

Genotype by environment interaction between aerated and non-aerated ponds and the impact of aeration on genetic parameters in Nile tilapia (*Oreochromis niloticus*)

Samuel B. Mengistu^{a,b,c}, Han A. Mulder^{a,*}, John A.H. Benzie^{b,d}, Hooi Ling Khaw^{b,1}, Hendrik-Jan Megens^a, Trong Quoc Trinh^b, Hans Komen^a

^a Wageningen University & Research Animal Breeding and Genomics, P.O. Box 338, 6700, AH, Wageningen, the Netherlands

^b WorldFish, Jalan Batu Maung, Batu Maung, 11960 Bayan Lepas, Penang, Malaysia

^c School of Animal and Range Sciences, College of Agriculture, P. O. Box 5, Hawassa University, Hawassa, Ethiopia

^d School of Biological Earth and Environmental Sciences, University College Cork, Cork, Ireland

ARTICLE INFO

Keywords:

Nile tilapia
Genotype by environment interaction
Harvest weigh
Thermal growth coefficient
Survival
Dissolved oxygen

ABSTRACT

A major problem in smallholder Nile tilapia (*Oreochromis niloticus*) farms is that the achieved production is much lower than under optimal management. One of the main environmental factors contributing to lower production is dissolved oxygen (DO), because the majority of Nile tilapia production takes place under smallholder farms with no aeration of ponds which leads to large DO fluctuations. On the contrary, breeding programs select fish in aerated ponds. Aerating ponds is currently not an option for smallholder farmers because either it is too expensive or they lack access to cheap electricity supply. Therefore, it is crucial to know the genetic correlation between aerated and non-aerated ponds to optimize breeding programs to select fish that perform well in ponds with fluctuating DO levels. The objectives of this study were 1) to investigate the presence of genotype by environment (GxE) interaction between aerated and non-aerated earthen ponds using a design that minimized common environmental effects and 2) the impact of (non-)aeration on genetic parameters. The experimental fish were mass-produced using natural group spawning and nursed in four 30m² hapas. A random sample of fingerlings from each hapa was tagged and randomly distributed to aerated and non-aerated ponds for a grow-out period of 217 and 218 days. Body weight and photographs were taken on five consecutive time points during grow-out. Of the stocked fish, 2063 were genotyped-by-sequencing. A genomic relationship matrix was built using 11,929 SNPs to estimate genetic parameters with ASReml. No-aeration significantly reduced mean harvest weight (HW), survival and thermal growth coefficient (TGC) compared to aeration. Substantial heritabilities (0.14–0.45) were found for HW, TGC, surface area (SA) and body shape, expressed as ellipticity, and low heritabilities (0.03–0.04) for survival in aerated and non-aerated ponds. In both ponds, the environmental effect common to full sibs was not significant. Genetic coefficients of variation were 20–23% lower and heritabilities were 19–25% lower in the non-aerated pond compared to the aerated pond, for HW, TGC and survival. Genetic correlations between ponds for HW, standard length, height, SA and TGC were 0.81, 0.80, 0.74, 0.78 and 0.78, respectively. In summary, some GxE interaction between aerated and non-aerated ponds was found and no-aeration decreased genetic coefficients of variation and heritabilities compared to aerated ponds. Breeding programs are recommended to use half sib information from non-aerated farms or to set up a reference population for genomic selection in a non-aerated environment either on-station or in farms.

1. Introduction

A major problem in smallholder Nile tilapia (*Oreochromis niloticus*) farms is that the achieved production is much lower than under optimal management. This difference in performance is called yield gap

(Mengistu et al., 2019). One of the main environmental factors contributing to yield gap is dissolved oxygen (DO) (Mengistu et al., 2019), because the majority of Nile tilapia production takes place under smallholder farms with no aeration of ponds. No-aeration leads to large daily DO fluctuations. On the contrary, many breeding programs, like

* Corresponding author.

E-mail address: han.mulder@wur.nl (H.A. Mulder).

¹ Current address: GenoMar Genetics Group, Tjuvholmen Alle 11, 0252 Oslo, Norway.

Genetically Improved Farmed Tilapia (GIFT), are selecting fish that have been growing-out in aerated ponds, which means that there is a difference between the selection environment and the majority of the production environments that do not use aeration. A simple solution seems to be that small-holder farms should aerate their ponds, but for smallholder farmers aeration is either too expensive or they lack access to cheap and stable electricity supply (Mengistu et al., 2019). Therefore, it is crucial to optimize breeding programs to select fish that perform well in non-aerated ponds with fluctuating DO levels.

The key parameter for optimization of breeding programs in the presence of GxE is the genetic correlation between environments (Falconer, 1952; Mulder and Bijma, 2005). The presence of GxE leads to reranking, meaning that the best genotype in one environment is not the best genotype in another environment. In a number of studies in Nile tilapia, genetic correlations between fertilized pond with/without feed supplement, cage culture with feed supplement/commercial pellet feed and rice-fish culture (Bentsen et al., 2012), between cage and pond environments (Khaw et al., 2012) and between low and high input environments (Trong et al., 2013a) were estimated and found to be lower than unity (Bentsen et al., 2012; Khaw et al., 2012; Trong et al., 2013a). However, genetic correlation estimates for production traits between aerated and non-aerated ponds are lacking. It is hypothesized that no aeration may hinder fish from expressing their genetic potential and result in lower additive genetic variance compared to aerated ponds and that may also lead to genotype by environment interaction between aerated and non-aerated ponds.

Unbiased estimates of the genetic correlation between environments are crucial to optimize breeding programs, but the classical experiments using pedigree relationships and prolonged separate full sib family rearing are prone to yield biased estimates of the genetic correlation. The key problem is the presence of common environmental effects and properly disentangling those common environment effects from genetic effects. Nile tilapia families for selective breeding programs are traditionally produced by single pair mating, in a mating ratio of one male to two females (Komen and Trong, 2014). The separate rearing of full-sib families until tagging size (3-5 g) introduces common environmental effects, which explain 10–20% of the phenotypic variance (Gjerde et al., 2012; Thoa et al., 2016). Rearing until tagging size could typically take 2–3 months (Trong et al., 2013a). Full sib families resemble each other, because they share the same common environment and share part of their genes. To get more accurate and unbiased estimates of the genetic parameters, Lozano-Jaramillo et al. (2019) recommended 1:5 male to female mating ratios in the presence of common environmental effects and 1:1 ratios when there is no common environmental effect. However, it is practically difficult to attain a 1:5 mating ratio. In addition, prolonged communal rearing programs may increase the genetic correlation of traits in different environments and mask detection of GxE interaction, especially for harvest weight which is clearly the sum of the communal growth and the growth in the grow-out period (Dupont-Nivet et al., 2010). Therefore, to reduce the presence of common environmental effects and reduce the bias in the estimated genetic correlation, it is crucial to shorten the period of family production and use communal full sib rearing to get more accurate GxE estimates. One solution is to produce fry by natural group mating, e.g. groups of 7 males and 15 females or 12 males and 25 females (Fessehaye et al., 2006; Trong et al., 2013b) and later genotype the fish to estimate genetic parameters based on a genomic relationship matrix (GREML, VanRaden, 2008; Goddard et al., 2011), so removing the need for separate full sib family rearing until tagging size. Such a design with natural group mating and communal rearing is hypothesized to reduce the contribution of common environmental effects to the phenotypic variance. To test these hypotheses, the main objective of this paper was to investigate whether GxE exists between aerated and non-aerated earthen ponds and the impact of (non-)aeration on genetic parameters. The study was performed by designing a GxE experiment that minimized common environmental effects in Nile tilapia, based on natural group spawning, and G-BLUP parameter estimates, using GBS.

2. Material and methods

2.1. Family production and nursery

The experiment was carried out in the Aquaculture Extension Centre, Department of Fisheries, Jitra, Kedah State, Malaysia. The fish used in this experiment were mass produced from generation 16 of the GIFT breeding program as follows: the male and female breeders were separately conditioned for two weeks in cages in an earthen pond (3 × 3 × 1 m, mesh size 1 cm) before stocking them in mating hapas. Mating was done in four hapas (each 30 m²) in a 500 m² earthen pond, which was aerated by a paddlewheel. Eighteen males and 50 female breeders were stocked for 15 days in each of the mating hapas. In total 72 males and 200 females were used.

On the sixteenth day the breeders were removed, and the fry were kept in the same hapas for nursing for a duration of 60 days. Fry were fed commercial feed with 43% crude protein and 5% crude fat at a rate of 10–15% of body weight. The feed was divided into three portions and the fry were fed three times a day.

2.2. Grow-out and pond management

After 60 days of nursery, the fingerlings from each hapa were transferred into four aerated tanks and conditioned for three days before tagging. Feeding was stopped one day before tagging. The fingerlings from the same nursery hapa were combined in one aerated tank and a random sample of fingerlings was anaesthetised using clove oil and individually tagged using PIT (Passive Integrated Transponder) tags. At tagging, each fish was photographed, a 1 cm² fin clip sample collected and PIT tag number and body weight (BW) were recorded. The fin clip samples were preserved in 95% ethanol. The tagging, weight recording, fin clip sample collection and photographing was done in four consecutive days. A random sample of an equal number of individually tagged fingerlings from each nursery hapa was stocked in two earthen ponds. Totally 1570 fish were stocked in each pond with a stocking density of 3 fish/m².

The size of each of the ponds was 511 m² with a water depth of 1 to 1.2 m. The only treatment difference between the ponds was aeration. One of the ponds was aerated using a paddle wheel and blower to create a normoxic environment. The second pond had no aerator to create diurnal dissolved oxygen (DO) fluctuations, which is a typical feature of earthen ponds where green algae are the main source of oxygen (Fig. 1).

During the grow-out period feeding management was kept the same in both ponds. Commercial feed with 30% crude protein and 5% crude fat at a rate of 5% of body weight per day were used. After 2 months this was reduced to 3% of their body weight per day. The feeding rate was adjusted approximately every three weeks based on a sample of ~100 fish. It was also adjusted based on total biomass and number of fish recorded at each of three interval measurements (55/56 days, 104/105 days, and 167/168 days). The feed was divided into two portions and fed in the morning from 9:00 to 10:00 and afternoon from 15:00 to 16:00. Some morning feedings were skipped due to cloudy weather conditions that dropped the DO level in the non-aerated pond below 2 mg/l. At these concentrations, it was observed that fish no longer eat.

2.3. Records

Body weight and a photograph of each fish were recorded at stocking, at 55/56 days, 104/105 days, 167/168 days and at harvest, which was after 217 and 218 days of grow-out from the non-aerated pond and aerated pond, respectively. Body weight (g) of fish in each pond was recorded using a digital scale with an precision of one decimal of a gram. Next, each fish was photographed (Olympus OMD EM5 and EM10MKII) together with a metric (cm) ruler and unique labels. Sex was determined at 104/105 days, when the fish was on average 150 g. Survival was recorded as the number of days from stocking until

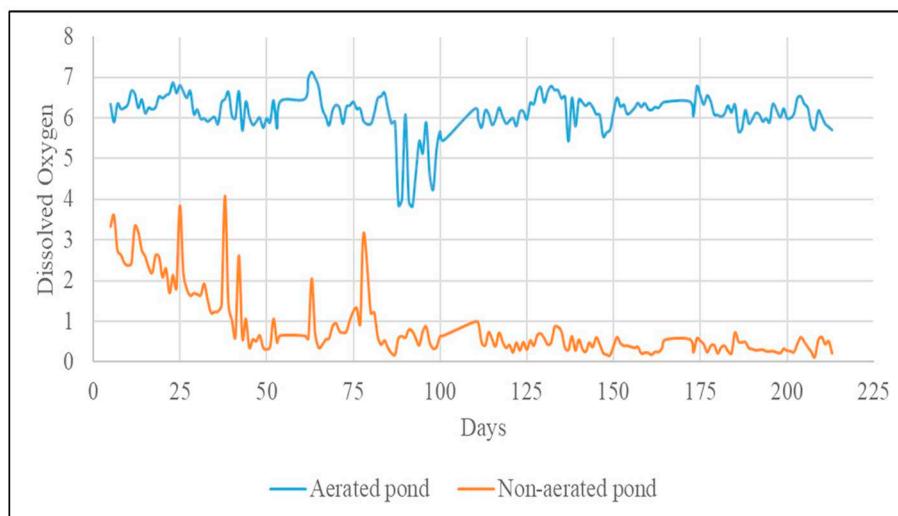


Fig. 1. Morning dissolved oxygen in the aerated and the non-aerated pond.

the last measurement the fish was observed alive. In total, fish were measured at stocking, at harvest and three times in between but the focus of this article is on traits at final harvest and survival days.

Standard length (SL) and body height (body depth) (H) measurements of fish were obtained from the picture of each fish taken at stocking and harvest. In total 2063 photographs that were taken at stocking and 1512 photographs that were taken at harvest were loaded into tpsUtil software (Rohlf, 2017b) and digitized for six landmarks using tpsDIG 2.30 (Rohlf, 2017a). The landmarks were as follows: landmarks 1 and 2 were on the 0 and 20 cm marks on the ruler which was photographed together with the fish for scaling. The landmarks 3 and 4 were used to measure SL, the distance between the tip of the snout to the base of caudal fin. The landmarks 5 and 6 were the dorsal and ventral landmarks where the distance is maximum. These landmarks were used to calculate H, the maximum dorso-ventral distance (see Fig. 2 for the landmarks). To get the distance between the Cartesian coordinates, they were analysed in R software using geomorph-package version 3.0.7 (Adams et al., 2018) and the true distance was computed based on the reference scale. Photographs for 174 fish at harvest from the non-aerated pond were either missing or the quality was poor. For these fish, the standard length and height were estimated based on linear regression of SL or H on body weight of fish from the same pond: (height = 6.41 + 0.007 * body weight, and length = 17.04 + 0.012 * body weight).

Body surface area (SA) of Nile tilapia is similar to the area of an ellipse and was calculated as:

$$SA = \frac{1}{4}\pi * SL * H \quad (1)$$

Ellipticity (Ec) (Blonk et al., 2010) was calculated as:

$$Ec = \frac{(SL - H)}{(SL + H)} \quad (2)$$

Thermal growth coefficient (TGC) (Jobling, 2003) was computed as:

$$TGC = [(\sqrt[3]{W_t} - \sqrt[3]{W_0}) / (T \times t)] \times 1000 \quad (3)$$

where W_t is harvest weight, W_0 is stocking weight, T is temperature in °C and t is time in days.

The trait survival days was recorded during three interval measurements and at harvest. The trait survival days utilizes the data better than a binomial 0–1 trait for survived and dead fish, because it accounts for the assumed day of mortality (Ellen et al., 2008; Wommongkol et al., 2018).

Dissolved oxygen and water temperature were measured three times a day just before the sunrise, at noon and just before sunset using EcoSense® DO200A. Each time the measurement was taken at three

random locations at about two meters from the pond side. Ammonia and pH were measured every week using DR3900 spectrophotometer and EcoTester pH 1, respectively.

2.4. DNA extraction, genotyping, alignment, variant calling and filtering

DNA was isolated from fin clips using the QIAamp DNeasy® 96 Blood and Tissue kit (QIAGEN #69581) following company specifications. DNA yield and quality were checked by full-spectrum spectrophotometer NanoDrop 2000 (Thermo Scientific) and Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). After qualification and quantification, DNA samples were subjected to GBS to identify single nucleotide polymorphisms (SNPs) across the genome.

GBS is a reduced representation approach that uses restriction enzymes to fragment the genome with subsequent size-selection. Individual DNA samples were digested by the *ApeKI* restriction enzyme and adapters added to both ends of the DNA segments (one end containing a unique barcode to allow pooling of samples in the sequencing process). Fragments of a size between 170 and 350 bp were amplified by polymerase chain reaction (PCR) and subsequently sequenced of these libraries on an Illumina HiSeq 2000 sequencing machine.

The cleaned sequence reads were aligned to the Nile tilapia (*Oreochromis niloticus*) reference genome GCF_001858045.1 (Conte et al., 2017, <https://www.ncbi.nlm.nih.gov/genome/197>) by using the BWA-mem algorithm (Version 0.1.75) (Li and Durbin, 2009), for every fish of 2171 individually. Initial alignments were re-aligned using GATK IndelRealigner (Van der Auwera et al., 2013), and subsequently sorted and indexed by samtools version 0.1.19 (Li et al., 2009). For efficient, parallelized, genotype calling using FreeBayes (Garrison and Marth, 2012) the genome was divided in 100 kb regions, initially filtered for SNPs to require to have a genotype call rate of at least 70% and a heterozygosity of at least 15%. Variants were further filtered to overlap with expected fragment sizes (in the range of 170–350 bp) based on in silico prediction of *ApeKI* restriction sites. Variants of all regions were concatenated and the final dataset consisted of 42,293 single nucleotide polymorphisms (SNPs). Further stringent quality control was applied using PLINK version 1.9 (Purcell et al., 2007; Purcell, 2018) parameters including the requirement that at least 90% of SNPs were successfully genotyped on all animals, SNP were required to have a minor allele frequency of above 2%, and a genotype call rate for individual fish was required to be at least 70% across all SNPs. The final dataset of 2063 individuals and 11,929 SNPs that passed the quality control thresholds was used for further analyses. The SNPs were distributed throughout the whole genome (Fig. 3 and Fig. 4).



Fig. 2. Nile tilapia picture with landmarks 1:6. Landmarks 1 and 2 marks a reference scale 20 cm length, landmarks 3 and 4 snout and base of caudal fin, respectively, landmarks 5 and 6 used to measure height (maximum dorso-ventral length) measure height of fish.

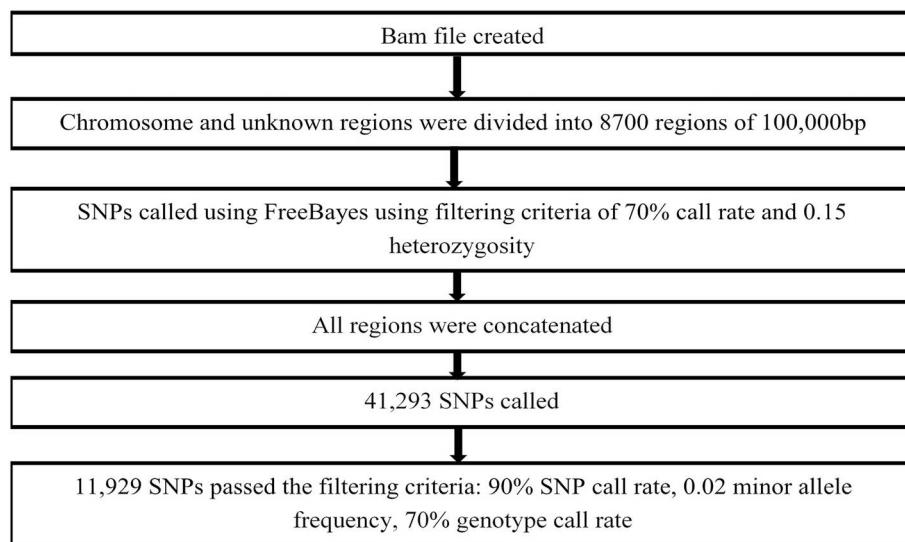


Fig. 3. Filtering and quality control flow chart.

2.4.1. Genomic relationship matrix

We computed a genomic relationship matrix (GRM) based on 11,293 SNPs using calc_grm software (Calus and Vandapel, 2016) using the vanraden2 option. The mean of the diagonal of the computed GRM was 0.83 (GRM1). By definition, the mean of the diagonal of a GRM should be one or higher and furthermore the GRM needs to be invertible. The lower than average diagonal is partly due to the high proportion of missing markers in GBS-data. Therefore, two adjusted GRMs were used:

GRM2: GRM1 was adjusted using the number of non-missing alleles per individual for self-relatedness (diagonal elements) and using the

non-missing alleles found on the two individuals for the off-diagonal elements. The adjustment factor k_{ij} for each element of GRM 1 was calculated as:

$$k_{ij} = N_{SNP_all}/(N_{SNP_all} - N_{SNP_missing}) \quad (4)$$

where N_{SNP_all} was the total number SNP-loci used (11,293 SNPs) and $N_{SNP_missing}$ was the number of SNP-loci missing in a particular individual for diagonal elements of the GRM. For off-diagonal elements of the GRM (i.e. the genomic relationship between two animals), $N_{SNP_missing}$ was the number of SNP loci missing in at least one of the two individuals; $N_{SNP_all} - N_{SNP_missing}$ was therefore the number of SNP-loci

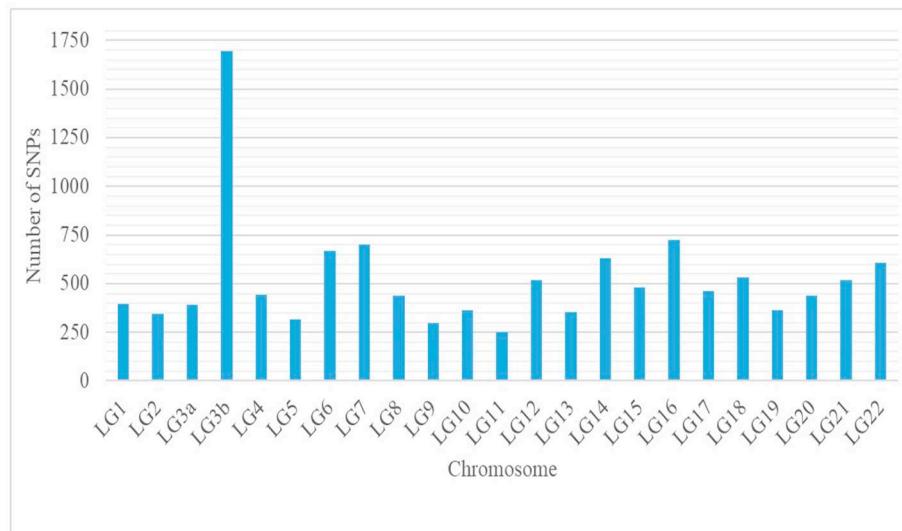


Fig. 4. Single nucleotide polymorphisms (SNPs) distribution.

with genotypes for both individuals.

The element of GRM2 was calculated as:

$$G_{2ij} = k_{ij} G_{ij} \quad (5)$$

GRM3: GRM 2 was multiplied with an extra adjustment factor to make the average of the diagonal elements equal to 1, because the average diagonal element in GRM2 was 0.842. GRM3 was calculated as

$$G_{3ij} = \frac{1}{0.842} G_{2ij} \quad (6)$$

Based on preliminary analysis the genetic parameter estimates from the three GRMs were similar (see Table 1 for the variance component estimates for harvest weight using the different GRMs). The analysis presented in this paper is based on GRM3.

2.5. Statistical analysis

2.5.1. Estimation of phenotypic and genetic parameters within pond

Firstly, variance components and heritabilities for HW, SL, H, SA, Ec, TGC and survival days within ponds were estimated using univariate models by residual maximum likelihood (REML), fitting an animal model with a genomic relationship matrix using ASReml version 4.1 (Gilmour et al., 2015). The model used was:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e} \quad (7)$$

where, \mathbf{y} is the vector of one trait from HW, SL, H, SA, Ec, TGC and survival days, \mathbf{b} is the vector of fixed effects, which were stocking weight, nursery hapas (1–4) and sex (female, male and unknown); \mathbf{a} is

Table 1

Additive genetic (σ_a^2) and residual (σ_e^2) variances and heritability estimates (h^2) and standard error (se) for harvest weight using different genomic relationship matrixes (GRMs) in aerated and non-aerated ponds.

Pond	GRM	σ_a^2	σ_e^2	h^2 (se)
Aerated	GRM1 ^a	10,744.9	28,525.9	0.27 (0.07)
	GRM2 ^b	10,013.4	28,809.4	0.26 (0.07)
	GRM3 ^c	8435.7	28,809.4	0.23 (0.06)
Non-aerated	GRM1	3367.2	12,371.8	0.21 (0.07)
	GRM2	3239.0	12,409.0	0.21 (0.07)
	GRM3	2728.67	12,409.0	0.18 (0.06)

^a GRM1 calculated using calc_grm software using vanraden2 option.

^b Calculated using Eqs. (4) and (5).

^c Calculated using Eq. (6).

the vector of random additive genetic effects with $\mathbf{N} = (\mathbf{0}, \mathbf{G}\sigma_a^2)$ where \mathbf{G} is the genomic relationship matrix and σ_a^2 is additive genetic variance, \mathbf{e} is the vector of residual effects with $\mathbf{N} = (\mathbf{I}\sigma_e^2)$ where \mathbf{I} is the identity matrix and σ_e^2 is the residual variance. The \mathbf{X} and \mathbf{Z} are design matrices assigning phenotypic values to the levels of fixed effects and additive genetic effects. Heritability (h^2) of each trait was computed as the ratio of additive genetic variance and phenotypic variance (σ_p^2), $h^2 = \frac{\sigma_a^2}{\sigma_p^2}$. For all traits, linear mixed models were used; for the trait survival days a linear mixed model violates the normality assumption, but a linear mixed model has been used before for such a trait and tends to yield similar results as more complex threshold models (Ellen et al., 2008; Wonmongkol et al., 2018).

Secondly, phenotypic and genetic correlations between different traits measured on the same individual within pond were estimated using bivariate linear models. For all the bivariate models, the fixed effects were the same as in the univariate models. The additive genetic effects were normally distributed as $\mathbf{N} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \mathbf{G} \otimes \begin{bmatrix} \sigma_{a,T1}^2 & r_{a,T12}\sigma_{a,T1}\sigma_{a,T2} \\ r_{a,T21}\sigma_{a,T1}\sigma_{a,T2} & \sigma_{a,T2}^2 \end{bmatrix} \right)$ where $\sigma_{a,T1}^2$ ($\sigma_{a,T2}^2$) being the additive genetic variance of trait 1 (trait 2) and $r_{a,T12(21)}$ is the additive genetic correlation between trait 1 and 2. The residual effects were normally distributed as $\mathbf{N} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \mathbf{I} \otimes \begin{bmatrix} \sigma_{e,T1}^2 & r_{e,T12}\sigma_{e,T1}\sigma_{e,T2} \\ r_{e,T21}\sigma_{e,T1}\sigma_{e,T2} & \sigma_{e,T2}^2 \end{bmatrix} \right)$ where $\sigma_{e,T1}^2$ ($\sigma_{e,T2}^2$) is the residual variances for trait 1 (trait 2) and $r_{e,T12(21)}$ is the residual correlation between trait 1 and 2.

2.5.2. Estimation of GxE between ponds

To investigate the degree of GxE between aerated and non-aerated ponds, the r_g between the same traits measured on different individuals in the aerated and non-aerated ponds were estimated with a bivariate model. The model used was:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e} \quad (8)$$

where, \mathbf{y} is the vector of either HW, SL, H, SA, Ec, TGC or survival days measured on different individuals in the aerated and non-aerated ponds. The environmental covariance was set to zero, because individual fish cannot be tested in two environments at the same time. The fixed effects and the genetic variance-covariance matrix are the same as described above except for the residual variance-covariance matrix (\mathbf{R}) $\begin{bmatrix} \sigma_{e,Ap}^2 & 0 \\ 0 & \sigma_{e,NAp}^2 \end{bmatrix}$ where $\sigma_{e,Ap}^2$ ($\sigma_{e,NAp}^2$) is the residual variance

for a trait in aerated pond (non-aerated pond).

The r_g between the same traits were not estimable as the bivariate model did not converge when the whole data set ($N = 2063$) was used. Alternatively, a subset of data based on clustering fish using genomic relationships was used and analysed with GREML (Chu et al. (2019)).

Individuals that were poorly linked to others were removed because this could result in fewer fish with a better genetic connectedness and alleviate the model convergence problem. STRUCTURE software version 2.3.4 (Pritchard et al., 2000; Falush et al., 2003; Falush et al., 2007; Hubisz et al., 2009) was used to cluster fish in each hapa into 10 groups, creating a total of 40 groups with closer than average relationship. After the probability that each fish belonged to each group (prob.) was obtained, four probability thresholds (≥ 0.5 , ≥ 0.6 , ≥ 0.65 and ≥ 0.70) were used to exclude less related individuals from each group. The individuals from each hapa that passed the screening were merged which resulted in 1309, 1012, 827 and 802 fish with a probability of ≥ 0.5 , ≥ 0.6 , ≥ 0.65 and ≥ 0.7 , respectively. Bivariate analyses using the 1309, 1012, 827 and 802 fish data sets were undertaken. Only the data set with 802 fish that passed the ≥ 0.7 threshold probability converged and these parameters are reported here.

3. Results

3.1. Descriptive statistics

Dissolved oxygen and temperature were recorded daily. The average morning DO and temperature in the aerated pond were 6.0 mg/l and 28.6 °C, respectively. The average morning DO and temperature levels in the non-aerated pond were 0.9 mg/l and 27.3 °C, respectively. The morning DO in the non-aerated pond decreased over time as fish got bigger and was much lower than the minimum requirement of 3.0 mg/l (Fig. 1). The average dissolved oxygen around 1:00 pm and 6:00 pm were both above the 5 mg/l in both ponds. The average un-ionized ammonia (UIA) was 0.03 mg/l in both ponds. The average pH was 7.4 and 7.3 in the aerated and the non-aerated pond, respectively.

The number of fish harvested from the aerated pond and non-aerated pond were 1005 and 899, respectively. Survival in the aerated and non-aerated pond was 64.0% and 57.2%, respectively. Survival was significantly ($p < .001$) higher in the aerated pond than the non-aerated pond. Feed conversion ratio (FCR) was 1.73 and 2.31 in the aerated and non-aerated pond, respectively (Table 2).

Descriptive statistics of harvest weight, for each sex and combined, are presented in Table 3 and in the Supplementary Table S1. The mean weights in the four hapas were different, but the average stocking weights in the aerated (25.4 g) and non-aerated (24.8 g) ponds, were similar. However, the coefficient of variation was somewhat higher in the non-aerated pond (53.9%) compared to the aerated pond (51.8%), due to random sampling effects (Table 3). The mean harvest weight was 781.4 g for the aerated pond and 578.5 g for the non-aerated pond. Males were 36.6% and 26.6% heavier than females at harvest (Table 3) in the aerated and non-aerated pond respectively. The coefficient of variation (CV) for harvest weight for females was higher in both aerated (31.1%) and non-aerated ponds (25.0%) than for males in both aerated (22.3%) and non-aerated ponds (19.4%). Survival, expressed as days to (assumed) mortality, was higher in the aerated pond (199.9 +/− 47.6) than in the non-aerated pond (190.2 +/− 53.5; Table 3). There was no

difference in shape (Ec) between the two ponds (mean value 0.4). In summary, pond aeration leads to a higher mean harvest weight, higher survival and better FCR compared to non-aeration.

3.2. Phenotypic and genetic parameters estimation within ponds

Estimates of variance components and heritability (h^2) from univariate models for HW, SL, H, SA, Ec, TGC and survival days in the aerated and non-aerated ponds are presented in Table 5. For all traits, variance estimates were lower in the non-aerated pond. The genetic coefficients of variation in the non-aerated pond were 9.7 to 47.2% lower compared to the aerated pond (Table 5). The h^2 estimates for HW, H, SA, Ec and TGC in the aerated and non-aerated ponds were moderate to high, ranging from 0.14 to 0.45 with small standard errors (0.05 to 0.07). The common environmental effects to full sibs, which were in our case fixed rearing hapa effects, were not significant in both ponds. All h^2 were higher by 4.8 to 65.2% in the aerated pond compared to the non-aerated pond for all the traits, except Ec. The h^2 estimates for survival days were low and with large standard errors, 0.04 ± 0.03 and 0.03 ± 0.02 in the aerated and non-aerated ponds, respectively.

Phenotypic and genetic correlations are presented in Table 6 (aerated pond) and Table 7 (non-aerated pond). As expected, the phenotypic (r_p) and genetic correlations (r_g) between HW and body size traits (SL, H, SA) were high ($r_p = 0.91$ to > 0.99 , $r_g = 0.85$ to > 0.99). The genetic correlation between HW and SA was close to unity in both ponds ($r_g = 0.99$ for the non-aerated pond and $r_g > 0.99$ for the aerated pond). The estimated r_g between TGC and HW or SA was 0.97 in the aerated pond, and 0.84–0.86 in the non-aerated pond, indicating that SA and HW describe genetically the same trait.

The estimated r_g between Ec and HW were -0.25 ± 0.17 and -0.48 ± 0.17 in the aerated and the non-aerating ponds, respectively. A negative value means that genetically larger fish are rounder (ellipse value closer to zero) than smaller fish. The estimated r_g between Ec and TGC in both ponds were also low and negative (-0.26 to -0.38), indicating that fish that grow fast are more round.

The estimated r_g between survival days and traits such as HW, body size traits (SL, H and SA) and TGC in the aerated pond were low and negative with large standard errors (-0.00 to -0.29 , Table 6). Taking into account the large standard errors, these genetic correlations suggest that fish that have genetically higher HW, body size and TGC had less survival days than fish with lower HW, body size and TGC. Similarly, the estimated r_g between survival days and SL (-0.50 ± 0.24) and survival and H (-0.14 ± 0.37) in the non-aerated pond were low and negative with large standard errors (Table 7), suggesting that survival days reduces genetically with increasing HW and body size. Longer fish has less survival days in the non-aerated pond than the aerated pond. Genetic correlations between survival and Ec in both ponds and between survival and traits such as HW, Ec and TGC in the non-aerated pond could not be estimated due to model convergence problems. Phenotypic correlations between survival and other traits at harvest were not estimated because all fish that survived until harvest had the same survival days. Because of lack of variation in survival days in the survived fish, there were no residual correlations. Although the standard errors were large, the negative r_g between survival days and HW, body size traits and TGC suggest that heavier, large and fast-growing fish had a lower survival rate and therefore fewer survival days, especially in the non-aerated pond.

3.3. GxE between ponds

Genetic correlations of HW, TGC, survival, and body size traits between the aerated and the non-aerated ponds are given in Table 8. The bivariate model did not converge when the full dataset was used ($N = 2063$). After clustering and removing all fish with low relationships (giving $N = 802$) most genetic correlations could be estimated with a bivariate model (Table 8). Correlations were generally high, with

Table 2

Total number of stocked and harvested fish, survival percentage and feed conversion ratio (FCR) in aerated and non-aerated pond.

Pond	Number of fish stocked	Number of fish harvested	Survival (%)	FCR
Aerated	1570	1005	64.0	1.73
Non-aerated	1572	899	57.2	2.31

Table 3

Number of fish (N), mean body weights at stocking and harvest (g), mean survival days, coefficient of variation (CV, %) and standard deviation (SD) of genotyped fish per pond.

Environment	Sex	At stocking				At harvest				Survival days			
		N	Mean body weight	SD	CV	N	Mean body weight	SD	CV	N	Mean survival days	SD	CV
Aerated pond	Male	484	26.6	14.2	53.1	451	952.1	211.9	22.3	484	212.1	23.2	10.9
	Female	459	24.1	11.8	48.9	434	603.3	187.6	31.1	459	212.6	23.12	10.9
	Not determined	83	26.1	13.9	53.3					83	58		
	All sex	1026	25.4	13.2	51.7	885	781.4	265.5	34.0	1026	199.9	47.6	23.8
Non-aerated pond	Male	440	24.9	14.0	56.2	388	669.9	129.6	19.4	440	209.1	20.7	9.9
	Female	468	24.7	13.1	53.0	413	492.6	123.3	25.0	468	208.8	20.7	9.9
	Not determined	129	25.1	12.5	50.0					129	58		
	All sex	1037	24.8	13.4	543.0	801	578.5	154.34	26.7	1037	190.2	53.5	28.1

Table 5

Additive genetic (σ_a^2) and residual (σ_e^2) variances, genetic coefficient of variation (GCV) and heritability (h^2) and its standard error (se) estimates from univariate models for harvest weight (HW), standard length (SL), body height (H), surface area (SA), ellipticity (Ec), thermal growth coefficient (TGC) and survival days in aerated (A) and non-aerated (NA) ponds.

Trait	Pond	σ_a^2	σ_e^2	GCV	h^2 (se)
HW	A	8435.72	28,809.40	0.118	0.23 (0.06)
	NA	2728.57	12,409.00	0.090	0.18 (0.06)
SL	A	0.91	2.97	0.037	0.23 (0.06)
	NA	0.22	2.43	0.019	0.08 (0.05)
H	A	0.25	0.92	0.045	0.21 (0.06)
	NA	0.14	0.56	0.036	0.20 (0.06)
SA	A	288.02	1157.20	0.072	0.20 (0.06)
	NA	106.28	672.10	0.052	0.14 (0.06)
Ec	A	0.00014	0.00024	0.030	0.37 (0.07)
	NA	0.00012	0.00015	0.027	0.45 (0.07)
TGC	A	0.0042	0.0121	0.064	0.26 (0.06)
	NA	0.0019	0.0073	0.051	0.21 (0.06)
Survival days	A	21.93	472.38	0.023	0.04 (0.03)
	NA	12.39	376.34	0.019	0.03 (0.02)

high standard errors. The correlations for survival and Ec were not estimable as the model did not converge. The genetic correlation between ponds for HW, SL, H, SA and TGC were 0.81 ± 0.30 , 0.80 ± 0.27 , 0.74 ± 0.33 , 0.78 ± 0.34 and 0.78 ± 0.22 , respectively, indicating some degree of GxE interaction.

4. Discussion

The objectives of this study were to investigate the presence of GxE between aerated (A) and non-aerated (NA) earthen ponds and the impact of (non)-aeration on genetic parameters by designing a GxE experiment that could minimize common environmental effects. GxE may manifest itself as heterogeneity of variances and re-ranking. In the next three sections, the novel aspects of the experimental design, the impact of (non)-aeration and GxE and the implications of our study for genetic

improvement programs are discussed.

4.1. The experimental approach

Novelties in this experiment were mass production of families to minimize common environmental effects, use of genomic relationships based on GBS and use of digital image analysis (DIA) to measure standard length and height. A common problem in GxE studies in tilapia is the estimation of common environmental effects. In our experiment based on mass spawning and genomic relationships there were very small and not significant common environmental effects, which were in this case the rearing hapa effects. The experimental approach is therefore a solution to minimize common environmental effects and has clearly advantages in addition to minimizing common environmental effects. In the experiment families were produced in only 15 days using mass spawning which is much shorter than the 2–3 months required by classical family production (Trøng et al., 2013a). Another benefit of mass spawning was that it required less labour and infrastructure. The main disadvantage of mass spawning was that it was not possible to have control over the mating and the number of families produced. A small number of sires may have contributed to a large portion of the offspring (Fessehaye et al., 2006). However, the use of genomic information allows the estimation of relationships between and within families, which creates additional power for estimation of the genetic parameters (Visscher et al., 2014).

In this study, the family production and family group communal nursery time (70 days) was much shorter than the grow-out length (217 to 218 days). In earlier studies that investigated GxE interactions (Eknath et al., 2007; Khaw et al., 2009b; Trøng et al., 2013a; Omasaki et al., 2016), genetic correlations have been estimated using pedigree relationships on animals that experienced a prolonged common environment prior to testing. In these studies, the period of grow-out in production environment was less than the period of time in which families were reared separately in circumstances leading to communal environmental effects. Shorter grow-out periods relative to those for family production could lead to higher genetic correlations due to full

Table 6

Genetic correlations above the diagonal, phenotypic correlations below the diagonal, heritabilities on the diagonal and standard errors in brackets for harvest weight (HW), standard length (SL), height (H), surface area (SA), ellipticity (Ec) and days of survival of fish in aerated pond.

	HW	SL	H	SA	Ec	TGC	Survival
HW		0.95 (0.02)	0.95 (0.02)	. > 0.99	-0.25 (0.17)	0.97 (0.01)	-0.07 (0.31)
SL	0.92 (0.01)		0.79 (0.07)	0.94 (0.02)	0.07 (0.18)	0.92 (0.03)	-0.00 (0.32)
H	0.93 (0.01)	0.88 (0.01)		0.95 (0.02)	-0.54 (0.13)	0.93 (0.03)	-0.29 (0.35)
SA	> 0.99	0.96 (0.00)	0.97 (0.00)		-0.26 (0.17)	0.97 (0.01)	-0.15 (0.33)
Ec	-0.40 (0.04)	-0.17 (0.04)	-0.60 (0.03)	-0.42 (0.03)		-0.26 (0.16)	^a
TGC	0.97 (0.00)	0.93 (0.01)	0.93 (0.01)	0.96 (0.00)	-0.41 (0.04)		-0.16 (0.29)
Survival	^b	^b	^b	^b	^b	^b	

^a The parameters could not be estimated due to model convergence problem.

^b Phenotypic correlations could not be estimated due to no residual variance. All fish with harvest weight measurement had the same survival days.

Table 7

Genetic correlations above the diagonal and phenotypic correlations below the diagonal and standard errors in brackets for harvest weight (HW), standard length (SL), height (H), surface area (SA), ellipticity (Ec) and days of survival of fish in non-aerated pond.

	HW	SL	H	SA	Ec	TGC	Survival
HW		0.85 (0.08)	0.92 (0.03)	0.99 (0.01)	-0.48 (0.16)	0.86 (0.05)	^a
SL	0.91 (0.01)		0.61 (0.17)	0.86 (0.07)	0.04 (0.26)	0.73 (0.13)	-0.50 (0.24)
H	0.94 (0.00)	0.86 (0.01)		0.93 (0.03)	-0.79 (0.10)	0.78 (0.08)	-0.14 (0.37)
SA	0.97 (0.00)	0.94 (0.00)	0.97 (0.00)		-0.52 (0.17)	0.84 (0.07)	^a
Ec	-0.38 (0.04)	-0.11 (0.04)	-0.58 (0.03)	-0.38 (0.04)		-0.38 (0.16)	^a
TGC	0.95 (0.00)	0.87 (0.01)	0.91 (0.01)	0.93 (0.01)	-0.38 (0.04)		^a
Survival	^b	^b	^b	^b	^b	^b	

^a The parameters could not be estimated due to model convergence problem.

^b Phenotypic correlations could not be estimated due to no residual variance. All fish with harvest weight measurement had the same survival days.

Table 8

Genetic correlation between the aerated and the non-aerated ponds for harvest weight (HW), standard length (SL), height (H), surface area (SA), ellipticity (Ec), thermal growth coefficient (TGC) and survival days.

Trait	Genetic correlation
HW	0.81 ± 0.30
SL	0.80 ± 0.27
H	0.74 ± 0.33
SA	0.78 ± 0.34
Ec	^a
TGC	0.78 ± 0.22
Survival days	^a

^a The genetic correlations were not estimable.

sibs spending more time together than in different environments, which may mask GxE (Dupont-Nivet et al., 2010). Our estimates of genetic correlations were, however, not very different from previous studies estimating GxE between various environments.

Estimation of genetic correlations between the two ponds using a GRM based on SNPs from GBS, was not trivial, because ASReml did not converge when the whole data set ($N = 2063$) was used. Three issues may have played a role in this convergence problem: 1. limited sample size, 2. missingness and genotyping errors in GBS data and 3. mass spawning and unbalanced family size. The whole data set with approximately 1000 animals per environment is on the low side for estimating GxE (Sae-Lim et al., 2010; Lozano-Jaramillo et al., 2019). However, genotypic information can largely decrease the standard error on estimated genetic correlations. For instance, when using the equations by Visscher et al. (2014), the standard error on the genetic correlation becomes ~0.1 when assuming that the number of independent chromosomal segments is 500, heritability is 0.3 and the true genetic correlation is 0.8. This suggests that the convergence problem may not be just low sample size. Secondly, missingness and genotyping errors in GBS data could play a role. The coverage of the GBS data may have yielded a low number of informative SNP genotypes in pairs of individuals. For instance, in the extreme case, when using a threshold of 30% of missing SNP per individual, a pair of two individuals each with 30% missing SNP genotypes and with no overlap in the missing SNPs, could have only 40% of the SNPs used to estimate the genetic relationship between these two individuals. This in itself may create extra noise in the genomic relationships between individuals, because for each pair of individuals different SNPs were used to estimate the genomic relationship. Another sign of the limited quality of the GBS data is the fact that the average of the diagonal elements in the G-matrix was below 1, indicating an excessive amount of heterozygosity across loci which is likely due to poor SNP calling on some loci. Similarly, Pérez-Enciso (2014) and Dodds et al. (2015) reported in simulated and real data distortion of elements of the G-matrix based on low-coverage GBS data. Thirdly, mass spawning may lead to large differences in family size. Fessehaye et al. (2006) found that some

families were present in small quantities while other families were very abundant. In the study, the parents were not genotyped and parentage assignment was not possible. Therefore, families could not be equalized in terms of numbers of individuals. However, by clustering animals in groups based on high molecular coancestry, individuals were identified that were poorly related with all other individuals based on genomic relationships. This indeed resulted in an analysis that converged. The genetic correlations in the present study were estimated based on 802 fish which is on the lower side of what was considered an optimal design for estimating genetic correlation using stochastic simulation (Lozano-Jaramillo et al., 2019). On the other hand, Ødegård and Meuwissen (2012) showed with simulation that with a low number of families and large family size heritabilities can already be estimated quite accurately just based on within-family relationships using genomics, which was supported by analytical work by Hill (2013). For new experiments, precision of estimated genetic correlations can be improved by genotyping parents to construct the pedigree and remove small families, using a SNP array data instead of GBS and increasing the size of the experiment.

Another novelty in this experiment was the use of DIA to calculate SA and Ec in Nile tilapia. The genetic correlation between HW and SA, and TGC and SA were strong and positive, while the genetic correlations between HW and Ec were negative in both ponds and only significantly deviating from zero in the non-aerated pond (*t*-test) (Lynch and Walsh, 1998). Therefore, HW, TGC and Ec could be improved by selecting on the correlated SA, which can be automated using DIA and reduces handling stress. Another advantage of automated DIA is that it allows for multiple measurements to be taken over time so that the moment of mortality can be recorded with much greater precision than based on only 3 interval measurements and the final harvest as in the current experiment. The use of DIA is time-efficient and can be stored for later use (Blonk et al., 2010), moreover, it is less stressful for the fish than the manual method and avoids recording errors.

4.2. The impact of (non)aeration and genotype by environment interactions

Aerating the pond had a positive impact on harvest weight, survival and FCR. The mean harvest body weight was 26% higher in the aerated pond than in the non-aerated pond. Similarly, survival of fish was higher in the aerated pond. The FCR in the aerated pond was 1.73 which is high but an acceptable value, while in the non-aerated pond FCR was too high at 2.31 (Craig, 2009). Aerating ponds kept dissolved oxygen level always above 5 mg/l, while in the non-aerated pond the dissolved oxygen dropped to < 1 mg/l during the night. Dissolved oxygen is one of the main factors that affect FCR and growth of Nile tilapia (Mengistu et al., 2019). Under hypoxia (3 mg/l), Nile tilapia significantly underperform in terms of FCR and growth compared to normoxia (5 mg/l) (Tran et al., 2016; Mengistu et al., 2019). In summary, our results confirm the findings of Mengistu et al. (2019) that aerating ponds results in a higher mean harvest weight, survival, and lower FCR.

In the presence of large environmental differences between selection and production environments, GxE is expected. In the present experiment there were heterogeneity of variances between the aerated and non-aerated ponds and genetic correlations less than unity between ponds. The additive genetic variances and heritabilities for the different traits were lower in the non-aerated pond. For instance, the genetic coefficient of variation for HW, TGC and survival in the non-aerated pond was lower by 23.7%, 20.3% and 40.9%, respectively. Eknath et al. (2007) found 46 to 79% lower heritabilities for harvest weight in low input environments compared to high input environment. Non-aeration inhibited the fish from expressing their full genetic potential for growth. This inhibition of growth resulted also in re-ranking of fish as indicated by the genetic correlations around 0.8 in this study, although the genetic correlation did not significantly deviate from 1 (*t*-test) (Lynch and Walsh, 1998). To the best of our knowledge, this is the first study publishing genetic correlations between aerated and non-aerated ponds for the traits investigated. However, there have been GxE studies using pedigree relationships for HW of Nile tilapia between different environments. Our estimate of 0.81 ± 0.30 was similar to those estimates of GxE in other studies: 0.76–0.99 between different pond environments (Eknath et al., 2007), 0.86–0.94 between nucleus, cage and low input pond (Trong et al., 2013a), 0.74 between mixed sex and mono sex Nile tilapia (Omasaki et al., 2016), and 0.74 between low input and high input pond environments (Khaw et al., 2009a). Our result for TGC 0.78 ± 0.22 was similar with 0.77 between the nucleus breeding environment and low input environment for daily growth coefficient (DGC) reported by Trong et al. (2013a) and higher than the 0.59 between mono-sex and mixed sex Nile tilapia for DGC reported by Omasaki et al. (2016). In summary, our study reports for the first time GxE between aerated and non-aerated ponds that results in both heterogeneity of variance and heritability as well as re-ranking of genotypes.

4.3. Implications for genetic improvement programs

From a genetic improvement perspective, the question is how to select fish in aerated ponds that perform better in both aerated and non-aerated ponds. In many pond production environments, farmers usually do not aerate, while the selection environment usually consists of aerated ponds. Therefore, the improvement in performance in a non-aerated pond is a correlated response. A correlated response is less than a direct response when a genetic correlation is less than one, assuming heritabilities are similar in the two environments (Falconer, 1990). Selecting under environments that are not similar with a production environment limits scope of selection of alleles of genes that are responsible for better performance in a production environment (Hammond, 1947). With estimated genetic correlations of about 0.8 in this study and assuming that the true genetic correlation is close to the estimated value, it is clear that the genetic improvement in the nucleus based on aerated ponds will not be fully expressed in production environments without aeration. The advantage of selection in aerated ponds is that the heritability of most traits is higher, resulting in a higher accuracy of selection than when selection would be performed in non-aerated ponds. Nevertheless, if the breeding goal is to increase performance in non-aerated production environments and selection has to be undertaken in an aerated pond, it is advised to use half-sib information from on-station non-aerated ponds (Brascamp et al., 1985; Mulder and Bijma, 2005). This half sib information reduces the reduction in selection response due to GxE (Brascamp et al., 1985; Mulder and Bijma, 2005). However, half sib information requires pedigree data and pedigree data is often lacking in non-aerated ponds in farms. In the experiment, the common environmental effect was successfully reduced by using natural mating. Genotyping with GBS was however less reliable, probably due to a high rate of genotyping errors. We recommend therefore to collect genomic and phenotypic information from production environments and genotyping with a SNP-chip, to set up

reference populations for genomic selection programs to increase response in commercial environments (Mulder, 2016).

5. Conclusions

Substantial additive genetic variance was found for HW, TGC, survival, body shape and body size measurements in aerated and the non-aerated ponds indicating these traits respond to selective breeding. Mass spawning and use of genomic relationship enabled to minimize common environmental effects to full sibs. Non-aeration led to lower genetic variance and heritabilities. The estimated genetic correlations suggest some GxE for HW, standard length, height, surface area and TGC, although none of the genetic correlations was significantly deviating from unity due to large standard errors. To optimize breeding programs to breed fish that perform well in non-aerated ponds, breeding programs are recommended that use genotypic and phenotypic information from non-aerated on-station ponds to set up a reference population for genomic selection.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.aquaculture.2020.735704>.

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.

Acknowledgements

This publication was made possible through support provided principally by the Koepon Foundation, by the International Fund for Agricultural Development (IFAD), the CGIAR Research Program on Fish Agri-Food Systems (FISH) led by WorldFish. The program is supported by contributors to the CGIAR Trust Fund. We thank the WorldFish breeding team, and the Department of Fisheries Malaysia at Jitra for their technical support throughout the project.

References

- Adams, D.C., Collyer, M.L., Kaliotzopoulou, A., 2018. Geomorph: Software for geometric morphometric analyses. In: R Package Version 3.0.6.
- Bentsen, H.B., Gjerde, B., Nguyen, N.H., Rye, M., Ponsoni, R.W., Palada de Vera, M.S., 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* 338–341, 56–65.
- 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 (2010), 1–2.
- Brascamp, E.W., Merks, J.W.M., Wilmink, J.B.M., 1985. Genotype environment interaction in pig breeding programmes: methods of estimation and relevance of the estimates. *Livest. Prod. Sci.* 13, 135–146.
- Calus, M.P.L., Vandenplas, J., 2016. Calc_grm - a Program to Compute Pedigree, Genomic and Combined Relationship Matrices. ABGC, Wageningen UR Livestock Research.
- Chu, T.T., Bastiaansen, J.W.M., Berg, P., Komen, H., 2019. Optimized grouping to increase accuracy of prediction of breeding values based on group records in genomic selection breeding programs. *Genet. Sel. Evol.* 51, 1–12.
- Conte, M.A., Gammerdinger, W.J., Bartie, K.L., Penman, D.J., Kocher, T.D., 2017. A high quality assembly of the Nile Tilapia (*Oreochromis niloticus*) genome reveals the structure of two sex determination regions. *BMC Genomics* 18, 341.
- Craig, S., 2009. Understanding Fish Nutrition, Feeds, and Feeding (Publication 420–256). Virginia Polytechnic Institute and State University, pp. 4.
- Dodds, K.G., McEwan, J.C., Brauning, R., Anderson, R.M., van Stijn, T.C., Kristjansson, T., Clarke, S.M., 2015. Construction of relatedness matrices using genotyping-by-sequencing data. *BMC Genomics* 16, 1047.
- Dupont-Nivet, M., Karahan-Nomm, B., Vergnet, A., Merdy, O., Haffray, P., Chavanne, H., Chatain, B., Vandepitte, M., 2010. Genotype by environment interactions for growth in European seabass (*Dicentrarchus labrax*) are large when growth rate rather than weight is considered. *Aquaculture* 306, 365–368.
- Eknath, A.E., Bentsen, H.B., Ponsoni, R.W., Rye, M., Nguyen, N.H., Thodesen, J., Gjerde, B., 2007. Genetic improvement of farmed tilapias: composition and genetic parameters of a synthetic base population of *Oreochromis niloticus* for selective breeding.

Aquaculture 273, 1–14.

Ellen, E.D., Visscher, J., van Arendonk, J.A.M., Bijma, P., 2008. Survival of laying hens: genetic parameters for direct and associative effects in three purebred layer lines. *Poult. Sci.* 87, 233–239.

Falconer, D.S., 1952. The problem of environment and selection. *Am. Nat.* 86, 293–298.

Falconer, D.S., 1990. Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. *Genet. Res.* 56, 14.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164, 1567–1587.

Falush, D., Stephens, M., Pritchard, J.K., 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* 7, 574–578.

Fessehaye, Y., El-bialy, Z., Rezk, M.A., Crooijmans, R., Bovenhuis, H., Komen, H., 2006. Mating systems and male reproductive success in Nile tilapia (*Oreochromis niloticus*) in breeding hapas: a microsatellite analysis. *Aquaculture* 256, 148–158.

Garrison, E., Marth, G., 2012. Haplotype-Based Variant Detection from Short-Read Sequencing.

Gilmour, A.R., Gogel, G.J., Cullis, B.R., Welham, S.J., Thompson, R., 2015. ASReml User Guide Release 4.1 Functional Specification.

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* 342–343, 117–124.

Goddard, M.E., Hayes, B.J., Meuwissen, T.H.E., 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. *J. Anim. Breed. Genet.* 128, 409–421.

Hammond, J., 1947. Animal breeding in relation to nutrition and environmental conditions. *Biol. Rev.* 22, 195–213.

Hill, W.G., 2013. On estimation of genetic variance within families using genome-wide identity-by-descent sharing. *Genet. Select. Evol.* 45, 32.

Hubisz, M.J., Falush, D., Stephens, M., Pritchard, J.K., 2009. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* 9, 1322–1332.

Jobling, M., 2003. The thermal growth coefficient (TGC) model of fish growth: a cautionary note. *Aquac. Res.* 34, 581–584.

Khaw, H.L., Bovenhuis, H., Ponzoni, R.W., Rezk, M.A., Charo, H., Komen, J., 2009a. Genetic analysis of nile tilapia (*Oreochromis niloticus*) selection line reared in two input environments. *Aquaculture* 294 (2009), 1–2.

Khaw, H.L., Ponzoni, R.W., Hamzah, A., Kamaruzzaman, N., 2009b. Genotype by environmental interaction for live weight between two production environments in the GIFT strain (Nile tilapia, *Oreochromis niloticus*). Matching genetics and environment: a new look at an old topic. In: Proceedings of the 18th Conference of the Association for the Advancement of Animal Breeding and Genetics, Barossa Valley, South Australia, Australia, 28 September–1 October, 2009. Association for the Advancement of Animal Breeding and Genetics Armidale.

Khaw, H.L., Ponzoni, R.W., Hamzah, A., Abu-Bakar, K.R., Bijma, P., 2012. Genotype by production environment interaction in the GIFT strain of Nile tilapia (*Oreochromis niloticus*). *Aquaculture* 326–329, 53–60.

Komen, H., Trong, T.Q., 2014. Nile tilapia genetic improvement: achievements and future directions. In: Proc. The 10th International Symposium on Tilapia in Aquaculture - ISTA10, Jerusalem, Israel.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079.

Lozano-Jaramillo, M., Komen, J., Bastiaansen, J.W.M., Dessie, T., 2019. Predicting Breed by Environment Interaction Using Ecological Modelling. Wageningen University, Wageningen.

Lynch, M., Walsh, B., 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts, US.

Mengistu, S.B., Mulder, H.A., Benzie, J.A.H., Komen, H., 2019. 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.* 11.

Mulder, H.A., 2016. Genomic selection improves response to selection in resilience by exploiting genotype by environment interactions. *Front. Genet.* 7, 178.

Mulder, H.A., Bijma, P., 2005. Effects of genotype \times environment interaction on genetic gain in breeding programs. *J. Anim. Sci.* 83, 49–61.

Ødegård, J., Meuwissen, T.H.E., 2012. Estimation of heritability from limited family data using genome-wide identity-by-descent sharing. *Genet. Sel. Evol.* 44, 1–10.

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.

Pérez-Enciso, M., 2014. Genomic relationships computed from either next-generation sequence or array SNP data. *J. Anim. Breed. Genet.* 131, 85–96.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Purcell, S., 2018. PLINK. Version 1.9.

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

Rohlf, F.J., 2017a. tpsDig2, Version 2.30. Department of Ecology and Evolution, State University of New York at Stony Brook.

Rohlf, F.J., 2017b. tpsUtil, Version 1.74. Department of Ecology and Evolution, State University of New York at Stony Brook.

Sae-Lim, P., Komen, J., Kause, A., 2010. Bias and precision of estimates of genotype-by-environment interaction: a simulation study. *Aquaculture* 310, 66–73.

Thoa, N.P., Ninh, N.H., Knibb, W., Nguyen, N.H., 2016. Does selection in a challenging environment produce Nile tilapia genotypes that can thrive in a range of production systems? *Sci. Rep.* 6, 21486.

Tran, N.T.K., Dinh, N.T., Tin, N.H., Roem, A.J., Schrama, J.W., Verreth, J.A.J., 2016. Interaction between Dissolved Oxygen Concentration and Diet Composition on Growth, Digestibility and Intestinal Health of Nile Tilapia (*Oreochromis niloticus*).

Trong, T.Q., Mulder, H.A., van Arendonk, J.A.M., Komen, H., 2013a. 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* 384–387, 119–127.

Trong, T.Q., van Arendonk, J.A.M., Komen, H., 2013b. Genetic parameters for reproductive traits in female Nile tilapia (*Oreochromis niloticus*): I. Spawning success and time to spawn. *Aquaculture* 416–417, 57–64.

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K.V., Altshuler, D., Gabriel, S., DePristo, M.A., 2013. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* 43, 11.10.11–11.10.33.

VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423.

Visscher, P.M., Hemani, G., Vinkhuyzen, A.A.E., Chen, G.-B., Lee, S.H., Wray, N.R., Goddard, M.E., Yang, J., 2014. Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *PLoS Genet.* 10, e1004269.

Wommongkol, P., Sukhavachana, S., Ampolsak, K., Srisapoom, P., Suwanasoppe, T., Poompuang, S., 2018. Genetic parameters for resistance against *Flavobacterium columnare* in Nile tilapia *Oreochromis niloticus* (Linnaeus, 1758). *J. Fish Dis.* 41, 321–328.