Genome-wide association analysis of adaptation to oxygen stress in Nile tilapia (*Oreochromis niloticus*)

Xiaofei Yu[1]*, Hendrik-Jan Megens[1], Samuel B. Mengistu[1,2], John W.M. Bastiaansen[1], Han A. Mulder[1], John A.H. Benzie[3,4], Martien A.M. Groenen[1], Hans Komen[1]

1. Wageningen University & Research Animal Breeding and Genomics, The Netherlands

2. School of Animal and Range Sciences, College of Agriculture, Hawassa University, Hawassa Ethiopia

3. WorldFish Centre, Jalan Batu Maung, Bayan Lepas, Penang, Malaysia

4. School of Biological Earth and Environmental Sciences, University College Cork, Cork, Ireland

* Corresponding authors. E-mail: xiaofei.yu@wur.nl.

Abstract

Background: Tilapia is one of the most abundant species in aquaculture. Hypoxia is known to depress growth rate, but the genetic mechanism by which this occurs is unknown. In this study, two groups consisting of 3140 fish that were raised in either aerated (normoxia) or non-aerated pond (nocturnal hypoxia). During grow out, fish were sampled five times to determine individual body weight (BW) gains. We applied a genome-wide association study to identify SNPs and genes associated with the hypoxic and normoxic environments in the 17th generation of a Genetically Improved Farmed tilapia population.

Results: In the hypoxic environment, 36 SNPs associated with at least one of the five body weight measurements (BW1 till BW5), of which six, located between 19.48 Mb and 21.04 Mb on Linkage group (LG) 8, were significant for body weight in the early growth stage (BW1 to BW2). Further significant associations were found for BW in the later growth stage (BW3 to BW5), located on LG1 and LG8. Analysis of genes within the candidate genomic region suggested that MAPK and VEGF signalling were significantly involved in the later growth stage under the hypoxic environment. Well-known hypoxia-regulated genes such as *igf1rb*, *rora, efna3* and *aurk* were also associated with growth in the later stage in the hypoxic environment. Conversely, 13 linkage groups containing 29 unique significant and suggestive SNPs were found across the whole growth period under the normoxic environment. A meta-analysis showed that 33 SNPs were significantly associated with BW across the two environments, indicating a shared effect independent of hypoxic or normoxic environment. Functional pathways involved in nervous system development and organ growth in the early stage, and oocyte maturation in the later stage.

Conclusions: There are clear genotype-growth associations in both normoxic and hypoxic environments, although genome architecture involved changed over the growing period, indicating a transition in metabolism along the way. The involvement of pathways important

in hypoxia especially at the later growth stage indicates a genotype-by-environment interaction, in which MAPK and VEGF signalling are important components.

Keywords: Nile tilapia, growth, hypoxia, oxygen stress, GWAS, meta-analysis

Background

Tilapia is one of the most important species in aquaculture noted for their relative ease of culture and rapid growth. Tilapia is currently cultured in over 120 countries, mainly in the tropics and sub-tropics, with a production from 0.3 million tonnes in 1987 to closely 7 million tonnes in 2018, which makes it the second largest aquaculture species in the world [1]. Tilapia is a valuable protein source in developing and emerging economies. Due to its wide range of culturing conditions, tilapia is also an excellent model to study adaptive responses to environmental stresses [2]. One of the most important non-commercial breeding programs is the GIFT (Genetic Improvement of Farmed Tilapia) project, executed by WorldFish in Malaysia. It has sustained genetic gains for growth and body trait more than 10% per generation for more than six generations [3]. However, rapid growth potentially exacerbates existing limitations in the production environment. In non-aerated ponds, high stocking density can lead to an extreme hypoxic environment, especially at the end of the night (nocturnal hypoxia), when algae have higher rate of oxygen consumption than oxygen production. The extreme hypoxic environment can lead to lower feed intake, stagnated growth, and susceptibility to disease [4, 5]. The result is a higher mortality and lower yield than what could potentially be achieved [6]. The effects can be mitigated through mechanical aeration of ponds, but a daily fluctuation in oxygen availability is nevertheless inevitable.

Response to hypoxia is a highly complicated biological process that has received considerable scientific attention, both in fishes and in land vertebrates (e.g. high-altitude adaptation studies). Most of these response processes happen very early at the onset of hypoxia through the activation of pathways depending on proteins that are already present [7]. But in the longer term, adaptive responses to hypoxia are leading to different expression of genes. In mammals, studies in the past decades pointed to a central role of the hypoxia-Inducible factors (HIF) in the regulation of gene expression during hypoxia [8]. Other genes such as tyrosine hydroxylase (TH), phosphoglycerate kinase 1 (PGK1) and vascular endothelial growth factor (VEGF) are also important key actors [9]. Recent studies have described that fish have homologs of HIF-α and -β, which may show similar function to those in mammals in the hypoxic environment [9, 10]. Several other hypoxia-related proteins and signal pathways have been reported, such as AMP-activated protein kinase (AMPK), reactive oxygen species (ROS), mitogen-activated protein kinase (MAPK) and IGF-1/PI3K/AKT signalling, which have been reported to function in hypoxia adaptation of some fish species [11, 12].

Genetic adaptation to hypoxia is important for survival in many aquatic species, since variation in oxygen availability in water can vary far more, and far more rapidly, than in terrestrial ecosystems. Hypoxia is an important cause of economic losses in aquaculture. Understanding the genomic architecture of hypoxia adaptation could help to improve resilience through breeding programs for economically important species. So far, hypoxia tolerance has been studied in a limited number of fish species, including catfish [13, 14], Atlantic salmon [15], and tilapia [16], with the aim to identify QTLs for hypoxia-tolerant traits. Genome-wide association study (GWAS) has been regarded as a powerful tool to identify genetic markers associated with target traits, and a more complete gene network will provide the knowledge bases required for the aquaculture industry to make improvements

[17]. In hybrid catfish, Zhong et al. [18] revealed in total nine SNPs associated with dissolved oxygen (DO) level using a 250K SNP array. Analysis of the genes overlapping or close to those SNPs suggested that many of those genes were involved in the PI3K/AKT and VEGF pathways. In another study, Brennan et al. [19] aimed to identify population differences in hypoxia tolerance by calculating the amount of time for Killifish to lose equilibrium using GWAS. They found that variation in Hyaluronan synthase 1 (*has1*) influenced the production of hyaluronan, which can lead to a direct effect on hypoxia tolerance.

There are only a few studies that focused on genetic bases of either hypoxia tolerance or growth in Nile tilapia [16, 20], however, none of these investigated how hypoxia influences growth in Nile tilapia. The main objective of this study was to unravel the genomic architecture associated with phenotypic variation during adaptation to hypoxia or normoxia, and to elucidate the effect of hypoxia on the genetic regulation of growth.

RESULTS

Phenotype Statistics

Fish fry was produced from generation 16 of the GIFT breeding program. The experiment was carried out in an aerated (normoxic) and non-aerated (nocturnal hypoxic) ponds, each producing 1026 and 1037 fish that were involved in the analysis. Body weight of growing fish was measured at five time points (**Table 1**). The data show that the number of tilapia in both environments gradually decreased. This effect was more pronounced in the hypoxic environment, with a total loss from stocking to harvest period of 23% of the initial number of individuals, compared to 14% in the normoxic environment. The average body weight at five time points in the normoxic environment was significantly higher than those in the hypoxic

environment, with the exception of the first time point (BW1). Interestingly, the variation in body weight (CV) at each time point in the two separate environments decreased.

The estimated phenotypic correlations for body weight between the different time points in the two environments are shown in **Table 2**. Results show that phenotypic correlation between time points in the hypoxic and nomorxic environments was initially high (0.80 and 0.81 separately), but decreased with increasing time between measurements.

Table 1 Summary statistics of body weight across the whole growth period in Nile tilapia

Trait	Days	Environments	No.	Mean	Max	Min	SD	CV(%)	P value
BW1	0	Hypoxia	1037	24.8	77.0	3.6	13.4	54.0	0.14
	0	Normoxia	1026	25.4	77.1	2.9	13.1	51.7	
BW2	55	Hypoxia	1037	144.3	328.0	26.0	54.7	37.9	3.81E-07
	56	Normoxia	1026	159.1	394.3	30.2	63.1	39.7	
BW3	104	Hypoxia	907	265.9	498.3	70.5	73.3	27.6	4.17E-08
	105	Normoxia	941	289.4	650.5	63.3	92.5	32.0	
BW4	167	Hypoxia	885	426.4	805.3	117.0	118.9	27.9	2.20E-16
	168	Normoxia	903	533.6	1079.1	68.2	177.2	33.2	
BW5	217	Hypoxia	799	579.6	1003.4	135.5	154.4	26.6	2.20E-16
	218	Normoxia	885	780.9	1588.6	185.7	265.6	34.0	

Body weight (BW), days means the time survived in either hypoxia or normoxia, the number of surviving animals (No.), maximum (Max), minimum (Min), standard deviation (SD), coefficient of variation (CV).

Table 2 Phenotypic correlations of body weight across the whole growth period in different environments

Trait	BW1	BW2	BW3	BW4	BW5
BW1	--	0.81	0.61	0.32	0.22
BW2	0.80	--	0.77	0.44	0.32
BW3	0.59	0.80	--	0.66	0.52
BW4	0.29	0.46	0.68	--	0.83
BW5	0.15	0.31	0.56	0.85	--

The spearman's rank correlation coefficient of body weight in hypoxia is presented below diagonal, while the normoxia is above diagonal.

SNP statistic and Population structure

In total 27,090 SNPs that passed SNP minor allele frequency, genotype and individuate call rate criteria, were used for subsequent analysis. Those SNPs were found to be randomly distributed across the genome with a density of approximately 28 SNP per Mb. The highest number of SNPs (4,344) on LG 3 while LG11 had the lowest number of SNPs (630) (**Figure 1A**). A few windows on LG3 show a higher density of SNPs (**Figure 1B)**. Besides this exception, the distribution of SNPs is uniform with the linkage group physical length of the *Oreochromis niloticus* genome (GCF_001858045.1).

The PCA represents the genetic structure for individuals from the hypoxic and normoxic environments, respectively (**Figures 1C**, **1D and Supplementary Figures 2, 3**). In the hypoxic environment, the first three principal components (PCs) explains 47.0% of the total genotype-based variation and separate samples according to their family differences. PC1 accounts for 15.2% of the total genotype variation and separates families in hapa3 with other families. In the normoxic environment the first three components explained 39.8% of the total genotype variation, while the first component accounts for 15.3%. Moreover, the largest PC (PC1) of all samples separates disperse cluster from families in hapa3 again.

These results indicated that there was clear genetic variation caused by family differences in both environments. This was partially caused by the different distribution of the number of fish from the four rearing hapas under the normoxic and hypoxic environments. Additionally, the average body weight of fish in hapa3 was larger than that of other hapas, especially the mean body weight of male fish at the first time point was much higher in the normoxic environment than the hypoxic environment (**Supplementary Figure 1**), indicating that a few families with high body weight dominated in one environment but not the other.

Single environmental GWAS at five different time points

Significant SNPs were detected with a univariate GWAS by applying a linear mixed model. We observed that sex and hapa effects can explain part of the difference in body weight. Thus, these were treated as fixed factors in our analysis. Overall, five association analyses, one for

each time point where body weight was measured, were performed for each environment. The Manhattan plots for each of the five time points in the hypoxic and normoxic environments are shown in **Figures 2A** and **2B** respectively. In addition, Quantile-Quantile plots with genomic inflation factors were created to aid in estimating the influence of population structure on single environmental GWAS (shown in **Supplementary Figures 3 and 4)**. The P values of corrected thresholds for suggestive and genome-wide significant levels were 4.22 ($-\log_{10}(1/16504)$) and 5.52 ($-\log_{10}(0.05/16504)$), respectively.

In the hypoxic environment, the analyses showed 10 significant and 26 suggestive SNPs associated with BW1 to BW5 (**Supplementary Table 1**). Among those, six SNPs between 19.48 Mb and 21.04 Mb on LG8 attained genome-wide significance for BW1 to BW3. However, those SNPs were not significant during BW4 and BW5. Two SNPs (LG1: 30766342 and LG1:30766336) were significant associated with BW3 to BW5. Additionally, 16 SNPs above the suggestive level as defined above for BW1 to BW2 were found on LG8, LG18 and LG19, while 18 SNPs mostly found on LG1 and LG8, were found for BW4 to BW5. Interestingly, at BW3, SNPs on LG8 overlapped with BW1 and BW2, while SNPs on LG1 overlapped with BW4 and BW5, further confirming that there is a transition in genomic architecture associated with growth over time.

We also detected 2 significant and 27 suggestive SNPs across different growth stages in the normoxic environment (**Supplementary Table 2**). The suggestive peak at BW1 covered the same genomic region as that found for the hypoxic environment between 19.48 to 21.03 Mb on LG8. However, similar to the hypoxic environment, the significance of those SNPs declined from BW1 to BW3, a pattern also seen for the SNPs located on LG18 and LG22. A few SNPs on LG7 and LG15 also showed a signal near the suggestive level from BW3 to BW5, which could be potentially interesting, although they did not attain statistical significance.

Meta-analysis GWAS across two environments

A meta-analysis GWAS that combined the effects of 27,090 SNPs in common in the hypoxic and normoxic environments was performed. The results of the meta-analysis are shown in **Figure 3**. In total 33 SNPs were found to be significantly associated with five measurements of body weight during the whole growth stage. Clusters of significant SNPs were mostly found on LG8, LG18 and LG22 (**Supplementary Table 3**). Interestingly, six SNPs located between 19.48 and 21.03 Mb on LG8, three SNPs between 12.44 and 27.32 Mb on LG18 and three SNPs within 1kb at 35.25 Mb on LG22, were all significantly associated with body weight at time points BW1 and BW2. However, the P-values of those SNPs decreased in subsequent growth periods. Five SNPs between 30.54 and 31.19 Mb on LG1, and one SNP on LG15 (LG15:23051993), were associated with body weight from BW3 to BW5. Moreover, two SNPs on LG8 (LG8:4319661, LG8: 11800435) were significant at BW4 and BW5, Notably, the SNPs located on LG8 were found at a different region compared to SNPs on the same LG in hypoxic GWAS. Hence, associations for BW1 to BW2 were different from BW4 to BW5, although intriguingly BW3 shows both overlap to early and late growth stages, which could indicate that a transition in the pathways involved occurred around this stage.

Functional annotation analysis

Based on the SNP association pattern for five measurements across the whole growth stage, we defined the early stage as BW1 and BW2, while the later stage is BW3 to BW5. Through gene identification within the associated genomic regions, the functional processes and pathways were subsequently enriched for single environmental and across environmental GWAS, respectively. Considering that BW3 is the transition point, SNPs that overlapped with the early stage were excluded in the functional annotation for the later stage. The candidate

genes derived from single environment and across environment GWAS are shown in **Figures 4A and 4B**, where 15 and 25 genes from the and BW3 to BW5 respectively, were uniquely associated with body weight in the hypoxic environment while another 12 genes were unique to growth in the normoxic environment. It is also noteworthy that three genes (*raraa*, *rarab*, *bahcc1*) were significant for BW1 and BW2 for both single and across environmental GWAS.

During the early growth stage in the hypoxic environment, fourteen GO (Gene ontology) terms were found to be significantly overrepresented (**Supplementary table 4**), including central nervous system development and steroid hormone mediated signalling pathways. Six KEGG pathways were found at later growth stage (**Figure 4C)**, including MAPK and VEGF signalling pathways. Protein interaction network analysis showed *dock5*, *dock10*, *dock11*, *baiap2a*, *baiap2b*, *aurka* and *aurkb* strongly interacting with *rac1b* and *ppp3ca*, which all are proteins participating in MAPK and VEGF signalling (**Figure 4D)**.

For the early growth stage of the normoxic environment, retinoic acid receptor signalling pathway, apoptotic signalling pathway, liver development, signal transduction, steroid hormone mediated signalling pathway and brain development biological processes (**Supplementary table 5),** were significantly enriched, while two (retinoic acid receptor and steroid hormone mediated signalling pathways) overlapped with the same growth period in the hypoxia environment. However, in contrast to the hypoxic environment, we did not find significant terms during the later growth stage in the normoxic environment.

In the meta-analysis GWAS across the normoxic and hypoxic environments, nine GO terms, including retinoic acid receptor signalling pathway and steroid hormone mediated signalling pathway, were mostly enriched in the early growth stage. During the later growth stage, two pathways involved in oocyte meiosis and progesterone-mediated oocyte maturation process. Interestingly, none of hypoxia-related pathway was enriched. (**Supplementary table 6**).

Discussion

Hypoxia is one of the most important environmental factors for fish. Hypoxia tolerance represents the ability of fish species to endure low oxygen level and to maintain a sustainable metabolic rate at lower dissolved oxygen levels [21]. Growth is a key trait for aquaculture and can be assessed by weight gain in order to examine the impact of hypoxic conditions on fish production. For more than a half century, various, often divergent, claims have been made regarding the interaction between body size and hypoxia in teleost fish. Recent studies showed that small individuals are the least hypoxia tolerant within some fish species, such as Oscar cichlid [22, 23] and Red seabream [24]. In contrast, small fish chose lower oxygen levels more than large fish in Largemouth bass [25] and Yellow perch [26], however, this behaviour was suggested to allow the smaller fish to utilize the hypoxic zone as refuge from the bigger predators [27]. From these studies it is clear that selection for low oxygen is difficult to ascertain, indicating a clear added value of investigations into genetic consequences of selection, such as the present study.

In general, metabolic rate is highly affected by dissolved oxygen in the rearing environment. Generally, faster growing animals have a higher metabolic rate and therefore require more oxygen. As a consequence, hypoxia is expected to adversely affect fish growth and feed utilization [6]. On the other hand, larger individuals have a clear advantage over smaller ones in severe hypoxic environments because small fish will use up their glycogen reserves and reach lethal levels of anaerobic end-products much faster due to their higher metabolic rate [28]. Overall fish production declines and, due to physiological stress disease resistance decreases as a consequence of hypoxia [29]. It has been observed that larger Nile tilapia tolerated low DO levels better than small ones, thought partially due to the fact that fish immunity was higher in larger than smaller Nile tilapia [30]. Regardless of the

complexity of the relationship between hypoxia and growth, studies the genomic basis of hypoxia-growth interactions in Nile tilapia are sparse.

Our results suggest a number of genes and metabolic pathways involved in the adaptation to differences in dissolved oxygen in Nile tilapia. In the hypoxic environment, 14 significantly enriched processes were associated during the early growth stage, including nervous system development and animal organ development. *Rara* genes code for the retinoic acid receptor alpha, a transcription factor which regulates genes involved in cellular growth and differentiation [31]. In addition, *raraa* and *rarab* play an important role during zebrafish development [32]. Mediator of RNA polymerase II transcription subunit 24 (*med24*), an orthologue also found in human, mouse and zebrafish, is involved in nervous system development [33]. However, these genes and associated molecular pathways do not indicate a clear link with hypoxia when comparing to other fish studies, and rather might reflect a relation to general growth and developmental pathways.

During the later growth stage, the results of pathway enrichment suggest that candidate regions are significantly enriched for adherens junctions, oocyte meiosis, MAPK signalling pathway, VEGF signalling pathway, regulation of actin cytoskeleton and progesterone-mediated oocyte maturation. Among these six pathways, various studies in zebrafish, channel catfish, and sea bass have shown MAPK to be involved in low oxygen tolerance in fish [14, 34, 35]. VEGF signalling was shown to be essential for maintaining the vascular density and oxygen supply in tissue hypoxia [36]. Additionally, the VEGF pathway is also one of the targets of *HIF-1α*, which rapidly accumulates to activate a wide variety of genes involved in a series of responses to hypoxia [8, 37]. The candidate gene *igf1ra,* identified in this study, codes for IGF-1 receptor-a, a receptor of insulin-like growth factor that was reported to be a primary mediator of growth hormones [38]. The ephrin-A3 gene (*efna3*) is shown as a key functional mediator of hypoxic microenvironment and is regarded as a therapeutic target for

hypoxia-specific disease [39]. Retinoic acid receptor-related orphan receptor alpha (*rora*) was demonstrated to be an important mediator of *HIF-1α* activities in human [40]. Finally, the aurora kinase A (*aurka*) gene, a serine kinase in neuroblastoma that stimulate cell growth or migration, can up-regulate expression in human BE(2)-C cells under hypoxia [41]. Recently, Li et al. [16] also found that several regions were significantly related with hypoxia tolerance, including LG3, 4, 11, 14 and 22, especially two regions (LG4:15080000, LG11:24255000) are found to be adjacent with the peak in the hypoxic environment (BW5) of our study. Nevertheless, our results suggest that hypoxia has a non-significant effect on growth during the early growth stage, while, conversely, faster growing tilapia have higher tolerance to hypoxia in the later growing stage, reflected by survival probability. Interestingly, it has showed that tilapia exposure to a nocturnal hypoxia for 9 weeks led to a better growth performance than normoxia, which is related with a compensatory appetite later in the day [42]. Additionally, Roze et al. [43] has reported that fast growing fish display a better ability to maintain balance to acute hypoxia exposure than slow growing fish, by comparing two genetically different growth strains of Rainbow trout, suggesting a better hypoxia tolerance similar to the findings presented in our study.

In the normoxic environment, six biological processes were significantly enriched for BW1 and BW2, including retinoic acid receptor signalling pathway, apoptotic signalling pathway, liver development, signal transduction, steroid hormone mediated signalling pathway and brain development. Steroid hormone mediated and retinoic acid receptor signalling pathway overlapped with the same stage in the hypoxia environment, which seems mostly involved in general growth and development processes. The overlap in the early growth stage between normoxic and hypoxic environments may result shared conditions until the first time point. Another possibility is that hypoxia affected small fish less, and there still was sufficient dissolved oxygen as a result of lower overall demand environment. As fish

316 grew bigger, the metabolic impact of high growth on oxygen consumption and availability
317 may have become more pronounced [44].

318 For the later growth stage, 12 suggestive SNPs tagging regions containing 22 candidate
319 genes were identified. These included the gene coding for mitochondrial calcium uniporter
320 (*mcu*) that was reported to play a role in skeletal muscle growth and homeostasis [45]. The
321 genes coding for oncoprotein-induced transcript 3 (*oit3*) and MAP6 domain containing 1
322 (*map3d1*) were both reported to be related with calcium ion binding activity [46]. Yoshida et
323 al. [20] performed the first genome-wide association studies to unravel the genetic
324 architecture of harvest weight in a Nile tilapia population derived from a mixture of the 8th
325 generation GIFT and the wild strains from Egypt and Kenya. In that study, four regions were
326 identified that were significantly associated with harvest weight in LG12, 15, 18 and 22,
327 respectively. However, the genes lying in these regions were not significant in our study. One
328 of the reasons could be that the GIFT population has been selected on growth for many
329 generations and those regions have become fixed. This could also explain the limited number
330 of significant SNPs and candidate genes for growth observed in our study. However, it is also
331 likely the specific variants found by Yoshida et al. were never present in our population to
332 begin with.

333 The results from the meta-analysis show that five genes play a major role in growth and
334 development during the early growth stage, namely *raraa*, *rarab*, *med24*, *brms11a* and
335 *prpf38b*. Two of them (*raraa*, *rarab*) also showed significance for single GWAS in the
336 normoxic and hypoxic environments, respectively. *Prpf38b* only showed a major effect in the
337 hypoxic environment. The orthologues of this gene in human, zebrafish and mouse have been
338 shown to have a function in the central nerve system [47]. Development related genes found
339 in single GWAS, such as *raraa*, *rarab*, and *med24* were significantly associated in the meta-
340 analysis during the later stages. Nucleotide-binding protein 2 (*nump2*) was reported to be

associated with both *IGF1* and *IGFP*3 in a human GWA study [48]. Those results suggest that a few major QTL determine much of the growth rate. Even though growth rate is known to be determined by many genes [49], similarly in human [50] and cattle [51], it was found that a few genes were exceptionally important in explaining genetic variance.

Moreover, no pathway related to hypoxia tolerance was found in meta-analysis GWAS, which indicates some genes affect body weight in the hypoxic environment while different subset of genes are important for body weight under the normoxic environment (see in **Figure 4A and 4B**). This indicates genotype-by-environment interaction (GxE). However, a GxE analysis for growth rate in the normoxic versus hypoxic environment, based on a quantitative genetic analysis using a genomic relationship matrix derived from the genotyping dataset, showed that the genetic correlation was close to 0.8 [52]. This value suggests some degree of GxE and some reranking of genotypes. Furthermore, there was a large difference in body weight and its variance between environments, which suggests scaling GxE. The genetic correlation of 0.8 suggests that most fish that grow well in a normoxic environment, are also able to grow well in an environment where they experience nocturnal hypoxia. After all, Nile tilapia is a fresh water fish species that has evolved in environments where hypoxia (e.g. as a result of high temperatures, algal blooms or drought) are nocturnal events. Natural selection would favour animals that would be able to cope with these environments if larger fish would have higher reproductive success.

Conclusions

Clear associations between genotype and growth were found for both hypoxic and normoxic environments. The associated SNPs, and hence the underlying genomic architecture, however, changed over the growing period. Furthermore, the meta-analysis GWAS across two environments suggested that growth was not under the control by the same genes compared to

single environmental GWAS, which we interpret as a genotype-by-environment interaction. The functional annotation confirms that hypoxic stress pathways such as MAPK signalling pathway and VEGF signalling pathway play an important role during the later growth stage in the hypoxic environment. Our findings reveal the complexity of the genetic architecture of body weight gain under a variety of dissolved oxygen conditions in Nile tilapia, and provide an essential insight into how hypoxia affects body weight gain during the growth stage, which will benefit future tilapia breeding programmes in the context of genomic architecture.

METHODS

Animal Resource

3,140 fish used in this experiment were mass produced by broodstock samples from generation 16 of the GIFT breeding program of WorldFish at the Aquaculture Extension Centre in Jitra (Kedah, Malaysia). A full description of the experiment is given in Mengistu et al. [52]. Fish were allowed to mate naturally which was done in four hapas (net-enclosures, each $30m^2$) in a $500m^2$ earthen pond, aerated by a paddlewheel. Eighteen males and 50 female breeders were stocked for 15 days in each of the mating hapas. On the sixteenth day the breeders were removed, and fry were kept in the same hapas for nursing for a duration of 60 days.

After nursing, the fingerlings from each hapa were transferred into aerated tanks and tagged, using PIT (passive integrated transponder) tags. At tagging, a 1 cm^2 fin clip sample collected and PIT tag number and BW were recorded. Fin clip samples were preserved in 95% ethanol. A random sample of an equal number of individually tagged fingerlings from four nursery hapas was then stocked in two earthen ponds. A number of 1570 fish were stocked in each pond at a stocking density of 3 fish/m^2. 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 was without aerator to create nocturnal hypoxia environment between 11 pm to the next day 9 am from our DO measurements, which are a typical feature of earthen ponds where green algae are the main source of oxygen [52].

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 were used. After 2 months this was reduced to 3% of their body weight. 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 time intervals. 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 mornings feeding was skipped when the DO level in the hypoxic environment dropped to below 2mg/L. At these concentrations, it was observed that fish no longer eat. Body weight was measured at five time points (stocking, at 55/56 days, 104/105 days, 167/168 days, and at harvest, which was after 217 and 218 days growing in the hypoxic and normoxic environment, respectively. Fish were euthanized using clove oil at a dose of 400 ppm after the experiment. More details of the experiment can be found in Mengistu et al. [52].

Genotyping and quality control

We isolated genomic DNA 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). Individual DNA samples were digested by the ApeKI restriction enzyme, and 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. Raw sequence reads were trimmed to

remove adaptor sequence and low quality reads in Sickle (https://github.com/najoshi/sickle). The quality of each individual was evaluated by FastQC (version 1.6) [53]. Sequence mapping for 2171 individuals was performed using bwa -mem algorithm [54] aligning to the tilapia reference genome (GCF_001858045.1, https://www.ncbi.nlm.nih.gov/assembly/GCF_001858045.1/). Variant calling was done with FreeBayes (version 1.0.2) [55] with default setting excepts for the parameters: --min-base-quality 10, --haplotype-length 0 and --ploidy 2. The SNP data was further filtered by Plink (version1.9) [56] with the following exclusion criteria: Minor Allele Frequency < 2%, genotyping call-rate for SNPs < 80% and individual rate < 70%. Finally, a total of 2063 individuals and 27,090 SNPs were used for subsequent analyses.

Statistic description, Population Structure and Association Analysis

Basic statistics of all phenotypic data was handled in R (version 3.5.3). Body weight in our study is not completely following a normal distribution as estimated by Shapiro-Wilk test [57]. Therefore, we compare two paired groups at five time point using the Wilcoxon test. The phenotypic correlation was calculated by spearman's rank correlation coefficient method. Then, body weight was transformed to better fit the normal distribution by square root method [58]. To estimate the influence of factors such as hapa (early rearing environment) and sex in our experiment, they were tested in a linear model using Stepwise Algorithm [59] with the formula: with the formula: $y_{ij} = u + \alpha_i + \beta_j + \alpha_i * \beta_j + \varepsilon_{ij}$, while y is the body weight; u is the population mean; α_i is the effect of the i^{th} level of hapa; β_j is the effect of the j^{th} level of sex; ε is the random error effect. It suggested that hapa, sex and their interaction were significant with body weight. Therefore, residuals from the fixed effects model were used for the subsequent association analysis [60].

A principal component analysis (PCA) was performed to estimate population stratification before GWAS in Plink (version 1.9). The top five principal components were added as covariates and included in subsequent GWAS models as fixed effect to account for the sample structure in this association analysis. Considering the Bonferroni method being overly conservative, we corrected the threshold of genome-wide significant P-value using the SimpleM method [61]. A total of 16,504 independent effective tests were therefore obtained based on LD (linkage disequilibrium) characteristics. The significant and suggestive lines are 1 and 5% genome-wide significant divided by the SNPs number of independent SNPs in the association. Given the number of effective independent tests, the thresholds for genome-wide and suggestive significance value were calculated as 3.03E-06 (0.05/16504) and 6.06E-05 (1/16504), respectively.

We used a univariate GWAS by applying a linear mixed model to that passed filtering threshold in GEMMA [62]:

$$\boldsymbol{y} = \boldsymbol{W\alpha} + \boldsymbol{x\beta} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}$$

In this equation, **y** is the a vector of observation on body weight; $\boldsymbol{W}$ is a covariate matrix of fixed effects (top five PCs and included a column of 1s) used to adjust population structure; $\boldsymbol{\alpha}$ is a vector of the corresponding coefficient including the intercept; $\boldsymbol{x}$ is a vector of the marker genotypes and $\boldsymbol{\beta}$ is the corresponding vector of marker effects for the phenotypes; $\boldsymbol{\mu}$ is a vector of random effects on y, with $\boldsymbol{\mu} \sim \mathbf{MVN_n\ (0,\ \lambda\tau^{-1}\ K)}$; $\boldsymbol{\varepsilon}$ is the random residuals, with $\boldsymbol{\varepsilon} \sim \mathbf{MVN_n\ (0,\ \tau^{-1}\ I_n)}$; $\mathbf{MVN_n}$ is the n-dimensional multivariate normal distribution; $\boldsymbol{\lambda}$ is the ration between the two variance components; $\boldsymbol{\tau^{-1}}$ is the ratio between the two variance components; **K** is a known n × n relatedness matrix and $\mathbf{I_n}$ is an n × n identity matrix. We performed the Wald statistic for each SNP which means we tested the alternative hypothesis H_1: $\beta \neq 0$ compared to null hypothesis H_0: $\beta = 0$ for each SNP, which is one of common used in GWAS studies of quantitative traits [63] .

Meta-analysis is powerful to detect shared genetic architecture across traits [64]. Thus, an inverse-variance weighted (IVW) method based on fixed effect in Meta (Version 1.7) [65, 66] was applied to the combined normoxic and hypoxic population IVW, then summarized effect sizes utilizing the inverse variance of individual studies as weights. The weight (wi) for the ith environment was calculated by the following equation:

$$w_i = \frac{1}{s_i^2}$$

Here s_i is the standard error of the SNP effect in i^{th} environment GWAS. Then, the effect size and standard error for i^{th} environment GWAS were estimated by the following:

$$\frac{\sum_{i=1}^{2} w_i \beta_i}{\sum_{i=1}^{2} w_i}$$

$$s^2 = \frac{1}{\sum_{i=1}^{2} w_i}$$

The statistical significance was tested by constructing a z-score of IVW as follows:

$$Z = \frac{\beta}{s} = \frac{\sum_{i=1}^{2} w_j \beta_i}{\sqrt{\sum_{i=1}^{2} w\, i}}$$

Post-GWAS analysis

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. Candidate regions were defined as the genomic regions that located 20 kb upstream and downstream of the genome-wide significant SNPs. In order to identify candidate genes nearby the significant SNPs, we used the Custom Annotations function to create an annotation set with parameters (--distance 20000 --gene_phenotype --symbol) in Ensembl Variant Effect Predictor (VEP) [67].

All protein sequences of candidate genes were extracted through reference protein sequence with an inhouse python script, and were further used for functional enrichment analysis in STRING V11.0 [68]. The false discovery rate (FDR) adjusted p-value of 0.05 was used to detect significant enrichment.

Abbreviations

AMPK: AMP-activated protein kinase; AURKA: aurora kinase A; BW: body weight; CV: coefficient of variation; DO: dissolved oxygen; EFNA3: ephrin-A3; FDR: false discovery rate; GIFT: Genetically Improved Farmed tilapia; GO: Gene ontology; GWAS: Genome Wide Association Study; GxE: genotype-by-environment interaction; HAS1: Hyaluronan synthase; HIF: hypoxia-Inducible factors; IVW: inverse-variance weighted; KEGG: Kyoto Encyclopedia of Genes and Genomes; LD: linkage disequilibrium; LG: lineage group; MAPK: mitogen-activated protein kinase; MAP3D1: MAP6 domain containing 1; MCU: mitochondrial calcium uniporter; MED24: Mediator of RNA polymerase II transcription subunit 24; NUMP2: nucleotide-binding protein 2; OIT3: oncoprotein-induced transcript 3; PCA: principal component analysis; PC: principal component; PGK1: phosphoglycerate kinase 1; Q-Q: quantile-quantile; RORA: retinoic acid receptor-related orphan receptor alpha; ROS: reactive oxygen species; SD: standard deviation; SNP: single nucleotide polymorphism; TH: tyrosine hydroxylase; VEGF: vascular endothelial growth factor; VEP: Variant Effect Predictor.

Declarations

Ethics approval and consent to participate

The fish were derived from the Aquaculture Extension Centre of the Malaysian Department of Fisheries at Jitra , Kedah State, Malaysia (6°15’32°N; 100°25’47°E). This study was approved by the internal WorldFish ethics committee. All the parties agreed for this experiment.

Consent for publication

Not applicable.

Availability of data and materials

The genotyping and genotyping data is deposited to the WorldFish Dataverse (https://dataverse.harvard.edu/dataverse/worldfish), and will be public upon acceptance of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was undertaken as part of, and funded by, the CGIAR Research Program on Fish Agri-Food Systems (FISH) led by WorldFish. The original experiment was funded by the Koepon Foundation with additional support to WorldFish by the International Fund for Agricultural Development (IFAD). 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.

Authors' contributions

JAHB, HAM, HK and MAMG conceived and designed the experiments. SMB and HAM were responsible for collecting the trait data. XFY, JWMB and HJM contributed to data analysis. XFY and HJM contributed to manuscript writing. HAM, HK, JAHB and MAMG contributed to manuscript revision. All authors reviewed and approved to the final manuscript.

Acknowledgements

The authors gratefully acknowledge WorldFish breeding team for their field and technical assistance. We also thank all colleagues who provided help during analysis procedures from Animal Breeding and Genomics group.

References

1. FAO: The State of World Fisheries and Aquaculture 2020. Rome: Sustainability in action.
2. Guyon R, Rakotomanga M, Azzouzi N, Coutanceau JP, Bonillo C, D'Cotta H, Pepey E, Soler L, Rodier-Goud M, D'Hont A *et al*: A high-resolution map of the Nile tilapia genome: a resource for studying cichlids and other percomorphs. *BMC Genomics.* 2012;13:222.
3. Ponzoni RW, Nguyen NH, Khaw HL, Hamzah A, Bakar KRA, Yee HY: Genetic improvement of Nile tilapia (Oreochromis niloticus) with special reference to the work conducted by the WorldFish Center with the GIFT strain. *Reviews in Aquaculture.* 2011;3(1):27-41.
4. Portner HO: Oxygen- and capacity-limitation of thermal tolerance: a matrix for integrating climate-related stressor effects in marine ecosystems. *J Exp Biol.* 2010;213(6):881-893.
5. Fitzgibbon QP, Strawbridge A, Seymour RS: Metabolic scope, swimming performance and the effects of hypoxia in the mulloway, Argyrosomus japonicus (Pisces : Sciaenidae). *Aquaculture.* 2007;270(1-4):358-368.
6. Mengistu SB, Mulder HA, Benzie JAH, Komen H: A systematic literature review of the major factors causing yield gap by affecting growth, feed conversion ratio and survival in Nile tilapia (Oreochromis niloticus). *Reviews in Aquaculture.* 2020;12(2):524-541.
7. Michiels C: Physiological and pathological responses to hypoxia. *Am J Pathol.* 2004;164(6):1875-1882.
8. Majmundar AJ, Wong WHJ, Simon MC: Hypoxia-Inducible Factors and the Response to Hypoxic Stress. *Mol Cell.* 2010;40(2):294-309.
9. Nikinmaa M, Rees BB: Oxygen-dependent gene expression in fishes. *American journal of physiology Regulatory, integrative and comparative physiology.* 2005;288(5):R1079-1090.
10. Xiao W: The hypoxia signaling pathway and hypoxic adaptation in fishes. *Sci China Life Sci.* 2015;58(2):148-155.
11. Zhu CD, Wang ZH, Yan BA: Strategies for hypoxia adaptation in fish species: a review. *Journal of Comparative Physiology B-Biochemical Systems and Environmental Physiology.* 2013;183(8):1005-1013.

12. Du SNN, Mahalingam S, Borowiec BG, Scott GR: Mitochondrial physiology and reactive oxygen species production are altered by hypoxia acclimation in killifish (Fundulus heteroclitus). *J Exp Biol.* 2016;219(8):1130-1138.
13. Zhong X, Wang X, Zhou T, Jin Y, Tan S, Jiang C, Geng X, Li N, Shi H, Zeng Q *et al*: Genome-Wide Association Study Reveals Multiple Novel QTL Associated with Low Oxygen Tolerance in Hybrid Catfish. *Marine biotechnology.* 2017;19(4):379-390.
14. Wang XZ, Liu SK, Jiang C, Geng X, Zhou T, Li N, Bao LS, Li Y, Yao J, Yang YJ et al: Multiple across-strain and within-strain QTLs suggest highly complex genetic architecture for hypoxia tolerance in channel catfish. Molecular Genetics and Genomics. 2017;292(1):63-76.
15. Anttila K, Dhillon RS, Boulding EG, Farrell AP, Glebe BD, Elliott JAK, Wolters WR, Schulte PM: Variation in temperature tolerance among families of Atlantic salmon (Salmo salar) is associated with hypoxia tolerance, ventricle size and myoglobin level. *J Exp Biol.* 2013;216(7):1183-1190.
16. Li HL, Gu XH, Li BJ, Chen CH, Lin HR, Xia JH: Genome-Wide QTL Analysis Identified Significant Associations Between Hypoxia Tolerance and Mutations in the GPR132 and ABCG4 Genes in Nile Tilapia. *Marine biotechnology.* 2017;19(5):441-453.
17. Aquaculture Genomics G, Breeding W, Abdelrahman H, ElHady M, Alcivar-Warren A, Allen S, Al-Tobasei R, Bao L, Beck B, Blackburn H et al: Aquaculture genomics, genetics and breeding in the United States: current status, challenges, and priorities for future research. *BMC Genomics*. 2017;18(1):191.
18. Zhong XX, Wang XZ, Zhou T, Jin YL, Tan SX, Jiang C, Geng X, Li N, Shi HT, Zeng QF *et al*: Genome-Wide Association Study Reveals Multiple Novel QTL Associated with Low Oxygen Tolerance in Hybrid Catfish. *Marine biotechnology.* 2017;19(4):379-390.
19. Brennan RS, Healy TM, Bryant HJ, Van La M, Schulte PM, Whitehead A: Integrative Population and Physiological Genomics Reveals Mechanisms of Adaptation in Killifish. *Molecular biology and evolution.* 2018;35(11):2639-2653.
20. Yoshida GM, Lhorente JP, Correa K, Soto J, Salas D, Yanez JM: Genome-Wide Association Study and Cost-Efficient Genomic Predictions for Growth and Fillet Yield in Nile Tilapia (Oreochromis niloticus). *G3*. 2019;9(8):2597-2607.
21. Mandic M, Todgham AE, Richards JG: Mechanisms and evolution of hypoxia tolerance in fish. *Proc Biol Sci.* 2009;276(1657):735-744.
22. Almeida-Val VM, Val AL, Duncan WP, Souza FC, Paula-Silva MN, Land S: Scaling effects on hypoxia tolerance in the Amazon fish Astronotus ocellatus (Perciformes: Cichlidae): contribution of tissue enzyme levels. *Comp Biochem Physiol B Biochem Mol Biol.* 2000;125(2):219-226.
23. Sloman KA, Wood CM, Scott GR, Wood S, Kajimura M, Johannsson OE, Almeida-Val VM, Val AL: Tribute to R. G. Boutilier: the effect of size on the physiological and behavioural responses of oscar, Astronotus ocellatus, to hypoxia. *J Exp Biol.* 2006;209(Pt 7):1197-1205.
24. Ishibashi Y, Inoue K, Nakatsukasa H, Ishitani Y, Miyashita S, Murata O: Ontogeny of tolerance to hypoxia and oxygen consumption of larval and juvenile red sea bream, Pagrus major. *Aquaculture.* 2005;244(1-4):331-340.
25. Burleson ML, Wilhelm DR, Smatresk NJ: The influence of fish size on the avoidance of hypoxia and oxygen selection by largemouth bass. *J Fish Biol.* 2001;59(5):1336-1349.
26. Robb T, Abrahams MV: Variation in tolerance to hypoxia in a predator and prey species: an ecological advantage of being small? *J Fish Biol.* 2003;62(5):1067-1081.

27. Vejrik L, Matejickova I, Juza T, Frouzova J, Seda J, Blabolil P, Ricard D, Vasek M, Kubecka J, Riha M *et al*: Small fish use the hypoxic pelagic zone as a refuge from predators. *Freshwater Biol.* 2016;61(6):899-913.
28. Nilsson GE, Ostlund-Nilsson S: Does size matter for hypoxia tolerance in fish? *Biol Rev.* 2008;83(2):173-189.
29. Abdel-Tawwab M, Monier MN, Hoseinifar SH, Faggio C: Fish response to hypoxia stress: growth, physiological, and immunological biomarkers. *Fish Physiol Biochem.* 2019;45(3):997-1013.
30. Abdel-Tawwab M, Hagras AE, Elbaghdady HAM, Monier MN: Effects of dissolved oxygen and fish size on Nile tilapia, Oreochromis niloticus (L.): growth performance, whole-body composition, and innate immunity. *Aquacult Int.* 2015;23(5):1261-1274.
31. Wang XJ, Dasari S, Nowakowski GS, Lazaridis KN, Wieben ED, Kadin ME, Feldman AL, Boddicker RL: Retinoic acid receptor alpha drives cell cycle progression and is associated with increased sensitivity to retinoids in T-cell lymphoma. *Oncotarget.* 2017;8(16):26245-26255.
32. Hale LA, Tallafuss A, Yan YL, Dudley L, Eisen JS, Postlethwait JH: Characterization of the retinoic acid receptor genes raraa, rarab and rarg during zebrafish development. *Gene Expr Patterns.* 2006;6(5):546-555.
33. Harrison C, Wabbersen T, Shepherd IT: In Vivo Visualization of the Development of the Enteric Nervous System Using a Tg(-8.3bphox2b:Kaede) Transgenic Zebrafish. *Genesis.* 2014;52(12):985-990.
34. Tian Y, Wen HS, Qi X, Zhang XY, Li Y: Identification of mapk gene family in Lateolabrax maculatus and their expression profiles in response to hypoxia and salinity challenges. *Gene.* 2019;684:20-29.
35. Marques IJ, Leito JTD, Spaink HP, Testerink J, Jaspers RT, Witte F, van Den Berg S, Bagowski CP: Transcriptome analysis of the response to chronic constant hypoxia in zebrafish hearts. *Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology.* 2008;178(1):77-92.
36. Maitland ML, Lou XJ, Ramirez J, Desai AA, Berlin DS, McLeod HL, Weichselbaum RR, Ratain MJ, Altman RB, Klein TE: Vascular endothelial growth factor pathway. *Pharmacogenet Genom.* 2010;20(5):346-349.
37. Kurihara T, Westenskow PD, Friedlander M: Hypoxia-Inducible Factor (HIF)/Vascular Endothelial Growth Factor (VEGF) Signaling in the Retina. *Adv Exp Med Biol.* 2014;801:275-281.
38. Hakuno F, Takahashi SI: IGF1 receptor signaling pathways. *J Mol Endocrinol.* 2018;61(1):T69-T86.
39. Husain A, Chiu YT, Ho DW, Sze KM, Chan LK, Tsui YM, Wong CC, Ng IO: EFNA3, a key functional mediator of hypoxic microenvironment in hepatocellular carcinoma. *Cancer Research.* 2018;78(13).
40. Li H, Zhou L, Dai J: Retinoic acid receptor-related orphan receptor RORalpha regulates differentiation and survival of keratinocytes during hypoxia. *J Cell Physiol.* 2018;233(1):641-650.
41. Romain CV, Paul P, Lee S, Qiao J, Chung DH: Targeting aurora kinase A inhibits hypoxia-mediated neuroblastoma cell tumorigenesis. *Anticancer Res.* 2014;34(5):2269-2274.
42. Obirikorang KA, Acheampong JN, Duodu CP, Skov PV: Growth, metabolism and respiration in Nile tilapia (Oreochromis niloticus) exposed to chronic or periodic hypoxia. *Comp Biochem Physiol A Mol Integr Physiol.* 2020;248:110768.
43. Roze T, Christen F, Amerand A, Claireaux G: Trade-off between thermal sensitivity, hypoxia tolerance and growth in fish. *J Therm Biol.* 2013;38(2):98-106.

44. Mandic M, Regan MD: Can variation among hypoxic environments explain why different fish species use different hypoxic survival strategies? *J Exp Biol.* 2018;221(21).
45. Kwong JQ, Huo JZ, Bround MJ, Boyer JG, Schwanekamp JA, Ghazal N, Maxwell JT, Jang YC, Khuchua Z, Shi K *et al*: The mitochondrial calcium uniporter underlies metabolic fuel preference in skeletal muscle. *Jci Insight.* 2018;3(22).
46. Sprague J, Bayraktaroglu L, Clements D, Conlin T, Fashena D, Frazer K, Haendel M, Howe DG, Mani P, Ramachandran S *et al*: The Zebrafish Information Network: the zebrafish model organism database. *Nucleic acids research.* 2006;34(Database issue):D581-585.
47. Koscielny G, Yaikhom G, Iyer V, Meehan TF, Morgan H, Atienza-Herrero J, Blake A, Chen CK, Easty R, Di Fenza A *et al*: The International Mouse Phenotyping Consortium Web Portal, a unified point of access for knockout mice and related phenotyping data. *Nucleic acids research.* 2014;42(Database issue):D802-809.
48. Teumer A, Qi Q, Nethander M, Aschard H, Bandinelli S, Beekman M, Berndt SI, Bidlingmaier M, Broer L, Group CLW *et al*: Genomewide meta-analysis identifies loci associated with IGF-I and IGFBP-3 levels with impact on age-related traits. *Aging Cell.* 2016;15(5):811-824.
49. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chun AY, Estrada K, Luan J, Kutalik Z *et al*: Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature genetics.* 2014;46(11):1173-1186.
50. Couto Alves A, De Silva NMG, Karhunen V, Sovio U, Das S, Taal HR, Warrington NM, Lewin AM, Kaakinen M, Cousminer DL *et al*: GWAS on longitudinal growth traits reveals different genetic factors influencing infant, child, and adult BMI. *Sci Adv.* 2019;5(9):eaaw3095.
51. Seabury CM, Oldeschulte DL, Saatchi M, Beever JE, Decker JE, Halley YA, Bhattarai EK, Molaei M, Freetly HC, Hansen SL *et al*: Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics.* 2017;18(1):386.
52. Mengistu SB, Mulder HA, Benzie JA, Khaw HL, Megens H-J, Trinh TQ, Komen H: Genotype by environment interaction between aerated and non-aerated ponds and the impact of aeration on genetic parameters in Nile tilapia (*Oreochromis niloticus*). *Aquaculture*. 2020:735704.
53. Andrews, S. FastQC: a quality control tool for high throughput sequence data. (http://www.bioinformatics.babraham.ac.uk/projects/fastqc)
54. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-1760.
55. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. arXiv Prepr arXiv12073907. 2012;9.
56. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ *et al*: PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics.* 2007;81(3):559-575.
57. Villasenor Alva JA, Estrada EG: A generalization of Shapiro–Wilk's test for multivariate normality. *Communications in Statistics—Theory and Methods.* 2009;38(11):1870-1883.
58. McDonald JH: Handbook of biological statistics, vol. 2: sparky house publishing Baltimore, MD; 2009.
59. Neerchal NK, Morel J, Huang X, Moluh A: A stepwise algorithm for generalized linear mixed models. In: *SAS Global Forum, Washington, DC: 2014*; 2014: 1822-2014.

60. Gondro C, Van der Werf J, Hayes B: Genome-wide association studies and genomic prediction: Springer; 2013.
61. Gao XY, Becker LC, Becker DM, Starmer JD, Province MA: Avoiding the High Bonferroni Penalty in Genome-Wide Association Studies. *Genetic epidemiology*. 2010;34(1):100-105.
62. Zhou X, Stephens M: Genome-wide efficient mixed-model analysis for association studies. *Nature genetics.* 2012;44(7):821-U136.
63. Wu B, Pankow JS: Fast and Accurate Genome-Wide Association Test of Multiple Quantitative Traits. *Comput Math Methods Med*. 2018;2018:2564531.
64. Li YR, Li J, Zhao SD, Bradfield JP, Mentch FD, Maggadottir SM, Hou C, Abrams DJ, Chang D, Gao F et al: Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. *Nat Med*. 2015;21(9):1018-1027.
65. Liu JZ, Tozzi F, Waterworth DM, Pillai SG, Muglia P, Middleton L, Berrettini W, Knouff CW, Yuan X, Waeber G *et al*: Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nature genetics.* 2010;42(5):436-U475.
66. Lee CH, Cook S, Lee JS, Han B: Comparison of Two Meta-Analysis Methods: Inverse-Variance-Weighted Average and Weighted Sum of Z-Scores. *Genomics Inform.* 2016;14(4):173-180.
67. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F: The Ensembl Variant Effect Predictor. *Genome Biol.* 2016;17(1):122.
68. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P *et al*: STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research.* 2019;47(D1):D607-D613.

Figure Legends

Figure 1 SNP statistics with all individuals.

(A) Histogram of SNPs distribution across all linkage groups. (B) SNP density plots across all linkage groups. (C and D) 3D PC plot for origin of tilapia at BW1 in the hypoxic (C) and normoxic (D) environments using all SNPs that passed filtering, where each dot represents one individual.

Figure 2 Manhattan plots across the whole growth period in the hypoxic environment. (A) and normoxic environments (B). Each dot on this figure corresponds to a SNP within the dataset, while the orange and blue horizontal line represent the genome-wide significance (5.52) and suggestive significance threshold value (4.22), respectively. The Manhattan plots contain –log10 observed P-values for genome-wide SNPs (y-axis) plotted against their corresponding position on each chromosome (x-axis).

Figure 3 Manhattan plots of Meta-analysis GWAS across two environments. The orange and blue horizontal line represent the genome-wide significance (3.03E-06) and suggestive significance threshold value (6.06E-05) respectively

Figure 4 Functional annotation based on candidate genomic region associated with growth. (A) Venn diagram summarising the gene count of the early stage (BW1 to BW2) from hypoxia, normoxia and meta-analysis (cross normoxia and hypoxia). (B) Venn diagram summarizing gene count of later stage (BW3 to BW5) from hypoxia, normoxia and meta-analysis. (C) KEGG enrichment of candidate genes in later stage of hypoxia environment (D) protein association network among candidate genes in later stage of the hypoxia environment.

Supplementary information

Supplementary Figure 1. Body weight comparison amongst four hapas in the normoxic and hypoxic environments.

Supplementary Figure 2. Two-dimensional plots of all individuals using SNP markers in the hypoxic environment.

Supplementary Figure 3. Two-dimensional plots of all individuals using SNP markers in the normoxic environment.

Supplementary Figure 4. Quantile-quantile plots through the whole growth stage in the hypoxic environment.

Supplementary Figure 5. Quantile-quantile plots through the whole growth stage in the normoxic environment.

Supplementary Table 1. Information of genome-wide significant and suggestive SNPs in the hypoxic environment.

Supplementary Table 2. Information of genome-wide significant and suggestive SNPs in the normoxic environment.

Supplementary Table 3. Information of genome-wide significant SNPs across the two environments by meta-analysis.

Supplementary Table 4. Functional enrichment for the early stage in the hypoxic environment.

Supplementary Table 5. Functional enrichment for the early stage in the normoxic environment.

Supplementary Table 6. Functional enrichment for the whole growth stage across the two environments by meta-analysis.