia, Z., Peng, Y., Zhang, H., 2020. HECTD1 regulates the expression of SNAIL: implications for epithelial–mesenchymal transition. Int. J. Oncol. 56, 1186–1198. https://doi.org/10.3892/ijo.2020.5002.

Yu, X., Megens, H.J., Mengistu, S.B., Bastiaansen, J.W.M., Mulder, H.A., Benzie, J.A.H., Komen, H., 2021. Genome-wide association analysis of adaptation to oxygen stress in Nile tilapia (*Oreochromis niloticus*). BMC Genomics 22, 1–13. https://doi.org/10.1186/s12864-021-07486-5.

Zhang, Y., Gilbert, M.J.H., Farrell, A.P., 2019. Finding the peak of dynamic oxygen uptake during fatiguing exercise in fish. J. Exp. Biol. 222 https://doi.org/10.1242/jeb.196568.

Zhang, Y., Wang, Yuzhe, Li, Y., Wu, J., Wang, X., Bian, C., Kang, X., 2021. Genome-wide association study reveals the genetic determinism of growth traits in a Gushi-Anka F2 chicken population. Heredity 126, 293–307. https://doi.org/10.1038/s41437-020-00365-x.

Zhou, X., Stephens, M., 2012. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 44, 821–824.

Zorrilla, M., 2018. Contribution of Elastin to Cardiovascular Development in Zebrafish.