Application of genome-wide association studies in channel catfish to identify QTL associated with hypoxia tolerance

by

Xiaozhu Wang

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 6, 2017

Keywords: catfish, hypoxia tolerance, QTL, GWAS, genetic architecture

Approved by

Zhanjiang Liu, Chair, Professor of School of Fisheries, Aquaculture, and Aquatic Sciences
Rex Dunham, Professor of School of Fisheries, Aquaculture, and Aquatic Sciences
Nannan Liu, Professor of Department of Entomology and Plant Pathology
Charles Y. Chen, Associate Professor of Department of Crop, Soil and Environmental Sciences

Abstract

Catfish is the dominant aquaculture species in the United States. The production of the domestic catfish industry rapidly increased in the last two decades. However, as imported catfish from Southeast Asia have kept increasing, the domestic catfish industry faces a great challenge in these few years. To compete with the imported catfish, the domestic catfish industry need to reduce the production costs and enhance the production efficiency by improving various performance traits. However, little information is known about the genetic architecture controlling economically important traits, which hinders marker-assisted selection.

Hypoxia is common in aquaculture conditions because of high stocking densities. Although channel catfish (*Ictalurus punctatus*) is highly tolerant to low concentration of dissolved oxygen, hypoxia still causes enormous economic losses each year. Understanding how genetic architecture and environmental factors affect the hypoxia tolerance is of great interest to aquaculture. In this project, the aim was to investigate the effects of genetic background and environmental factors on hypoxia tolerance in channel catfish. Firstly, the effects of environmental factors such as gender and body size on hypoxia tolerance were investigated. Secondly, six strains of channel catfish were compared for their tolerance to hypoxic stress, including Marion Random, Marion Select, Thompson, Kansas Random, Kmix (*Kansas × Kansas Select*), and 103KS (*NWAC 103 × Kansas Select*). Thirdly, multiple within-strain and across-strain QTL were identified to be associated with the hypoxia tolerance in channel catfish by conducting a genome-wide association study (GWAS).

Finally, genes within the associated genomic regions were identified for their potential involvement in responses to hypoxia.

Six strains of channel catfish were compared for their hypoxia tolerance under a lethal level of dissolved oxygen (0.1 mg/L) by using survival analysis. 103KS and Marion S strains had significantly higher hypoxia tolerance, while Marion strain had the poorest hypoxia tolerance. In addition, effects of gender and body weight on hypoxia tolerance were also identified in catfish. Body weight was positively correlated with the tolerance to hypoxic stress when catfish were in relatively small size range (within 200 g). No significant difference of hypoxia tolerance was observed between female and male channel catfish.

To reveal the genetic architecture of hypoxia tolerance in catfish, a GWA study was conducted to identify QTL for hypoxia tolerance using the catfish 250K SNP array with channel catfish families from six strains. Multiple significant and suggestive QTL were identified both across strains and within strains. One significant QTL and four suggestive QTL were identified across strains. Six significant QTL and many suggestive QTL were identified within strains. There were rare overlaps among the QTL identified within the six strains, suggesting a complex genetic architecture of hypoxia tolerance. Overall, within-strain QTL explained larger proportion of phenotypic variation than across-strain QTL. Many of genes within these identified genomic regions have known functions for regulation of oxygen metabolism and involvement in hypoxia responses. Pathway analysis indicated that most of these genes were involved in MAPK and/or PI3K/AKT/mTOR signaling pathways that were known to be important for hypoxia-mediated angiogenesis, cell proliferation, apoptosis and survival.

Acknowledgments

I would like to express my sincere appreciation to my committee chair Professor, Dr. Zhanjiang Liu, who has the attitude and the substance of a real scientist. His passion and dedication in research deeply influence me on my own research. Without his guidance, encouragement, and persistent support this dissertation would not have been possible.

I would like to thank my committee members, Dr. Rex Dunham, Dr. Nannan Liu, and Dr. Charles Y. Chen, and my University reader Dr. Joseph Kloepper, for their guidance and assistance of my dissertation. Their suggestions make my dissertation more logical and deep my understanding of how to be a good researcher.

In addition, I am grateful for all the technical support from Dr. Huseyin Kucuktas and Ludmilla Kaltenboeck. My thank also goes to Dr. Shikai Liu, Dr. Chen Jiang, Dr. Yun Li, Dr. Xin Geng, Dr. Lisui Bao, Dr. Jun Yao, Dr. Qifan Zeng, Dr. Sen Gao, Ning Li, Tao Zhou, Yujia Yang, Qiang Fu, Zihao Yuan, Suxu Tan and all the other colleagues in the laboratory for their help, collaboration and friendship.

Finally, I would like to express my deepest appreciation to my beloved parents for their endless love and tremendous support.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Problem statement

Oxygen is essential for life of all aerobic organisms. They must adapt to low oxygen environments for survival when being exposed to hypoxic stress. As a matter of fact, the ability to survive hypoxic conditions determines the distribution of organisms. Land animals are living in relatively high oxygen conditions (~21%), and even so, animals living in high altitudes may still experience frequent exposures to hypoxia. In contrast, aquatic animals are routinely living in low oxygen conditions that has driven the evolutionary changes allowing them to survive in aquatic environments with relatively low dissolved oxygen (in the range of ~10 ppm). These aquatic animals, especially teleost fish, are ideal for hypoxia study due to their extremely efficiency in utilizing low concentration of dissolved oxygen in water.

With teleost fish, especially aquaculture species, understanding effects of environmental factors and genetic background on hypoxia tolerance is important because of several reasons: (1) they live in aquatic environments with high adaptability of a range of hypoxic conditions, making them the best natural models for understanding mechanisms of hypoxia responses; (2) improving hypoxia tolerance is of economic interest as hypoxia can cause enormous economic losses. Under aquaculture conditions, exposures to hypoxia can lead to a large number of mortalities due to high stocking density. Each year, hypoxia directly kills thousands of fish, leading to enormous losses in catfish industry. Aeration is essential to maintain the concentration of dissolved oxygen in the intense aquaculture systems; however, it is directly associated with energy cost. Furthermore,

exposure to hypoxia has been reported to cause depression in immune system, leading to increased susceptibility to diseases (Kvamme et al., 2013; Welker et al., 2007). Hypoxia also negatively affects fish behaviors, including predation, swimming and even reproductive behaviors (Pollock et al., 2007).

As the dominant species in the aquaculture industry of the United States, catfish is one of the most economically important agricultural commodities. The catfish industry provides large numbers of job positions for Alabama, Arkansas, Louisiana and Mississippi states. Catfish raised by farmers include different strains or families of channel catfish, two major strains of blue catfish, and channel catfish × blue catfish hybrid F1. Recent years, the continuous increases of imported catfish from Southeast Asia, feed and energy costs, have posed a formidable challenge to the domestic catfish industry. To maintain the domestic catfish industry, research has been focus on the improvement of production traits and reduction of production costs. Although channel catfish (*Ictalurus punctatus*) is highly tolerant to low concentration of dissolved oxygen, hypoxia still causes enormous mortalities. To improve hypoxia tolerance, the first step is to understand how genetic background and environmental factors affect the tolerance to hypoxia in catfish. In this dissertation, the environmental effectors and genetic architecture of hypoxia tolerance in catfish were investigated to set a foundation for future improvement of hypoxia tolerance in catfish.

1.2 Review of literatures

1.2.1 Mechanism of hypoxia responses

The ability to maintain O_2 homeostasis is essential to the survival of all invertebrate and vertebrate species. Physiological systems have evolved to ensure the optimal oxygenation of all cells in each organism. In humans and other vertebrates, the precise establishment and regulation

of systems, such as lungs, erythrocytes, heart and vasculature, provide a major basis for O_2 homeostasis. The O_2 homeostasis is the key to maintain the proper biological functions and survival of all organisms living an aerobic life style. For instance, organisms living at high altitude face hypobaric hypoxia due to the lowered barometric pressure, whereas they must adapt to the stress of limited oxygen availability and sustain aerobic metabolic processes. Not only is O_2 homeostasis essential for organism survival, but also hypoxia widely involves in the pathogenesis of major causes of mortality, such as cancer, chronic heart, and lung diseases. However, the definition of hypoxia is varied among different species due to their living environment and genetic background.

Response to hypoxia is a highly complex biological process. Under hypoxic conditions, organisms usually develop alternative strategies of energy metabolism, enzymatic activities and signal transduction pathways to survive. Because of the decrease of aerobic metabolic rates in hypoxia-sensitive cells under hypoxic conditions, the demand for glucose or glycogen for anaerobic glycolysis may drastically rise to make up for the energetic shortfall. However, because of energy insufficiency, ion and electrical potentials typically cannot be sustained, leading to the decouple of metabolic and membrane functions in effect (Hochachka, 1986). These problems can be concurred through a number of biological mechanisms in hypoxia-tolerant animals. Amongst these are metabolic arrest and stabilized membrane functions, which are effective strategies for hypoxia responses (Hochachka, 1986). For instance, metabolic arrest is achieved by means of a reversed or negative Pasteur effect (reduced or unchanging glycolytic flux at reduced O_2 availability). In spite of the lower energy turnover rates, coupling of metabolic and membrane function can be achieved by maintaining membranes of low permeability (probably via reduced densities of ion-specific channels).

It has been long believed that the control of hypoxia responses is at the post-translational level. Hypoxia inducible factors (HIFs), especially HIF-1 and HIF-2, are believed to be central mediators of the homeostatic responses that enable hypoxic cells to survive (Semenza, 2000). These proteins regulate a series of autonomous, autocrine, paracrine and endocrine effects with the overall goal of increasing oxygen delivery to tissues while decreasing their oxygen consumption (Semenza, 2012). HIF-1 is a heterodimer consisting of HIF-1α and HIF-1β with the basic helix-loop-helix (bHLH) and PER-ARNT-SIM homology (PAS) domains. HIF-1α and HIF-1β mRNA are expressed in most, if not all, human and rodent tissues (Wiener et al., 1996). HIF-1α protein, in spite of being continuously synthesized, is rapidly degraded under normoxic conditions. Under hypoxic conditions, HIF-1α rapidly accumulates and dimerizes to its partner HIF-1β. Upon recruitment of various transcriptional co-activators, which then binds to the hypoxia-response element (HRE) of effector genes, they then activate a wide variety of genes involved in both the cellular and systematic responses to hypoxia (Kietzmann et al., 2016). In addition to regulation at the post-translational level, HIF-1α gene was found to be transcriptionally induced by hypoxia in catfish (Geng et al., 2014). Similarly, the HIF-2α homologue is also regulated by oxygen tension, and dimerizes a transcription complex with HIF-1β (Keith et al., 2012). However, the biological functions of HIF-1α and HIF-2α are overlapped but unique (Liu and Simon, 2004). For instance, regulation of enzyme expression of the glycolytic pathway is a unique HIF-1α function, whereas erythropoiesis in vivo seems to be specifically controlled by HIF-2α; modulation of angiogenesis is a shared function of both HIF-1α and HIF-2α (Keith et al., 2012). Less is known about HIF-3α compared with the other HIF-α homologs. The inhibitory PAS domain protein (IPAS) was reported to be an alternatively spliced variant of HIF-3α and functions as a dominant-negative regulator of HIF-α, adding to the complexity in the regulation of hypoxia-inducible genes by the HIF (Makino

et al., 2002). It appears that HIF-1α generally signals the existence of hypoxia to the transcriptional machinery in the nucleus of all cells, whereas HIF-2α and HIF-3α function more specialized roles in O_2 homeostasis (Semenza, 2000).

The intracellular oxygen tension is not therefore the only upstream regulator of HIF proteins. Unlike hypoxia, which increases HIF-1α levels in all cell types, growth-factor stimulation induces HIF-1α expression in a cell type-specific manner. Interestingly, growth factors can control the expression of either HIF-1α or HIF-2α in a hypoxia-independent manner (Hu et al., 2003). Moreover, HIF activation can be regulated by tumor suppressors such as tumor protein p53 and von Hippel-Lindau, or oncogenes such as RAS, mammalian target of rapamycin (mTOR), phosphatidylinositol 3-kinase (PI3K), and AKT (Kietzmann et al., 2016). For instance, loss of p53 in tumor cells was found to enhance HIF-1α levels (Ravi et al., 2000). Heat shock protein 90 (Hsp90) can also interact directly with HIF-1α and has been suggested to promote a conformational change in HIF-1α, which lead to the dimerization with HIF-1β (Gradin et al., 1996). Growth factors, cytokines and other signaling molecules were also reported to stimulate HIF-1α synthesis via activation of the PI3K or mitogen-activated protein kinase (MAPK) pathways (Xia et al., 2012). The PI3K/Akt pathway, and its downstream target mTOR have been shown to increase levels of HIF-1α protein (Zhong et al., 2000). Several studies demonstrated that hypoxia induced increase of intracellular secondary messengers such as Ca^{2+} and mitochondrial reactive oxygen species (ROS) can also promote the accumulation of HIF-1α protein. For example, Ca^{2+}/CaM and its target genes such as CaM kinase II, calcineurin, and actin have been revealed in activating the expression of transcriptional activity of HIF-1α (Jung et al., 2010).

As to the downstream of HIF pathway, HIF-1 activity regulates more than 100 putative HIF target genes that are involved in many aspects of angiogenesis, cell survival, glucose

metabolism and invasion (Hong et al., 2004). For example, genes like the glucose transporter Glut-1 and carboanhydrase IX (CAIX) are involved in adaptation to low glucose levels; genes like vascular endothelial growth factor (VEGF) and erythropoietin (EPO) are key growth factors in erythropoiesis and angiogenesis; genes encoding cathepsin D, matrix metalloproteinase 2, urokinase plasminogen-activator receptor (uPAR) and vimentin play established roles in the pathophysiology of invasion; hypoxia and HIF-1 induce growth factors such as insulin-like factor-2 (IGF2) and transforming growth factor-α (TGFα) activate signal transduction pathways that lead to cell proliferation/survival and stimulates expression of HIF-1a itself (Semenza, 2000). In particular, VEGF is the main angiogenic factor induced by the HIF signalling pathway, which has important roles in various physiological and pathological conditions (Zelzer and Olsen, 2005). HIF-1 activates expression of these genes by binding to a 50-base pair cis-acting HRE located in their enhancer and promoter regions.

At the genomic level, living in the hypoxic conditions can place selective pressure for genotypes that are more tolerant to hypoxia. Most of genome level research of hypoxia tolerance has been conducted with plants. Submergence or flood is one of the major harmful abiotic stresses leading to crop losses due to waterlogging (hypoxic environment), thus waterlogging tolerance has been widely studied with soybean and major cereal crops, such as rice, wheat, barley, and maize. For instance, a single quantitative trait locus (QTL) was identified for waterlogging tolerance in soybean, which was responsible for better plant growth and grain yields, 11–18% and 47–180%, respectively (VanToai et al., 2001). A total of 20 putative quantitative trait loci (QTL) for waterlogging-response traits have been mapped in maize accounted individually 4.0–31.7 % of the phenotypic variation (Qiu et al., 2007). In two barley double haploids (DH) population, twenty QTL associated with waterlogging tolerance were identified and most of the them were associated

with leaf chlorosis and plant survival (Li et al., 2008). With animals, much less work has been conducted. However, QTL associated with tolerance to high altitude hypoxia were identified in human. High-altitude human populations such as Sherpas, Tibetans, Ethiopians, and Andeans were found to possess specific single-nucleotide polymorphism (SNP) markers associated with hypoxia tolerance, however, evolutionary processes may have acted differently on the colonizing populations to cause the different patterns of adaptation (Beall, 2006; Jha et al., 2016; Yi et al., 2010). For instance, one SNP in endothelial Per-Arnt-Sim domain protein 1 (EPAS1), a transcription factor involved in response to hypoxia, was reported to show a 78% frequency difference between Tibetan and Han samples, representing the fastest allele frequency change observed at any human gene to date (Yi et al., 2010). Identification of genetic architecture of hypoxia tolerance in aquaculture species is important, because it not only provides insight into the mechanism of hypoxia responses, but also sets the foundation for genetic improvement of this trait. However, no QTL studies have been conducted to identify genomic regions associated with hypoxia tolerance in any fish species so far.

1.2.2 Hypoxia study in teleost fish

Low dissolved oxygen (DO) concentration has always been a threat in the aquatic ecosystems. Natural phenomena such as weather, temperature or flow rate can cause low DO concentration in water. In addition, water pollution such as eutrophication and organic pollution, can lead to severe hypoxic conditions. Under the intense aquaculture conditions, due to high density stocking, hypoxia is a major concern although aerators can be used to maintain the DO concentration in the systems.

Teleost fish are extremely efficient in utilizing low levels of dissolved oxygen in water, making them an ideal model for hypoxia study. Many fish species have evolved various adaptive

strategies, allowing them to survive under hypoxic conditions. Behavioral adaptions include shifts from water breathing to surface breathing (Kramer and McClure, 1982) and lowering locomotor activities (Nilsson et al., 1993). Physiological adaptions include: (1) an enhanced ventilation in gill with extremely large surface areas to increase the effective respiratory area (Burggren and Cameron, 1980); (2) the metabolic processes are reduced or shutting down for energy saving (Jensen et al., 1993); (3) hypoxia can trigger cells to release catecholamine hormones, adrenaline and noradrenaline for response (Zhang et al., 2009); (4) hypoxia can change the enzymatic activities, such as glycolytic enzymes, malate dehydrogenase and lactate dehydrogenase (Nikinmaa and Rees, 2005).

The negative effects of exposures to hypoxia on fishes have been studied a lot in the past few years. Hypoxia may lead to the dramatically changes in fish behaviors (Pollock et al., 2007): (1) hypoxia may cause fish more vulnerable to predation and less efficient to forage for predators; (2) hypoxic conditions may lead to the failure of school formation; (3) hypoxia may lead to less or even no reproductive activities and cause the high mortality in fish eggs; (4) hypoxia can also cause a substantial reduction in embryo motility and heartbeat rate. The significant reduction of metabolic rate also was reported in fish such as paddlefish (Aboagye and Allen, 2014) and goldfish (Vanwaversveld et al., 1988). In addition, hypoxic stress may also affect the growth, yield and enteric septicemia (ESC) mortality of channel catfish (Burggren and Cameron, 1980; Welker et al., 2007).

Hypoxia tolerance is phylogenetically independently associated with three principal components accounting for 75% of phenotypic variation: routine O_2 consumption rate; mass-specific gill surface area; and whole blood haemoglobin (Hb)–O_2-binding affinity (Mandic et al., 2009). In addition, a number of environmental factors like gender, body size, and of course, genetic

background have been reported to affect hypoxia tolerance of fish. For instance, higher mortalities were reported for males than females with mosquitofish under the extreme hypoxic conditions (Cech Jr et al., 1985). Size as reflected in body weight is another factor to affect performance traits in fish. It is commonly noted that fish with larger body weight tend to be less tolerant to hypoxia due to smaller ratio of gill surface area to total body volume, as reported for largemouth bass (Burleson et al., 2001). However, opposite observations were also reported. For instance, Almeida-Val et al. (2000) and Sloman et al. (2006) reported that larger fish had a higher hypoxia tolerance than smaller ones for Oscar cichlid (*Astronotus ocellatus*). With regard to genetic background, strains were found to affect hypoxia tolerance of their hybrids in catfish (Dunham et al., 2014).

Hypoxia can alter the expression patterns of many genes as a part of the regulatory cascade to ameliorate pathological effects. Gene expression changes were reported in many fish species including zebrafish (Guan et al., 2011), rainbow trout (Faust et al., 2004), Atlantic salmon (Anttila et al., 2013; Niklasson et al., 2011) and catfish (Dunham et al., 2014). The function of HIF proteins in fish is similar as that in mammal. However, fish HIF-1α may accumulate even in normoxia due to the high occurrence of hypoxia in aquatic environments (Nikinmaa and Rees, 2005). VEGF and its receptors were regulated by HIF under the hypoxic condition in salmon (Vuori et al., 2004). Hypoxia was also reported to enhance mRNA levels of lactate dehydrogenase A and phosphoglycerate kinase 1, whereas inhibit hemoglobin α and β mRNAs significantly (Burggren and Cameron, 1980; Roesner et al., 2006). Two biological pathways, ubiquitin-proteasome pathway and phosphatidylinositol signaling pathway, were significantly dysregulated by hypoxia in medaka (*Oryzias latipes*) (Zhang et al., 2009). Tissue-specific expression patterns reflected the different metabolic roles of tissues in responses to hypoxic stress in the euryoxic fish, *Gillichthys mirabilis* (Gracey et al., 2001). Similarly, physiological differences in response to hypoxia were

also existed among tissues (heart, brain and eye) of zebrafish (Marques et al., 2008; Roesner et al., 2006). Moreover, the mechanism in response to short and long-term hypoxia exposure may be different by altering expressions of different genes (van der Meer et al., 2005). Taken together, hypoxia response in fish is a complex process, involving activation and regulation of multiple genes and signaling pathways.

1.2.3 Genetic and genomic research in catfish

Catfish is the major species in the aquaculture industry of the United States. With the development of next-generation sequencing technology, much progress has been made in recent years for integrated genetic and genomic analysis. As a non-model species, lots of efforts have been made to construct the catfish reference genome. A bacterial artificial chromosome (BAC) contig-based physical map was constructed (Xu et al., 2007) using the CHORI 212 BAC library (Wang et al., 2007). The existing physical maps and related BAC resources are useful for validation of reference genome sequences. The Illumina platform has been mostly used for the generation of the whole genome sequence of channel catfish. Recently, our lab generated the production of the reference genome sequence and its annotation (Liu et al., 2016). The reference genome sequence assembly for catfish was of good quality. The catfish reference genome sequence was assessed to be nearly complete as 99.7% of re-sequencing reads were mapped to the reference genome sequence, and 50% of the catfish genome sequence is included in only 31 of the largest scaffolds (Liu et al., 2016). The number of complete genes annotated in the catfish reference genome sequence is larger than that of any of the sequenced diploid fish species, including zebrafish (Liu et al., 2016). The vast majority of the reference genome sequence (99.1%) has been anchored to chromosomes, and the positions of 253,744 genetically mapped SNPs were fully concordant with those on the reference genome sequence with 4 exceptions (Zeng et al., 2017).

Molecular markers were identified in catfish by various studies. A large numbers of microsatellites have been developed and identified in channel catfish (Serapion et al., 2004; Somridhivej et al., 2008). As microsatellites are highly polymorphic, it is very time and money consuming to use very large numbers of microsatellites. With the rapid development of the next-generation sequencing, it became accessible to generate a large number of single nucleotide polymorphisms (SNPs) with aquaculture species, both technically and economically. In addition to whole genome sequencing, SNPs can be identified through genome re-sequencing or RNA-Seq projects. Using transcriptome sequencing and analysis, over two million putative SNPs were identified from channel catfish and almost 2.5 million putative SNPs were identified from blue catfish (Liu et al., 2011). Two third of these putative SNPs were filtered SNPs linked to 16,562 unique genes in channel catfish and 17,423 unique genes in blue catfish. To cover the genomic regions, almost 8 million high-quality genome-wide SNPs were generated using next-generation sequencing with genomic DNA samples from multiple families (Sun et al., 2014). These chromosome-level genome sequence assembly and genome scale SNPs have provided a way for detail analysis of any production traits.

Genetic linkage maps with different types of markers have been developed for catfish, including the genetic linkage map with microsatellites (Waldbieser et al., 2001), the genetic linkage map with AFLP (Liu et al., 2003), the genetic linkage map with gene-associated markers (Kucuktas et al., 2009), the genetic linkage map with mixed markers (Ninwichian et al., 2012), the genetic linkage map with SNP markers (Li et al., 2015a). Furthermore, the catfish 250K SNP array, including 103,185 gene-associated SNPs and 146,928 anonymous genomic SNPs, have been mapped to the linkage map of catfish by using Affymetrix Axiom technology (Liu et al., 2014).

Large numbers of transcriptome resources exist for catfish. Nearly 500,000 quality

expressed sequence tags (ESTs) were generated from 8 channel catfish and 4 blue catfish libraries using unique inter-specific approach (Wang et al., 2010). Large RNA-Seq datasets have been generated to characterize different expression patterns in response to environmental stress or diseases in various tissues of catfish. For instance, RNA-seq analysis was conducted to reveal the mucosal immune responses to *Edwardsiella ictaluri* (ESC) infection, and genes involved in the centrality of actin cytoskeletal polymerization/remodeling and junctional regulation in pathogen entry and subsequent inflammatory responses (Li et al., 2012). Similar RNA-seq analysis was conducted to identify the immune response to columnaris bacterial infection, revealing that the NF-κB suppression and IFN stimulation in the catfish gill may play an important role in host defense system (Sun et al., 2012). Expression signatures of genes involved in oxygen transport, protein synthesis, folding, and degradation were also revealed by RNA-Seq in response to heat stress in catfish (Liu et al., 2013). Moreover, RNA-Seq data allow the identification of a numbers of gene families in catfish and their potential functions in response to environmental stress like hypoxia or diseases. For example, to understand the immune response to ESC or columnaris disease in catfish, numerous studies have been conducted to identify the characteristics of immune genes, including chemokines (Peatman et al., 2006), cytokines (Yao et al., 2015), claudins (Sun et al., 2015), serpins (Li et al., 2015b), complement regulatory protein genes (Jiang et al., 2015), phosphoinositide 3-kinase family (Li et al., 2016), apolipoprotein genes (Yang et al., 2017), and Rho GTPase genes (Tan et al., 2017). Although the available reference genome and annotation of channel catfish make it a lot easier to access these gene families, the system studies of them still provide insight into the evolutionary patterns and biological function.

The development of the high-resolution genetic linkage map (Li et al., 2015a) and the catfish 250K SNP array (Liu et al., 2014) provide the foundation of QTL analysis, GWAS, fine

mapping and whole genome-based selection in catfish. Several analyses have been conducted to identify genomic regions associated with performance and production traits in catfish (Geng et al., 2016; Geng et al., 2015; Zhou et al., 2016). For instance, QTL for columnaris disease resistance (Geng et al., 2015), for head size (Geng et al., 2016), and for ESC disease resistance (Zhou et al., 2016) have been identified using the channel catfish × blue catfish interspecific system. However, no QTL studies have been conducted with hypoxia tolerance in catfish.

1.2.4 Introduction of QTL studies

The rapid advances in genotyping technology allow the access to a large number of SNPs on a genomic scale, and the use of SNPs not limits in model species but applies to other non-model species. The application of SNPs in various species make it possible for scientists to investigate the genetic architecture of a certain trait or disease and identify potential associated QTL and causal genes. Several statistical methods are available to investigate the genomic regions associated with different traits. Among them, there are three major methods including QTL mapping, bulk segregant analysis (BSA), and GWAS.

QTL mapping is a traditional but powerful method used to identify loci co-segregating with a given trait. QTL mapping can be conducted with various DNA markers such as AFLP, microsatellites, as well as SNPs. This allow QTL mapping to be widely used in any species. QTL mapping requires two or more strains of organisms as the parental population with the different genetic background with regard to the trait of interest. Moreover, genetic markers which are different in the parental lines should segregate with the contrasted phenotype. There are some fundamental limitations for QTL mapping: (1) as only family samples are used, the mapping resolution of QTL mapping is limited by the amount of recombination events within the pedigrees,

although this can be improved by several generations of intercrossing (Darvasi and Soller, 1995); (2) if the studied families contain no segregating alleles at some loci, these loci can never be detected leading to uncomprehensive discoveries; (3) QTL mapping has less power to identify common genetic variants with modest effects (Risch and Merikangas, 1996); (4) important quantitative traits usually not only have complex genetic architectures like epistatic interactions between QTL, but also genetic-environmental interactions such as genotype-by-sex, genotype-by-environment (Wang et al., 2005), whereas not all QTL studies were designed to detect such interactions (Mackay, 2001b); (5) the allele frequencies and combinations present in the sampled families may differ from those in the other populations, resulting the limited application of the identified QTL to the whole population (Korte and Farlow, 2013).

QTL mapping has been conducted for over 20 aquaculture species. QTL analysis for growth traits has been conducted in Asian seabass (Wang et al., 2006), rainbow trout (Wringe et al., 2010), tilapia (Cnaani et al., 2003), and catfish (unpublished data). QTL associated with sex determination have also been identified in tilapia (Eshel et al., 2012), salmon (Davidson et al., 2009), rainbow trout (Alfaqih et al., 2009), and catfish (unpublished data). For instance, a major QTL for sex determination was reported to locate between SSR markers ARO172 and ARO177 on LG 23 in tilapia (Eshel et al., 2012). Disease resistance is one of the most studied traits using QTL mapping as it is a major bottleneck in aquaculture. For example, infectious pancreatic necrosis (IPN) resistance was linked to three DNA markers on LG 21 in Atlantic salmon (Gheyas et al., 2010). In Japanese flounder, a single QTL was identified to control the resistance to lymphocystis disease (Fuji et al., 2006). Some QTL studies were also conducted with other performance traits like meat quality traits (Baranski et al., 2010; Derayat et al., 2007), salinity tolerance (Norman et al., 2011; Rengmark et al., 2007) and temperature tolerance (Cnaani et al., 2003).

Bulk segregant analysis (BSA) is another way to identify potential QTL for a certain trait. The basic idea of BSA is that phenotypic extremes should have dramatic differences in the loci associated with the trait when samples selected from phenotypic extremes and their genotypes are analyzed in bulk. Studies have been conducted to identify potential genomic regions associated with columnaris and ESC diseases in catfish using BSA (Peatman et al., 2013; Wang et al., 2013b). Although it is difficult to detect the associated loci when only comparing individuals with different phenotype performance, the comparison of pooled samples (bulk) with the phenotypic extremes increases the power to reveal the contrast in the genotype (Michelmore et al., 1991; Wang et al., 2013a). For example, when samples are grouped according to the contrasted trait, the frequency of the two marker alleles present within each of the two bulks should deviate significantly from the expected ratio in their specific population, resulting the identification of the correlation between genotype and phenotype (Quarrie et al., 1999). Some disadvantages limit the use of BSA include: (1) the major drawback is the imprecision caused by the genotype generated from the pooled; (2) if the family stratification exists, it is impossible to eliminate this effect due to the bulk analysis; (3) BSA can only detect the genetic effects of single locus, precluding any analysis of haplotype or gene-by-gene interaction effects. However, it is still broadly used because of its high efficiency, low cost, and analytical simplicity.

GWAS, also known as linkage disequilibrium (LD) mapping, allows the detection of QTL based on the strength of the correlation between mapped markers and the trait of interest. Similar to QTL mapping, GWAS is also a statistical method that links phenotypic data and genotypic data to explain the genetic basis that causes phenotypic variations. Unlike QTL mapping, which is well-suited for family-based samples, GWAS normally uses unrelated population samples and can potentially offer higher mapping resolution than QTL mapping by including all historic

recombination events among founders (Mackay, 2001a). In the natural populations that are utilized by GWAS, LD often decays more rapidly with increasing physical distances than in controlled crosses (Mackay and Powell, 2007).

GWAS has been a powerful tool for investigating the genetic architecture of important traits in human, crop, and some animals. For example, a number of GWA studies have been conducted to identify QTL associated with various human diseases like renal cell carcinoma (Su et al., 2013), allergic rhinitis (Andiappan et al., 2011), diabetes (Billings and Florez, 2010), and breast cancer (Fachal and Dunning, 2015). QTL associated with some other traits like human height, metabolic traits were also identified by GWAS (Sabatti et al., 2009; Yang et al., 2010). Many GWA studies have been conducted to reveal potential QTL for various plant traits such as grain morphology, root architecture, plant height, disease resistance (Biscarini et al., 2016; Huang and Han, 2014; Sun et al., 2016). Some GWA studies also successfully identified QTL associated with the interesting traits in mammal animals. For instance, Schoenebeck et al. (2012) found variation of BMP3 contributes to dog breed skull diversity in a GWA study. In Japanese Black cattle, a GWA study revealed three major QTL for carcass weight (Nishimura et al., 2012). GWAS in *Setaria italic* found that 916 varieties were phenotyped under five different environments and 512 loci were identified to be associated with 47 agronomic traits (Jia et al., 2013). However, GWAS has been barely utilized in aquaculture due to the insufficiency of genetic and genomic information. To our best knowledge, several GWA studies have been conducted in catfish to reveal the genetic architecture of head size (Geng et al., 2016), disease resistance (Geng et al., 2015; Zhou et al., 2017), and growth traits (unpublished data). The application of GWAS in aquaculture species will undoubtedly help connecting sequence diversity with phenotypic differences. With the advance of next-generation sequencing technologies, it is certain that more and more GWA

studies will be conducted with aquaculture species.

1.2.5 Methodology of GWAS

Over the last 10 years, GWAS has been widely applied in a number of organisms contributed by the development of high-density SNP arrays and DNA re-sequencing. It presents a powerful tool to reconnect the study trait back to its underlying genetics. In this section, a general introduction of the methodology of GWAS is reviewed concerning the special characteristics of aquaculture species.

The power of GWAS is the probability that a true association between a marker and the trait of interest is found to be significant by the designed study. It depends on many factors including the significance level α set by the experimenter, design of experiment, association test, effect size of each QTL, the allele frequency of the causal allele, the LD between the causal allele and the genotyped markers on the array, and sample size (Hayes, 2013). To achieve a higher statistical power which allows the identification of more associated QTL is the goal when we design a GWA study. The most obvious method to allow the higher power is to increase the sample size used in GWAS. Some software are available to calculate the sample size for unrelated individuals to ensure sufficient power (Gauderman and Morrison, 2006). Another way to increase the power is the use of selective genotyping, which genotypes individuals with the extremes of the phenotype (Lynch and Walsh, 1998). It requires less sample size but keeps the high power to detect QTL (Van Gestel et al., 2000). However, the tradeoff is that it may cause the potential overestimation of effect size. Moreover, including a statistician during planning phase is also recommended to ensure a solid and powerful design.

The ideal samples for GWAS should be homogenous in genetic background, highly contrasted in phenotype, and highly intercrossed to provide high mapping resolution, but no population stratification. Population stratification is generated from the different allele frequencies among subpopulations, especially when phenotypic variation exists among different subpopulations, and it always confounds the association test in the practical situations, leading to false positive results. For example, if the sample population contain two families with different resistance to one disease, it is then possible that the identified "associated" loci may be more associated with the family difference than with disease resistance. To eliminate the effect of population stratification, choose of the population samples should be more careful at the GWAS design stage.

Usually two types of population are widely used in GWAS. One sample type is the sample from non-manipulated natural populations with known phenotype, because it is easy to assume natural populations are unrelated. Using this kind of samples is cost- and time-effective, as there is no need for the breeding process with large number of founders to obtain higher generations. However, there are two major drawbacks for this sample type: (1) sometimes it is very difficult to collect the phenotype information for samples from natural population; (2) although samples from natural population can be assumed no population stratification existed, it is still possible, especially for aquaculture species, that subpopulations exist in a natural population as a large number of individuals could be derived from a very limited number of founders.

The other sample type is the family-based population sample. Family-based population design is more immune to population stratification because of the known founder information. It is unique for aquaculture species to use family-based samples in GWAS, because that: (1) many aquaculture species have high fecundities with thousands of progenies per spawn, saving

tremendous effort for reproduction when compared with livestocks; (2) even for samples with more than one full-sibling families, which happens in most practical experiments, the clear pedigree information of family-based population design makes the identification of population stratification much easier than the natural population design. For example, the interspecific hybrid catfish from mating female channel catfish with male blue catfish serves as a good model for detection of major QTL associated with columnaris disease resistance, because channel catfish is generally resistant to the disease while blue catfish is generally susceptible (Geng et al., 2015). Moreover, the higher generations of intercross hybrids produced by intermating F2 individuals for several generations can provide a higher resolution for the association mapping (Mott et al., 2000). This is because that the shorter haplotype blocks surrounding the gene of interest in the higher generations allow the identification of the casual genes within a small genomic region. Although many benefits of using family-based population sample, there are still some disadvantages limited the application of this sample type: (1) family-based population design has less power to detect the causal alleles that are homozygous in the subpopulation used in the association test but heterozygous in the whole population due to the limited number of founders used to generate the samples; (2) compared with the natural population which has more rounds of historical recombination, the limited numbers of recombination events in family-based population design may lead to low mapping resolution; (3) the generation of family-based population can take a long time when compared with the natural population, especially if the higher generations are needed; (4) the between-family stratification in both phenotype and genotype can be problematic and requires complex statistical methods to address.

After genotype calling procedure based on signal intensities generated by the SNP array for the alleles (Ziegler et al., 2008), quality control (QC) should be performed to avoid false results,

which includes sample-level QC and SNP-level QC. For sample-level QC, not only samples with low genotyping quality or a low call rate should be excluded from analysis. After visualizing sample structures by using principal component analysis (PCA) or cluster analysis based on identity by state (IBS) kinship matrix, the outliers with different ancestry should be removed (Gudbjartsson et al.). For SNP-level QC, SNPs with any Mendelian inheritance errors, low calling rate or low minor allele frequencies (MAF) should also be excluded. Moreover, quantile–quantile plots, which compare the obtained test statistics against what is expected under the null hypothesis of no association, are regularly used in SNP quality assessment. If unrelated samples are used, the SNPs severely out of Hardy–Weinberg equilibrium should be flagged, because disequilibrium can result from a true association, a potential genotyping error, or population stratification (Turner et al., 2011).

There are two types of study design based on the phenotypes used in GWAS: (1) qualitative trait design (case-control design), in this case, specified number of affected (case) and unaffected (control) individuals are genotyped for the trait of interest (Kang et al., 2010). For example, the disease resistance trait can be classified as "resistant group" and "susceptible group", which is a binary variable (Geng et al., 2015); (2) quantitative trait design, in this case, the trait need to have well-established quantitative measures. For example, the hypoxia tolerance trait was measured as the resistant time to hypoxic stress (Wang et al., 2017). One advantage to use quantitative trait design is that genetic effect size (the proportion of phenotypic variance explained by two alleles at a locus) can be easily calculated with quantitative traits, since the quantitative traits are measured by continuous numbers. For both study designs, to avoid potential confounding, the confounding factors of the trait are needed to adjust, including gender, age, body weight, experimental batch,

and known family structure (Dominik, 2013). Therefore, a correct and comprehensive adjustment for the phenotype of a trait is the key for the subsequent analysis.

The association analysis depends crucially on the study design. In the simplest case, where no population stratification exists, familiar methods such as logistic regression, χ^2 tests of association, and odds ratios may be suitable. However, population stratification almost always exists within the sample population, therefore correction of population stratification is a key step in the association analysis. When dealing with samples with population stratification, more complex strategies will be needed to detect population stratification and to infer genetic ancestry. A number of statistical strategies are established to detect population stratification, of which those commonly used in GWAS include the following: (1) genomic control, which estimates the degree of inflation of the test statistic under the assumption that existing population structure has a uniform influence throughout the entire genome (Devlin and Roeder, 1999); (2) PCA, which infers informative axes of genetic variation and locates each sample on this high-dimensional map of continuous genomic variation to reveal the genetic relationship of each sample (Price et al., 2006). Among PCA-based software packages that have been proposed, EIGENSTRAT is the most widely used (Price et al., 2006).

Mixed model can be used to correct population structure, family structure and cryptic relatedness (Yu et al., 2006). The general idea for the mixed model is to model phenotypes using a mixture of fixed effects and random effects. Fixed effects include the candidate SNP and optional covariates, such as gender, age, body weight or height; whereas random effects are based on a phenotypic covariance matrix, which summarize heritable and nonheritable random variation of the sample population (Yu et al., 2006). One important question is whether population structure should be modelled as part of the set of random effects, or as a separate fixed effect? Actually,

population structure is a fixed effect as its effect as a function of genetic ancestry is the same for all samples (Price et al., 2010). Modelling population structure as a fixed effect provides a higher level of certainty in correcting for stratification, however, additional model parameters by running PCA (or a similar method) are required to infer the genetic ancestry of each sample (Zhang et al., 2010).

There are two major types of mixed models: simple mixed models and linear mixed models. Simple linear models represent the phenotype Y as a function of fixed effects X:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{b} + \boldsymbol{\mathcal{E}}$$

where $\mathbf{X}$ denotes the genotype at the candidate marker in addition to optional covariates, such as gender or age, $\boldsymbol{b}$ denotes coefficients of fixed effects and $\boldsymbol{\mathcal{E}}$ is a normally distributed noise term that accounts for unexplained variation in Y.

To address population stratification, we can include principal component covariates from PCA in X to explicitly model the ancestry of each individual. However, PCA cannot account for relatedness between individuals, which may also cause inflation in test statistics.

Linear mixed models represent the phenotype Y as a function of fixed effects X plus random effects Z:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{u} + \boldsymbol{\mathcal{E}}$$

$$\mathbf{Var(u)} = \mathbf{K}\sigma_g^2$$

where $\mathbf{X}$ denotes a matrix of fixed effects including the genotype at the candidate marker in addition to optional covariates, such as gender or age; $\boldsymbol{\beta}$ denotes a coefficient vector of fixed effects; $\mathbf{Z}$ denotes a matrix of random additive genetic effects, $\boldsymbol{u}$ is a vector representing the coefficients of the random effect according to a kinship matrix $\mathbf{K}$ (σ_g^2 represent the parameter for

additive genetic variance). Thus, **Z** represents the heritable component of random variation and $\boldsymbol{\varepsilon}$ represents the non-heritable component of random variation.

Mixed models are used to be very computationally intensive. However, recent computational advances (such as the EMMAX and TASSEL software) have now made it possible for application (Kang et al., 2010; Zhang et al., 2010). An efficient mixed-model association method, Efficient Mixed-Model Association eXpedited (EMMAX), is one of the mostly applied method in GWAS. One of the key characters of EMMAX is that it includes a procedure to estimate the contribution of the kinship matrix to the phenotypes (Balding and Nichols, 1995). Unlike a traditional variance component model which uses identity by descent (IBD) coefficients estimated from the pedigree (Ober et al., 2001), EMMAX used IBS to estimate the genetic relatedness between the individuals. With structured model organism samples, using a haplotype-based IBS matrix or a simple IBS matrix is more robust for correcting the population structure resulting in a lower inflation factor than using the estimated IBD matrix (Kang et al., 2008; Zhao et al., 2007). Although EMMAX is designed for quantitative traits, the association test for qualitative traits can be approximately conducted using 0-1 quantitative response variable to represent the case-control status (Kang et al., 2010). Despite the advantage that mixed models could help eliminate false positives caused by population stratification, it is not guaranteed to correct all possible confounding population structures.

For family-based studies, in which individuals are ascertained from family pedigrees, transmission disequilibrium test (TDT) instead of mixed models can be used for correction of population stratification (Laird and Lange, 2006). When conducting TDT, the progenies in each family with a certain extreme phenotype of interest are selected, and genotyped along with the parents. The loci where parents are heterozygous will contribute to the association analysis: from

each parent, one allele must be transmitted to the progeny and the other one not; then the ratio of transmission to non-transmission over all families will be compared with the expected value of 1:1 (Mackay and Powell, 2007). TDT that focus on within-family information is immune to population stratification, as transmitted and untransmitted alleles have the same genetic ancestry, and such tests can be performed using the FBAT and QTDT software (Spielman et al., 1993). However, TDT is extremely susceptible to genotype errors in parents.

Setting a strict significance level for GWAS is as important as choosing statistic model for the association tests. A strict significance level can limit spurious false positive results found by chance in GWAS. The null hypothesis of a GAW study is that there is no association between a certain marker and the trait of interest. When P-value falls below a predetermined α value (significance level), which is usually 0.05 for single marker test, the null hypothesis will be rejected. Type I error may occur, which means the null hypothesis is falsely rejected with a probability of 5% when it is true in fact, leading a 5% chance of false positive results. However, when a multiple test is conducted in GWAS, hundreds of thousands of SNPs are tested for their association with the trait simultaneously, leading to the increase of cumulative likelihood of false positive results. In this case, there are two commonly used method to control the false positive results including: (1) Bonferroni correction, proposed by Duggal et al. (2008), in which the threshold of P-value (α value) for genome-wide significance can be calculated by converting α=0.05 to α=0.05/n, where n equals with the estimated number of independent markers and LD blocks (Duggal et al., 2008); (2) false discovery rate (FDR) correction, introduced by Hochberg and Benjamini (1990), in which all p-values are firstly ranked from the smallest to the largest, and then each p-value is adjusted according to its rank and the number of tests (Hochberg and Benjamini, 1990). FDR correction tolerates more false positive results when compared with the Bonferroni correction. For example,

a FDR=0.05 allows 5% of reported positives are false positives, while the Bonferroni correction α=0.05 only requires the whole family of positives to be true positives with the certainty of 95% (Qu et al., 2010).

Two types of P value plots have emerged as the standard presentation of results of GWAS: (1) −log10(P value) genome-wide association plots (Manhattan plots), which represent the P values of the entire GWAS on a genomic scale. All P values are represented in genomic order by chromosome/linkage group and position on the chromosome/linkage group (x-axis). The value on the y-axis represents the −log10 of the P value; and (2) quantile-quantile (QQ) plots, which graphically represent the deviation of the observed P values from the null hypothesis. The observed P values for each SNP are sorted from largest to smallest and plotted against expected values from a theoretical χ^2-distribution. In the typical QQ plots, the majority of the points are on or near the middle line between the x-axis and the y-axis as most of the observed values should be corresponded to the expected values; whereas some points may move towards the y-axis, indicating that the observed P values are more significant than expected under the null hypothesis. However, if an early separation of the observed from the expected occurs, this means that many moderately significant P values are more significant than expected under the null hypothesis. This result may indicate that population stratification existed in the sample population lead to this false positive result (Ehret, 2010).

GWAS has contributed to the identification of genomic regions or causal genes associated with various diseases in human and economic traits in crop and animals. These findings not only deep the understanding of genetic architecture of a certain trait, but also provide insight into the effects of genetic variants on the biological process and function. Although there are only a limited number of GWA studies in aquaculture species, the rapid development of genotyping technologies,

especially high-density SNP arrays, will accelerate the application of GWAS in aquaculture species.

1.3 Rationale and significance

The catfish industry provides employment opportunities for tens of thousands of producers, processors, service providers, marketers, retailers, and restaurant owners. It has a huge impact on the society of the United State. To meet the challenge from international import, the improvement of the breeding program to obtain better catfish breeding is required to reduce economic costs and increase production efficiency.

Hypoxia has always been a major concern in aquaculture. To reduce the economic loss caused by hypoxia, studies need to be conducted for better understanding of the mechanism of hypoxia responses. This dissertation is the first comprehensive study in both environmental and genetic effectors of hypoxia tolerance in catfish. In this dissertation, effects of both genetic and environmental factors on hypoxia tolerance were identified in channel catfish. It will provide insight into the genetic architecture of hypoxia tolerance, and set the foundation for future genetic improvement of hypoxia tolerance in catfish.

References

Aboagye, D.L., Allen, P.J., 2014. Metabolic and locomotor responses of juvenile paddlefish *Polyodon spathula* to hypoxia and temperature. Comparative biochemistry and physiology. A: comparative physiology 169, 51-59.

Alfaqih, M.A., Brunelli, J.P., Drew, R.E., Thorgaard, G.H., 2009. Mapping of five candidate sex-determining loci in rainbow trout (*Oncorhynchus mykiss*). BMC genetics 10, 2.

Almeida-Val, V.M., Val, A.L., Duncan, W.P., Souza, F.C., Paula-Silva, M.N., Land, S., 2000. Scaling effects on hypoxia tolerance in the Amazon fish *Astronotus ocellatus* (Perciformes: Cichlidae): contribution of tissue enzyme levels. Comparative biochemistry and physiology. B: biochemistry & molecular biology 125, 219-226.

Andiappan, A.K., Wang, D.Y., Anantharaman, R., Parate, P.N., Suri, B.K., Low, H.Q., Li, Y., Zhao, W., Castagnoli, P., Liu, J.J., Chew, F.T., 2011. Genome-wide association study for atopy and allergic rhinitis in a singapore chinese population. PloS one 6, e19719.

Anttila, K., Dhillon, R.S., Boulding, E.G., Farrell, A.P., Glebe, B.D., Elliott, J.A.K., Wolters, W.R., Schulte, P.M., 2013. Variation in temperature tolerance among families of Atlantic salmon (*Salmo salar*) is associated with hypoxia tolerance, ventricle size and myoglobin level. Journal of experimental biology 216, 1183-1190.

Balding, D.J., Nichols, R.A., 1995. A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity, Human identification: the use of DNA markers. Springer, pp. 3-12.

Baranski, M., Moen, T., Vage, D.I., 2010. Mapping of quantitative trait loci for flesh colour and growth traits in Atlantic salmon (*Salmo salar*). Genetics selection evolution 42, 17.

Beall, C.M., 2006. Andean, Tibetan, and Ethiopian patterns of adaptation to high-altitude hypoxia. Integrative and comparative biology 46, 18-24.

Billings, L.K., Florez, J.C., 2010. The genetics of type 2 diabetes: what have we learned from GWAS? Annals of the New York academy of sciences 1212, 59-77.

Biscarini, F., Cozzi, P., Casella, L., Riccardi, P., Vattari, A., Orasen, G., Perrini, R., Tacconi, G., Tondelli, A., Biselli, C., Cattivelli, L., Spindel, J., McCouch, S., Abbruscato, P., Vale, G., Piffanelli, P., Greco, R., 2016. Genome-wide association study for traits related to plant and grain morphology, and root architecture in temperate rice accessions. PloS one 11, e0155425.

Burggren, W.W., Cameron, J.N., 1980. Anaerobic metabolism, gas-exchange, and acid-base-balance during hypoxic exposure in the channel catfish, *Ictalurus punctatus*. Journal of experimental zoology 213, 405-416.

Burleson, M., Wilhelm, D., Smatresk, N., 2001. The influence of fish size size on the avoidance of hypoxia and oxygen selection by largemouth bass. Journal of fish biology 59, 1336-1349.

Cech Jr, J.J., Massingill, M.J., Vondracek, B., Linden, A.L., 1985. Respiratory metabolism of mosquitofish, *Gambusia affinis*: effects of temperature, dissolved oxygen, and sex difference. Environmental biology of fishes 13, 297-307.

Cnaani, A., Hallerman, E.M., Ron, M., Weller, J.I., Indelman, M., Kashi, Y., Gall, G.A.E., Hulata, G., 2003. Detection of a chromosomal region with two quantitative trait loci, affecting cold tolerance and fish size, in an F-2 tilapia hybrid. Aquaculture 223, 117-128.

Darvasi, A., Soller, M., 1995. Advanced intercross lines, an experimental population for fine genetic mapping. Genetics 141, 1199-1207.

Davidson, W.S., Huang, T.K., Fujiki, K., von Schalburg, K.R., Koop, B.F., 2009. The sex determining loci and sex chromosomes in the family salmonidae. Sex development 3, 78-87.

Derayat, A., Houston, R.D., Guy, D.R., Hamilton, A., Ralph, J., Spreckley, N., Taggart, J.B., McAndrew, B.J., Haley, C.S., 2007. Mapping QTL affecting body lipid percentage in Atlantic salmon (*Salmo salar*). Aquaculture 272, 250-251.

Devlin, B., Roeder, K., 1999. Genomic control for association studies. Biometrics 55, 997-1004.

Dominik, S., 2013. Descriptive statistics of data: understanding the data set and phenotypes of interest, genome-wide association studies and genomic prediction. Springer, pp. 19-35.

Duggal, P., Gillanders, E.M., Holmes, T.N., Bailey-Wilson, J.E., 2008. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. BMC Genomics 9, 516.

Dunham, R.A., Ramboux, A.C.R., Perera, D.A., 2014. Effect of strain on tolerance of low dissolved oxygen of channel X blue catfish hybrids. Aquaculture 420, 25-28.

Ehret, G.B., 2010. Genome-wide association studies: contribution of genomics to understanding blood pressure and essential hypertension. Current hypertension reports 12, 17-25.

Eshel, O., Shirak, A., Weller, J.I., Hulata, G., Ron, M., 2012. Linkage and physical mapping of sex region on LG23 of Nile Tilapia (*Oreochromis niloticus*). G3: genes, genomes, genetics 2, 35-42.

Fachal, L., Dunning, A.M., 2015. From candidate gene studies to GWAS and post-GWAS analyses in breast cancer. Current opinion in genetics and development 30, 32-41.

Faust, H.A., Gamperl, A.K., Rodnick, K.J., 2004. All rainbow trout (*Oncorhynchus mykiss*) are not created equal: intra-specific variation in cardiac hypoxia tolerance. Journal of experimental biology 207, 1005-1015.

Fuji, K., Kobayashi, K., Hasegawa, O., Coimbra, M.R.M., Sakamoto, T., Okamoto, N., 2006. Identification of a single major genetic locus controlling the resistance to lymphocystis disease in Japanese flounder (*Paralichthys olivaceus*). Aquaculture 254, 203-210.

Gauderman, W., Morrison, J., 2006. QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies. Available at http://hydra.usc.edu/gxe/.

Geng, X., Feng, J.B., Liu, S.K., Wang, Y.P., Arias, C., Liu, Z.J., 2014. Transcriptional regulation of hypoxia inducible factors alpha (HIF-alpha) and their inhibiting factor (FIH-1) of channel catfish (*Ictalurus punctatus*) under hypoxia. Comparative biochemistry and physiology. B: biochemistry & molecular biology 169, 38-50.

Geng, X., Liu, S.K., Yao, J., Bao, L.S., Zhang, J.R., Li, C., Wang, R.J., Sha, J., Zeng, P., Zhi, D.G., Liu, Z.J., 2016. A genome-wide association study identifies multiple regions associated with head size in catfish. G3: genes, genomes, genetics 6, 3389-3398.

Geng, X., Sha, J., Liu, S.K., Bao, L.S., Zhang, J.R., Wang, R.J., Yao, J., Li, C., Feng, J.B., Sun, F.Y., Sun, L.Y., Jiang, C., Zhang, Y., Chen, A.L., Dunham, R., Zhi, D.G., Liu, Z.J., 2015. A genome-wide association study in catfish reveals the presence of functional hubs of related genes within QTL for columnaris disease resistance. BMC Genomics 16, 196.

Gheyas, A.A., Haley, C.S., Guy, D.R., Hamilton, A., Tinch, A.E., Mota-Velasco, J.C., Woolliams, J.A., 2010. Effect of a major QTL affecting IPN resistance on production traits in Atlantic salmon. Animal genetics 41, 666-668.

Gracey, A.Y., Troll, J.V., Somero, G.N., 2001. Hypoxia-induced gene expression profiling in the euryoxic fish *Gillichthys mirabilis*. Proceedings of the national academy of sciences 98, 1993-1998.

Gradin, K., McGuire, J., Wenger, R.H., Kvietikova, I., Whitelaw, M.L., Toftgard, R., Tora, L., Gassmann, M., Poellinger, L., 1996. Functional interference between hypoxia and dioxin signal transduction pathways: competition for recruitment of the Arnt transcription factor. Molecular and cellular biology 16, 5221-5231.

Guan, B., Ma, H., Wang, Y.P., Hu, Y.L., Lin, Z.P., Zhu, Z.Y., Hu, W., 2011. Vitreoscilla hemoglobin (VHD) overexpression increases hypoxia tolerance in zebrafish (*Danio rerio*). Marine biotechnology 13, 336-344.

Gudbjartsson, D.F., Walters, G.B., Thorleifsson, G., Stefansson, H., Halldorsson, B.V., Zusmanovich, P., Sulem, P., Thorlacius, S., Gylfason, A., Steinberg, S., 2008. Many sequence variants affecting diversity of adult human height. Nature genetics 40, 609-615.

Hayes, B., 2013. Overview of statistical methods for genome-wide association studies (GWAS), genome-wide association studies and genomic prediction. Springer, pp. 149-169.

Hochachka, P.W., 1986. Defense strategies against hypoxia and hypothermia. Science 231, 234-241.

Hochberg, Y., Benjamini, Y., 1990. More powerful procedures for multiple significance testing. Statistics in medicine 9, 811-818.

Hong, S.-S., Lee, H., Kim, K.-W., 2004. HIF-1α: a valid therapeutic target for tumor therapy. Cancer research and treatment 36, 343-353.

Hu, C.J., Wang, L.Y., Chodosh, L.A., Keith, B., Simon, M.C., 2003. Differential roles of hypoxia-inducible factor 1 alpha (HIF-1 alpha) and HIF-2 alpha in hypoxic gene regulation. Molecular cellular and biology 23, 9361-9374.

Huang, X.H., Han, B., 2014. Natural variations and genome-wide association studies in crop plants. Annual review of plant biology 65, 531-551.

Jensen, F.B., Nikinmaa, M., Weber, R.E., 1993. Environmental perturbations of oxygen transport in teleost fishes: causes, consequences and compensations, Fish ecophysiology. Springer, pp. 161-179.

Jha, A.R., Zhou, D., Brown, C.D., Kreitman, M., Haddad, G.G., White, K.P., 2016. Shared genetic signals of hypoxia adaptation in drosophila and in high-altitude human populations. Molecular biology and evolution 33, 501-517.

Jia, G., Huang, X., Zhi, H., Zhao, Y., Zhao, Q., Li, W., Chai, Y., Yang, L., Liu, K., Lu, H., 2013. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). Nature genetics 45, 957-961.

Jiang, C., Zhang, J.R., Yao, J., Liu, S.K., Li, Y., Song, L., Li, C., Wang, X.Z., Liu, Z.J., 2015. Complement regulatory protein genes in channel catfish and their involvement in disease defense response. Developmental & comparative immunology 53, 33-41.

Jung, H.J., Kim, J.H., Shim, J.S., Kwon, H.J., 2010. A novel Ca2+/Calmodulin antagonist HBC inhibits angiogenesis and down-regulates hypoxia-inducible factor. The journal of biological chemistry 285, 25867-25874.

Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.Y., Freimer, N.B., Sabatti, C., Eskin, E., 2010. Variance component model to account for sample structure in genome-wide association studies. Nature genetics 42, 348-354.

Kang, H.M., Zaitlen, N.A., Wade, C.M., Kirby, A., Heckerman, D., Daly, M.J., Eskin, E., 2008. Efficient control of population structure in model organism association mapping. Genetics 178, 1709-1723.

Keith, B., Johnson, R.S., Simon, M.C., 2012. HIF1 alpha and HIF2 alpha: sibling rivalry in hypoxic tumour growth and progression. Nature reviews cancer 12, 9-22.

Kietzmann, T., Mennerich, D., Dimova, E.Y., 2016. Hypoxia-inducible factors (HIFs) and phosphorylation: impact on stability, localization, and transactivity. Frontiers in cell and developmental biology 4, 11.

Korte, A., Farlow, A., 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant methods 9, 29.

Kramer, D.L., McClure, M., 1982. Aquatic surface respiration, a widespread adaptation to hypoxia in tropical freshwater fishes. Environmental biology of fishes 7, 47-55.

Kucuktas, H., Wang, S.L., Li, P., He, C.B., Xu, P., Sha, Z.X., Liu, H., Jiang, Y.L., Baoprasertkul, P., Somridhivej, B., Wang, Y.P., Abernathy, J., Guo, X.M., Liu, L., Muir, W., Liu, Z.J., 2009. Construction of genetic linkage maps and comparative genome analysis of catfish using gene-associated markers. Genetics 181, 1649-1660.

Kvamme, B.O., Gadan, K., Finne-Fridell, F., Niklasson, L., Sundh, H., Sundell, K., Taranger, G.L., Evensen, O., 2013. Modulation of innate immune responses in Atlantic salmon by chronic hypoxia-induced stress. Fish and shellfish immunology 34, 55-65.

Laird, N.M., Lange, C., 2006. Family-based designs in the age of large-scale gene-association studies. Nature reviews genetics 7, 385-394.

Li, C., Zhang, Y., Wang, R.J., Lu, J.G., Nandi, S., Mohanty, S., Terhune, J., Liu, Z.J., Peatman, E., 2012. RNA-seq analysis of mucosal immune responses reveals signatures of intestinal barrier disruption and pathogen entry following Edwardsiella ictaluri infection in channel catfish, *Ictalurus punctatus*. Fish and shellfish immunology 32, 816-827.

Li, H.B., Vaillancourt, R., Mendham, N., Zhou, M.X., 2008. Comparative mapping of quantitative trait loci associated with waterlogging tolerance in barley (*Hordeum vulgare L.*). BMC genomics 9, 1.

Li, Y., Liu, S., Qin, Z., Waldbieser, G., Wang, R., Sun, L., Bao, L., Danzmann, R.G., Dunham, R., Liu, Z., 2015a. Construction of a high-density, high-resolution genetic map and its integration with BAC-based physical map in channel catfish. DNA research 22, 39-52.

Li, Y., Liu, S.K., Qin, Z.K., Yao, J., Jiang, C., Song, L., Dunham, R., Liu, Z.J., 2015b. The serpin superfamily in channel catfish: Identification, phylogenetic analysis and expression profiling in mucosal tissues after bacterial infections. Developmental & comparative immunology 49, 267-277.

Li, Z., Yao, J., Xie, Y., Geng, X., Liu, Z., 2016. Phosphoinositide 3-kinase family in channel catfish and their regulated expression after bacterial infection. Fish and shellfish immunology 49, 364-373.

Liu, L.P., Simon, M.C., 2004. Regulation of transcription and translation by hypoxia. Cancer biology and therapy 3, 492-497.

Liu, S., Sun, L., Li, Y., Sun, F., Jiang, Y., Zhang, Y., Zhang, J., Feng, J., Kaltenboeck, L., Kucuktas, H., 2014. Development of the catfish 250K SNP array for genome-wide association studies. BMC research notes 7, 135.

Liu, S.K., Wang, X.L., Sun, F.Y., Zhang, J.R., Feng, J.B., Liu, H., Rajendran, K.V., Sun, L.Y., Zhang, Y., Jiang, Y.L., Peatman, E., Kaltenboeck, L., Kucuktas, H., Liu, Z.J., 2013. RNA-Seq reveals expression signatures of genes involved in oxygen transport, protein synthesis, folding, and degradation in response to heat stress in catfish. Physiological genomics 45, 462-476.

Liu, S.K., Zhou, Z.C., Lu, J.G., Sun, F.Y., Wang, S.L., Liu, H., Jiang, Y.L., Kucuktas, H., Kaltenboeck, L., Peatman, E., Liu, Z.J., 2011. Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array. BMC genomics 12, 53.

Liu, Z.J., Karsi, A., Li, P., Cao, D.F., Dunham, R., 2003. An AFLP-based genetic linkage map of channel catfish (*Ictalurus punctatus*) constructed by using an interspecific hybrid resource family. Genetics 165, 687-694.

Liu, Z.J., Liu, S.K., Yao, J., Bao, L.S., Zhang, J.R., Li, Y., Jiang, C., Sun, L.Y., Wang, R.J., Zhang, Y., Zhou, T., Zeng, Q.F., Fu, Q., Gao, S., Li, N., Koren, S., Jiang, Y.L., Zimin, A., Xu, P., Phillippy, A.M., Geng, X., Song, L., Sun, F.Y., Li, C., Wang, X.Z., Chen, A.L., Jin, Y.L., Yuan, Z.H., Yang, Y.J., Tan, S.X., Peatman, E., Lu, J.G., Qin, Z.K., Dunham, R., Li, Z.X., Sonstegard, T., Feng, J.B., Danzmann, R.G., Schroeder, S., Scheffler, B., Duke, M.V., Ballard, L., Kucuktas, H., Kaltenboeck, L., Liu, H.X., Armbruster, J., Xie, Y.J., Kirby, M.L., Tian, Y., Flanagan, M.E., Mu, W.J., Waldbieser, G.C., 2016. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. Nature communications 7, 11757.

Lynch, M., Walsh, B., 1998. Genetics and analysis of quantitative traits. Sinauer Sunderland.

Mackay, I., Powell, W., 2007. Methods for linkage disequilibrium mapping in crops. Trends in plant science 12, 57-63.

Mackay, T.F., 2001a. The genetic architecture of quantitative traits. Annual review of genetics 35, 303-339.

Mackay, T.F., 2001b. Quantitative trait loci in *Drosophila*. Nature reviews genetics 2, 11-20.

Makino, Y., Kanopka, A., Wilson, W.J., Tanaka, H., Poellinger, L., 2002. Inhibitory PAS domain protein (IPAS) is a hypoxia-inducible splicing variant of the hypoxia-inducible factor-3 alpha locus. The journal of biological chemistry 277, 32405-32408.

Mandic, M., Todgham, A.E. and Richards, J.G., 2009. Mechanisms and evolution of hypoxia tolerance in fish. Proceedings of the Royal Society of London B: biological sciences, 276, 735-744.

Marques, I.J., Leito, J.T.D., Spaink, H.P., Testerink, J., Jaspers, R.T., Witte, F., van Den Berg, S., Bagowski, C.P., 2008. Transcriptome analysis of the response to chronic constant hypoxia in zebrafish hearts. Comparative biochemistry and physiology. B: biochemistry & molecular biology 178, 77-92.

Michelmore, R.W., Paran, I., Kesseli, R., 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proceedings of the national academy of sciences 88, 9828-9832.

Mott, R., Talbot, C.J., Turri, M.G., Collins, A.C., Flint, J., 2000. A method for fine mapping quantitative trait loci in outbred animal stocks. Proceedings of the national academy of sciences 97, 12649-12654.

Nikinmaa, M., Rees, B.B., 2005. Oxygen-dependent gene expression in fishes. American journal of physiology. Regulatory, integrative and comparative physiology 288, 1079-1090.

Niklasson, L., Sundh, H., Fridell, F., Taranger, G.L., Sundell, K., 2011. Disturbance of the intestinal mucosal immune system of farmed Atlantic salmon (*Salmo salar*), in response to long-term hypoxic conditions. Fish and shellfish immunology 31, 1072-1080.

Nilsson, G.E., Rosen, P., Johansson, D., 1993. Anoxic depression of spontaneous locomotor activity in crucian carp quantified by a computerized imaging technique. Journal of experimental biology 180, 153-162.

Ninwichian, P., Peatman, E., Liu, H., Kucuktas, H., Somridhivej, B., Liu, S.K., Li, P., Jiang, Y.L., Sha, Z.X., Kaltenboeck, L., Abernathy, J.W., Wang, W.Q., Chen, F., Lee, Y., Wong, L.L., Wang, S.L., Lu, J.G., Liu, Z.J., 2012. Second-Generation Genetic Linkage Map of Catfish and Its Integration with the BAC-Based Physical Map. G3: genes, genomics, genetics 2, 1233-1241.

Nishimura, S., Watanabe, T., Mizoshita, K., Tatsuda, K., Fujita, T., Watanabe, N., Sugimoto, Y., Takasuga, A., 2012. Genome-wide association study identified three major QTL for carcass weight including the PLAG1-CHCHD7 QTN for stature in Japanese Black cattle. BMC genetics 13, 40.

Norman, J.D., Danzmann, R.G., Glebe, B., Ferguson, M.M., 2011. The genetic basis of salinity tolerance traits in Arctic charr (*Salvelinus alpinus*). BMC genetics 12.

Ober, C., Abney, M., McPeek, M.S., 2001. The genetic dissection of complex traits in a founder population. The american journal of human genetics 69, 1068-1079.

Peatman, E., Bao, B.L., Peng, X., Baoprasertkul, P., Brady, Y., Liu, Z.J., 2006. Catfish CC chemokines: genomic clustering, duplications, and expression after bacterial infection with *Edwardsiella ictaluri*. Molecular genetics and genomics 275, 297-309.

Peatman, E., Li, C., Peterson, B.C., Straus, D.L., Farmer, B.D., Beck, B.H., 2013. Basal polarization of the mucosal compartment in Flavobacterium columnare susceptible and resistant channel catfish (*Ictalurus punctatus*). Molecular immunology 56, 317-327.

Pollock, M.S., Clarke, L.M.J., Dube, M.G., 2007. The effects of hypoxia on fishes: from ecological relevance to physiological effects. Environmental reviews 15, 1-14.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., Reich, D., 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nature genetics 38, 904-909.

Price, A.L., Zaitlen, N.A., Reich, D., Patterson, N., 2010. New approaches to population stratification in genome-wide association studies. Nature reviews genetics 11, 459-463.

Qiu, F.Z., Zheng, Y.L., Zhang, Z.L., Xu, S.Z., 2007. Mapping of QTL associated with waterlogging tolerance during the seedling stage in maize. Annals of botany 99, 1067-1081.

Qu, H.Q., Tien, M., Polychronakos, C., 2010. Statistical significance in genetic association studies. Clinical and investigative medicine 33, E266-270.

Quarrie, S.A., Lazić-Jančić, V., Kovačević, D., Steed, A., Pekić, S., 1999. Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize. Journal of experimental botany 50, 1299-1306.

Ravi, R., Mookerjee, B., Bhujwalla, Z.M., Sutter, C.H., Artemov, D., Zeng, Q.W., Dillehay, L.E., Madan, A., Semenza, G.L., Bedi, A., 2000. Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1 alpha. Gene development 14, 34-44.

Rengmark, A.H., Slettan, A., Lee, W.J., Lie, O., Lingaas, F., 2007. Identification and mapping of genes associated with salt tolerance in tilapia. Journal of fish biology 71, 409-422.

Risch, N., Merikangas, K., 1996. The future of genetic studies of complex human diseases. Science 273, 1516-1517.

Roesner, A., Hankeln, T., Burmester, T., 2006. Hypoxia induces a complex response of globin expression in zebrafish (*Danio rerio*). Journal of experimental biology 209, 2129-2137.

Sabatti, C., Service, S.K., Hartikainen, A.L., Pouta, A., Ripatti, S., Brodsky, J., Jones, C.G., Zaitlen, N.A., Varilo, T., Kaakinen, M., Sovio, U., Ruokonen, A., Laitinen, J., Jakkula, E., Coin, L., Hoggart, C., Collins, A., Turunen, H., Gabriel, S., Elliot, P., McCarthy, M.I., Daly, M.J., Jarvelin, M.R., Freimer, N.B., Peltonen, L., 2009. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. Nature genetics 41, 35-46.

Schoenebeck, J.J., Hutchinson, S.A., Byers, A., Beale, H.C., Carrington, B., Faden, D.L., Rimbault, M., Decker, B., Kidd, J.M., Sood, R., 2012. Variation of BMP3 contributes to dog breed skull diversity. PLoS genetics 8, e1002849.

Semenza, G.L., 2000. HIF-1: mediator of physiological and pathophysiological responses to hypoxia. Journal of applied physiology 88, 1474-1480.

Semenza, G.L., 2012. Hypoxia-inducible factors in physiology and medicine. Cell 148, 399-408.

Serapion, J., Kucuktas, H., Feng, J.A., Liu, Z.J., 2004. Bioinformatic mining of type I microsatellites from expressed sequence tags of channel catfish (*Ictalurus punctatus*). Marine biotechnology 6, 364-377.

Sloman, K.A., Wood, C.M., Scott, G.R., Wood, S., Kajimura, M., Johannsson, O.E., Almeida-Val, V.M., Val, A.L., 2006. Tribute to R. G. Boutilier: the effect of size on the physiological and behavioural responses of oscar, *Astronotus ocellatus*, to hypoxia. The journal of experimental biology 209, 1197-1205.

Somridhivej, B., Wang, S.L., Sha, Z.X., Liu, H., Quilang, J., Xu, P., Li, P., Hue, Z.L., Liu, Z.J., 2008. Characterization, polymorphism assessment, and database construction for microsatellites from BAC end sequences of channel catfish (*Ictalurus punctatus*): A resource for integration of linkage and physical maps. Aquaculture 275, 76-80.

Spielman, R.S., McGinnis, R.E., Ewens, W.J., 1993. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). American journal of human genetics 52, 506-516.

Su, T., Han, Y.F., Yu, Y.W., Tan, X.J., Li, X.P., Hou, J.G., Du, Y., Shen, J., Wang, G.P., Ma, L.Y., Jiang, S., Zhang, H.W., Cao, G.W., 2013. A GWAS-identified susceptibility locus on chromosome 11q13.3 and its putative molecular target for prediction of postoperative prognosis of human renal cell carcinoma. Oncology letters 6, 421-426.

Sun, C.M., Wang, B.Q., Yan, L., Hu, K.N., Liu, S., Zhou, Y.M., Guan, C.Y., Zhang, Z.Q., Li, J.N., Zhang, J.F., Chen, S., Wen, J., Ma, C.Z., Tu, J.X., Shen, J.X., Fu, T.D., Yi, B., 2016. Genome-Wide Association Study Provides Insight into the Genetic Control of Plant Height in Rapeseed (*Brassica napus L.*). Frontiers in plant science 7, 1102.

Sun, F.Y., Peatman, E., Li, C., Liu, S.K., Jiang, Y.L., Zhou, Z.C., Liu, Z.J., 2012. Transcriptomic signatures of attachment, NF-kappa B suppression and IFN stimulation in the catfish gill following columnaris bacterial infection. Developmental & comparative immunology 38, 169-180.

Sun, L.Y., Liu, S.K., Bao, L.S., Li, Y., Feng, J.B., Liu, Z.J., 2015. Claudin multigene family in channel catfish and their expression profiles in response to bacterial infection and hypoxia as revealed by meta-analysis of RNA-Seq datasets. Comparative biochemistry and physiology. D genomics and proteomics 13, 60-69.

Sun, L.Y., Liu, S.K., Wang, R.J., Jiang, Y.L., Zhang, Y., Zhang, J.R., Bao, L.S., Kaltenboeck, L., Dunham, R., Waldbieser, G., Liu, Z.J., 2014. Identification and analysis of genome-wide snps provide insight into signatures of selection and domestication in channel catfish (*Ictalurus punctatus*). PloS one 9, e109666.

Tan, S., Yao, J., Zhou, T., Liu, S., Yuan, Z., Tian, C., Li, Q., Liu, Z., 2017. Identification, annotation and expression analysis of 29 Rho GTPase genes from channel catfish (*Ictalurus punctatus*) after bacterial infections. Developmental & comparative immunology 67, 445-451.

Turner, S., Armstrong, L.L., Bradford, Y., Carlson, C.S., Crawford, D.C., Crenshaw, A.T., Andrade, M., Doheny, K.F., Haines, J.L., Hayes, G., 2011. Quality control procedures for genome-wide association studies. Current protocols in human genetics 68, 1-19.

van der Meer, D.L.M., van den Thillart, G.E.E.J.M., Witte, F., de Bakker, M.A.G., Besser, J., Richardson, M.K., Spaink, H.P., Leito, J.T.D., Bagowski, C.P., 2005. Gene expression profiling of the long-term adaptive response to hypoxia in the gills of adult zebrafish. American journal of physiology. Regulatory, integrative and comparative physiology 289, 1512-1519.

Van Gestel, S., Houwing-Duistermaat, J.J., Adolfsson, R., van Duijn, C.M., Van Broeckhoven, C., 2000. Power of selective genotyping in genetic association analyses of quantitative traits. Behavior genetics 30, 141-146.

VanToai, T.T., St Martin, S.K., Chase, K., Boru, G., Schnipke, V., Schmitthenner, A.F., Lark, K.G., 2001. Identification of a QTL associated with tolerance of soybean to soil waterlogging. Crop Science 41, 1247-1252.

Vanwaversveld, J., Addink, A.D.F., Vandenthillart, G., Smit, H., 1988. Direct calorimetry on free swimming goldfish at different oxygen levels. Journal of thermal analysis and calorimetry 33, 1019-1026.

Vuori, K.A., Soitamo, A., Vuorinen, P.J., Nikinmaa, M., 2004. Baltic salmon (*Salmo salar*) yolk-sac fry mortality is associated with disturbances in the function of hypoxia-inducible transcription factor (HIF-1α) and consecutive gene expression. Aquatic toxicology 68, 301-313.

Waldbieser, G.C., Bosworth, B.G., Nonneman, D.J., Wolters, W.R., 2001. A microsatellite-based genetic linkage map for channel catfish, *Ictalurus punctatus*. Genetics 158, 727-734.

Wang, C.M., Lo, L.C., Zhu, Z.Y., Yue, G.H., 2006. A genome scan for quantitative trait loci affecting growth-related traits in an F1 family of Asian seabass (*Lates calcarifer*). BMC genomics 7, 274.

Wang, R., Sun, L., Bao, L., Zhang, J., Jiang, Y., Yao, J., Song, L., Feng, J., Liu, S., Liu, Z., 2013a. Bulk segregant RNA-seq reveals expression and positional candidate genes and allele-specific expression for disease resistance against enteric septicemia of catfish. BMC genomics 14, 929.

Wang, R.J., Sun, L.Y., Bao, L.S., Zhang, J.R., Jiang, Y.L., Yao, J., Song, L., Feng, J.B., Liu, S.K., Liu, Z.J., 2013b. Bulk segregant RNA-seq reveals expression and positional candidate genes and allele-specific expression for disease resistance against enteric septicemia of catfish. BMC genomics 14.

Wang, S.L., Peatman, E., Abernathy, J., Waldbieser, G., Lindquist, E., Richardson, P., Lucas, S., Wang, M., Li, P., Thimmapuram, J., Liu, L., Vullaganti, D., Kucuktas, H., Murdock, C., Small, B.C., Wilson, M., Liu, H., Jiang, Y.L., Lee, Y., Chen, F., Lu, J.G., Wang, W.Q., Xu, P., Somridhivej, B., Baoprasertkul, P., Quilang, J., Sha, Z.X., Bao, B.L., Wang, Y.P., Wang, Q., Takano, T., Nandi, S., Liu, S.K., Wong, L.L., Kaltenboeck, L., Quiniou, S., Bengten, E., Miller, N., Trant, J., Rokhsar, D., Liu, Z.J., Consortium, C.G., 2010. Assembly of 500,000 inter-specific catfish expressed sequence tags and large scale gene-associated marker development for whole genome association studies. Genome biology 11, R8.

Wang, S.L., Xu, P., Thorsen, J., Zhu, B.L., de Jong, P.J., Waldbieser, G., Kucuktas, H., Liu, Z.J., 2007. Characterization of a BAC library from channel catfish *Ictalurus punctatus*: Indications of high levels of chromosomal reshuffling among teleost genomes. Marine biotechnology 9, 701-711.

Wang, W.Y., Barratt, B.J., Clayton, D.G., Todd, J.A., 2005. Genome-wide association studies: theoretical and practical concerns. Nature reviews genetics 6, 109-118.

Wang, X., Liu, S., Jiang, C., Geng, X., Zhou, T., Li, N., Bao, L., Li, Y., Yao, J., Yang, Y., Zhong, X., Jin, Y., Dunham, R., Liu, Z., 2017. Multiple across-strain and within-strain QTL suggest highly complex genetic architecture for hypoxia tolerance in channel catfish. Molecular genetics and genomics 292, 63-76.

Welker, T.L., McNulty, S.T., Klesius, P.H., 2007. Effect of sublethal hypoxia on the immune response and susceptibility of channel catfish, *Ictalurus punctatus*, to enteric septicemia. Journal of the world aquaculture society 38, 12-23.

Wiener, C.M., Booth, G., Semenza, G.L., 1996. In vivo expression of mRNAs encoding hypoxia-inducible factor 1. Biochemical and biophysical research communications 225, 485-488.

Wringe, B.F., Devlin, R.H., Ferguson, M.M., Moghadam, H.K., Sakhrani, D., Danzmann, R.G., 2010. Growth-related quantitative trait loci in domestic and wild rainbow trout (*Oncorhynchus mykiss*). BMC genetics 11, 63.

Xia, Y., Choi, H.K., Lee, K., 2012. Recent advances in hypoxia-inducible factor (HIF)-1 inhibitors. European journal of medicinal chemistry 49, 24-40.

Xu, P., Wang, S.L., Liu, L., Thorsen, J., Kucuktas, H., Liu, Z.J., 2007. A BAC-based physical map of the channel catfish genome. Genomics 90, 380-388.

Yang, J.A., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., Goddard, M.E., Visscher, P.M., 2010. Common SNPs explain a large proportion of the heritability for human height. Nature genetics 42, 565-569.

Yang, Y., Fu, Q., Zhou, T., Li, Y., Liu, S., Zeng, Q., Wang, X., Jin, Y., Tian, C., Qin, Z., 2017. Analysis of apolipoprotein genes and their involvement in disease response of channel catfish after bacterial infection. Developmental & comparative immunology 67, 464-470.

Yao, J., Mu, W.J., Liu, S.K., Zhang, J.R., Wen, H.S., Liu, Z.J., 2015. Identification, phylogeny and expression analysis of suppressors of cytokine signaling in channel catfish. Molecular immunology 64, 276-284.

Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Cuo, Z.X.P., Pool, J.E., Xu, X., Jiang, H., Vinckenbosch, N., Korneliussen, T.S., Zheng, H.C., Liu, T., He, W.M., Li, K., Luo, R.B., Nie, X.F., Wu, H.L., Zhao, M.R., Cao, H.Z., Zou, J., Shan, Y., Li, S.Z., Yang, Q., Asan, Ni, P.X., Tian, G., Xu, J.M., Liu, X.A., Jiang, T., Wu, R.H., Zhou, G.Y., Tang, M.F., Qin, J.J., Wang, T., Feng, S.J., Li, G.H., Huasang, Luosang, J.B., Wang, W., Chen, F., Wang,

Y.D., Zheng, X.G., Li, Z., Bianba, Z.M., Yang, G., Wang, X.P., Tang, S.H., Gao, G.Y., Chen, Y., Luo, Z., Gusang, L., Cao, Z., Zhang, Q.H., Ouyang, W.H., Ren, X.L., Liang, H.Q., Zheng, H.S., Huang, Y.B., Li, J.X., Bolund, L., Kristiansen, K., Li, Y.R., Zhang, Y., Zhang, X.Q., Li, R.Q., Li, S.G., Yang, H.M., Nielsen, R., Wang, J., Wang, J.A., 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. Science 329, 75-78.

Yu, J., Pressoir, G., Briggs, W.H., Bi, I.V., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature genetics 38, 203-208.

Zelzer, E., Olsen, B.R., 2005. Multiple roles of vascular endothelial growth factor (VEGF) in skeletal development, growth, and repair. Current topics in developmental biology 65, 169-187.

Zeng, Q., Fu, Q., Li, Y., Waldbieser, G., Bosworth, B., Liu, S., Yang, Y., Bao, L., Yuan, Z., Li, N. and Liu, Z., 2017. Development of a 690 K SNP array in catfish and its application for genetic mapping and validation of the reference genome sequence. Scientific reports, 7, 40347.

Zhang, Z., Ersoz, E., Lai, C.Q., Todhunter, R.J., Tiwari, H.K., Gore, M.A., Bradbury, P.J., Yu, J., Arnett, D.K., Ordovas, J.M., Buckler, E.S., 2010. Mixed linear model approach adapted for genome-wide association studies. Nature genetics 42, 355-360.

Zhang, Z.P., Ju, Z.L., Wells, M.C., Walter, R.B., 2009. Genomic approaches in the identification of hypoxia biomarkers in model fish species. Journal of experimental marine biology and ecology 381, 180-187.

Zhao, K., Aranzana, M.J., Kim, S., Lister, C., Shindo, C., Tang, C., Toomajian, C., Zheng, H., Dean, C., Marjoram, P., Nordborg, M., 2007. An Arabidopsis example of association mapping in structured samples. PLoS genetics 3, e4.

Zhong, H., Chiles, K., Feldser, D., Laughner, E., Hanrahan, C., Georgescu, M.M., Simons, J.W., Semenza, G.L., 2000. Modulation of hypoxia-inducible factor 1 alpha expression by the epidermal growth factor/phosphatidylinositol 3-kinase/PTEN/AKT/FRAP pathway in human prostate cancer cells: Implications for tumor angiogenesis and therapeutics. Cancer research 60, 1541-1545.

Zhou, T., Liu, S., Geng, X., Jin, Y., Jiang, C., Bao, L., Yao, J., Zhang, Y., Zhang, J., Sun, L., Wang, X., Li, N., Tan, S., Liu Z., 2017. GWAS analysis of QTL for enteric septicemia of catfish and their involved genes suggest evolutionary conservation of a molecular mechanism of disease resistance. Molecular genetics and genomics 292, 231-242.

Ziegler, A., König, I.R., Thompson, J.R., 2008. Biostatistical aspects of genome‑wide association studies. Biometrical Journal 50, 8-28.

Chapter 2 Effects of strain and body weight on hypoxia tolerance of channel catfish

2.1 Abstract

Understanding how genetic background and environmental factors affect hypoxia tolerance is of great interest in aquaculture. Although channel catfish (*Ictalurus punctatus*) is highly tolerant to low concentration of dissolved oxygen, hypoxia still causes enormous mortalities. In this project, the effects of strain, body weight, and gender on hypoxia tolerance in channel catfish exposed to a lethal concentration of dissolved oxygen (0.1 mg/L). The variation in hypoxia tolerance, assessed as the time to loss of equilibrium, of channel catfish from six strains (103KS, Kansas, KMix, Marion, Marion S, and Thompson) was examined. Catfish (15–179 g) showed a large variation in resistant time to hypoxia, ranging from 8 to 104 min, and both strain and body weight contributed significantly to this variation ($P < 0.05$). 103KS and Marion S strains had higher hypoxia tolerance than the other strains, while the Marion strain had the poorest hypoxia tolerance ($P < 0.05$). In addition to genetic background, body weight positively correlated with hypoxia tolerance, but there were no significant differences between female and male catfish in hypoxia tolerance. The results indicate that genetic background and body weight are important factors that contribute variations in low-oxygen tolerance.

2.2 Background

Hypoxia is common in aquaculture because of high stocking densities. Although fish are efficient in oxygen utilization in water, exposures to hypoxia can cause high mortalities. In many

instances, although fish can survive hypoxic exposures, more adverse effects of hypoxia are realized through reduced metabolic rate (Aboagye and Allen, 2014), reduced growth rate and feed conversion efficiency (Buentello et al., 2000), and increased susceptibility to diseases (Welker et al., 2007). All these negative effects of hypoxia can lead to slow response to other environmental stresses and low survival ability for fish.

Several factors have been reported to affect hypoxia tolerance of fish including gender, body size, as well as genetic background. Sexual dimorphism is common in various performance traits of fish including low-oxygen tolerance. For instance, males were reported to have higher mortality rate than females with mosquitofish (*Gambusia affinis*) under extreme hypoxic conditions (Cech Jr et al., 1985). Fish size as reflected in body weight is another factor to affect performance traits in fish. It is commonly noted that fish with larger body weight tend to be less tolerant to hypoxia due to smaller ratio of gill surface area to total body volume, as reported for largemouth bass (Burleson et al., 2001). However, opposite observations were also reported. For instance, Almeida-Val et al. (2000) and Sloman et al. (2005) reported that larger fish had a higher level of tolerance to hypoxia in Oscar cichlid (*Astronotus ocellatus*). With regard to genetic background, strains were found to contribute to the variation in low-oxygen tolerance of channel × blue hybrid catfish (Dunham et al., 2014).

Channel catfish is the major aquaculture species in the United States. Genetic background was found to have large effects on growth rate, feed conversion efficiency, disease resistance and many other performance traits (Gjerde et al., 2011; Henryon et al., 2002). Identification of multiple QTL in different channel catfish strains suggested a complex genetic architecture of hypoxia tolerance (Wang et al. 2017). To investigate the effects of both genetic and environmental factors

on hypoxia tolerance, this study was conducted to evaluate the effects of strain, body weight, and gender on the variation in low-oxygen tolerance using survival analysis.

2.3 Materials and methods

2.3.1 Ethics statement

All experiments involving the handling and treatment of fish were approved by the Institutional Animal Care and Use Committee (IACUC) at Auburn University. Tissue samples were collected after euthanasia. All animal procedures were carried out according to the Guide for the Care and Use of Laboratory Animals and the Animal Welfare Act in the United States.

2.3.2 Experimental fish

One-year-old channel catfish from six strains were used in this study. Information of all experimental fish is summarized in Table 1. All catfish were produced at the E. W. Shell Fisheries Research Center, Alabama Agricultural Experiment Station, Auburn University, Alabama. The ancestries of the channel catfish strains are detailed in a study by Dunham and Smitherman (1984). In brief, Kansas strain was originated from Ninnescah River in Kansas, while Kansas S has been selected for body weight for eight generations. KMix strain was obtained by crossing Kansas S with Kansas. 103KS strain was obtained by crossing NWAC 103 with Kansas S. Thompson strain was from Mississippi with an early spawning capability. Marion strain was originated from Marion National Fish Hatchery, while Marion S strain was obtained by mass selection from the original Marion population for body weight.

Experimental fish were moved to the Auburn University Hatchery Challenge Facility one week before hypoxia challenge. All fish were anesthetized with MS-222 before being weighed and

injected with passive integrated transponder (PIT) tags using intramuscular injection. Gender of the experimental fish was determined by visual examination of body shape, and genital shape. All these procedures were conducted 1 week before the hypoxia challenge. The information of PIT tag number, body weight, gender, and strain of each individual was recorded. Then the fish were randomly mixed together and kept in two 305 × 90 cm tanks at 20 °C with aerated flow-through water. The experimental fish were fed once daily, and the concentration of dissolved oxygen (DO) in water was monitored twice a day using a YSI dissolved oxygen meter.

Table 1. Summary of experimental fish used in this study.

	Female			Male		
Strains	NO. of fish	Body weight (g, mean±SD[*])	Time before losing balance (min)	NO. of fish	Body weight (g, mean±SD[*])	Time before losing balance (min)
103KS	45	96.7±31.1	62.0	36	102.9±28.9	61.6
Kansas	32	57.5±20.1	42.6	51	70.2±18.8	43.5
KMix	57	83.1±25.8	49.6	25	80.8±24.3	47.6
Marion	20	52.8±19.7	35.9	27	55.4±22.7	33.9
Marion S	21	82.3±22.8	60.6	22	99.0±26.6	67.5
Thompson	40	101.6±29.2	51.5	17	114.8±26.5	61.9

SD is the standard deviation for body weight of each strain.

2.3.3 Hypoxia challenge

All experimental fish were acclimated at ambient temperature of 20°C in the aerated flow-through water for 72 h before treatment. Feeding was stopped 24 h prior to the hypoxia challenge. During the experiment, water temperature was allowed to fluctuate naturally from 20 to 21 °C. The aerator was turned off, and sodium sulfite was added in water to reduce the DO concentration (Boyd and Tucker, 2012). The DO concentration was reduced from ~ 9.0 to 0.1 mg/L in 1 h. Then DO concentration was kept at 0.1 mg/L by adding additional sodium sulfite, if needed, and constantly monitored by the YSI dissolved oxygen meter. After DO concentration reached a lethal

concentration of 0.1 mg/L (starting point for measurement of resistant time), fish were monitored for the signs of losing equilibrium. Fish were removed immediately after loss of equilibrium. The resistant time and sequence of individuals were recorded along with their PIT tag number. After sampling, fish were returned to well-oxygenated water for recovery.

2.3.4 Statistical analysis

Hypoxia tolerance of channel catfish was statistically analyzed with survival analysis using the 'survival' package in R software (Miller Jr, 2011; Therneau, 2013). To examine the variations among stains and gender, Kaplan-Meier curves were constructed for the six strains and two gender groups using Kaplan-Meier analysis (Kaplan and Meier, 1958). Comparison of Kaplan-Meier curves was based on the log-rank test (Bland and Altman, 2004). Regression analysis was conducted within each of the six strains to determine the effect of body weight on hypoxia tolerance without the compounding effect of strain.

Cox proportional-hazards regression model (Cox PH model) was performed with all predictor variables (strain, body weight, gender) and their interactions simultaneously to identify the variables that significantly affected hypoxia tolerance (time before losing balance) (Lin and Wei, 1989). Proportional hazards assumption was checked before model fitting using 'Rcmdr' package in R (Fox et al., 2009). Four different Cox PH models were developed. First model was a full model with all predictor variables (strain, gender and body weight) and their interactions. Then three additional models were developed: one model without the 'strain' variable, one model without the 'body weight' variable, and one model without the 'gender' variable. To determine which variables and interactions were significantly contributed to hypoxia tolerance, the likelihood ratio test was conducted to compare these three models with the full model. In addition, auto model

selection procedure was conducted to select an appropriate model. The selected stratified Cox PH model was used to determine the effect size of strains on resistant time to hypoxia in catfish. All tests were regarded as statistically significant when $P \leq 0.05$.

2.4 Results and discussion

The DO concentration in water was reduced to 0.1 mg/L and kept for the remaining period of the hypoxia challenge. Under the hypoxic condition, the experimental fish reduced their movement and occasionally swim up to the surface of water to capture atmospheric air by gulping. The first fish lost its balance at 8 min after DO concentration reached 0.1 mg/L, while the last fish lost balance at 104 min. A broad range of variations in resistant time to hypoxia was observed in channel catfish.

Large variation in resistant time to hypoxia was identified among six strains by comparing Kaplan-Meier curves of these strains (Figure 1). Apparently, the most tolerant strains are Marion S and 103KS, followed by Thompson, KMix, Kansas, and Marion strains. To determine if these differences were statistically significant, we conducted pairwise comparison of these Kaplan-Meier curves using log-rank test. Based on the χ^2 statistics of log-rank test, 103KS and Marion S strains had significantly higher levels of tolerance to hypoxia than Kansas, KMix and Marion strains (Table 2). Thompson strain had a significantly higher level of hypoxia tolerance when compared with Kansas and Marion strains. Similarly, to determine if different levels of hypoxia tolerance existed in different gender, survival curves of females and males were established (Figure 2). Results of log-rank test for survival curves of different gender showed no significant difference between males and females ($P = 0.471$), indicating that gender is not important for hypoxia tolerance.
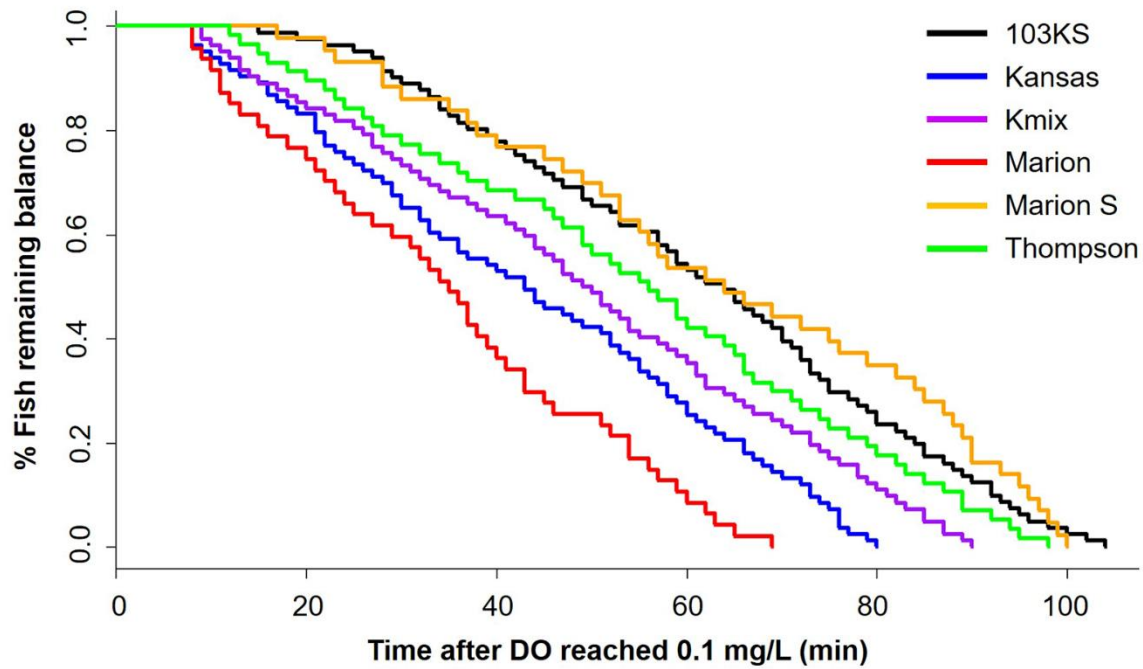
Figure 1. Comparison of hypoxia tolerance among six channel catfish strains.

Table 2. Pairwise comparison of hypoxia tolerance among six channel catfish strains using log-rank test.

Strains	103KS	Kansas	KMix	Marion	Marion S	Thompson
103KS		26.68	12.19	36.48	0.48	5.05
Kansas	<.0001*		2.49	0.03	25.94	10.45
KMix	0.0072*	0.8390		3.14	10.24	2.20
Marion	<.0001*	1.0000	0.6971		36.50	14.39
Marion S	1.0000	<.0001*	0.0204*	<.0001*		3.08
Thompson	0.3118	0.0182*	0.8920	0.0022*	0.7100	

The numbers above the diagonal are the Chi-square values for pairwise comparison, while the numbers below the diagonal are the corresponding P-values. Asterisks (*) indicated significant differences ($P \leq 0.05$).
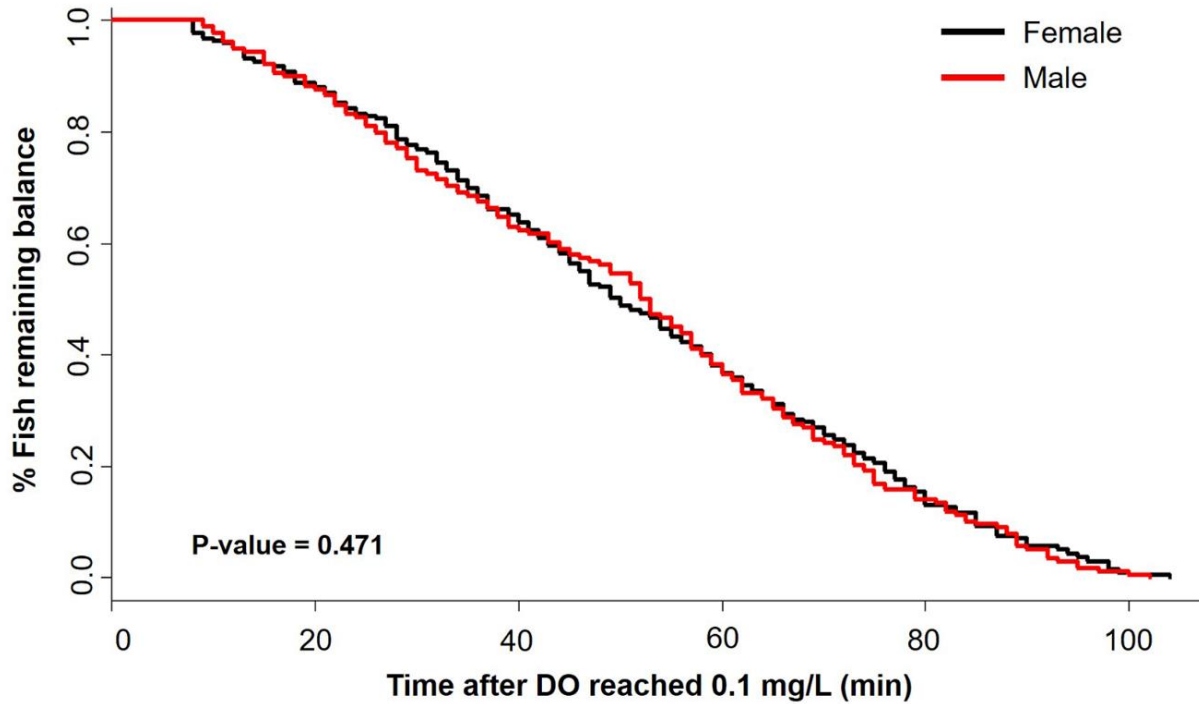
Figure 2. Comparison of hypoxia tolerance between female and male channel catfish.

To determine the effect of body weight on hypoxia tolerance without the compounding effect of strain, regression analysis was conducted within each of the six strains. The results are presented in Figure 3, body weight had a positive correlation with resistant time to hypoxia within each of six strains, although the correlation coefficient R varied between 0.23 (Marion strain) to 0.53 (KMix strain). This result suggests that fish with larger body weight tend to be more tolerant to hypoxic stress, but the extent of body weight effect on hypoxia tolerance is varied among strains.
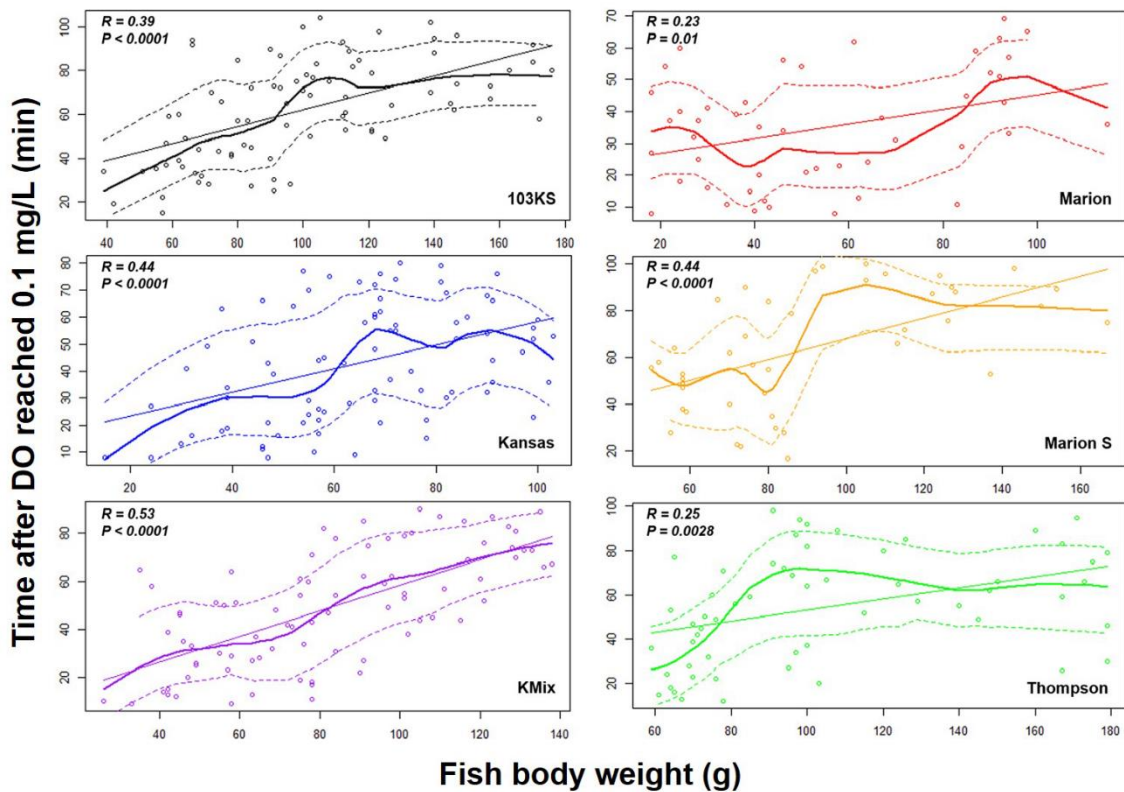
Figure 3. Correlation of body weight with hypoxia tolerance within strains. Linear regression analysis was conducted using body weight within strains and time before losing balance for the six strains.

To further verify the effects of strain, gender, and body weight on low-oxygen tolerance, both Cox PH model and likelihood ratio test were conducted. Results of likelihood ratio test of four Cox PH model demonstrated that (1) both strain and body weight had significantly effects on resistant time to hypoxia with a P value of 0.002 and 3.997e−15, respectively, whereas (2) gender had no significant effect on resistant time to hypoxia with a P value of 0.574. These results further confirmed the results from log-rank test on different Kaplan-Meier curves. Taken together, strain and body weight are two important effectors which significantly affect hypoxia tolerance in catfish.

To determine the effect size of strain and body weight on resistant time to hypoxia in catfish, model selection with stepwise procedure was conducted to obtain a fitted Cox PH model (Burnham and Anderson, 2004). The results of the Cox PH model are listed in Table 3. The most fitted model only included strain and body weight, indicating that strain and body weight independently contributed to the observed variation in resistant time to hypoxia ($P < 0.05$). The stratified Cox PH model was used to estimate effect size of strain as the variable body weight failed the assumption of proportionality (Grambsch and Therneau, 1994; Therneau and Grambsch, 2000). Based on the hazard ratio, effect size of strain on resistant time to hypoxia was calculated (Table 3). Using KMix strain as a reference, Marion S strain had a 44% higher chance, 103KS strain had a 40% higher chance, Kansas strain had a 16% higher chance, Thompson strain had a 6% higher chance to maintain balance, while Marion strain had a 1.19-fold lower chance to maintain balance under hypoxic conditions. The 103KS and Marion S strain had significantly higher hypoxia tolerance than KMix strain with a P-value of 0.02 and 0.05, respectively. Marion strain had significantly lower level of hypoxia tolerance ($P = 0.02$). These results suggested that 103KS and Marion S strains may be good choice for the breeding programs to improve hypoxia tolerance.

Table 3. Analysis of strain effects on hypoxia tolerance with stratified Cox Proportional-Hazards Model by considering body weight variations among strains. KMix strain was used as a reference in this model, and all the values were obtained by comparing with KMix strain.

Strain	Coefficient	Hazard ratio	P-value
103KS	-0.5043	0.6039	0.03
Kansas	-0.1722	0.8418	0.44
Marion	0.7855	2.1935	0.02
Marion S	-0.5723	0.5642	0.05
Thompson	-0.0608	0.9410	0.81

Surprisingly, this study demonstrated that body weight was positively correlated with the resistant time to hypoxia in channel catfish. This finding is the opposite of what was reported with channel × blue catfish that larger fish were less tolerant to hypoxic stress (Dunham et al., 2014). Several factors may explain this difference: (1) different catfish were used in these two studies; channel catfish from six strains were used in this study, while Dunham et al. (2014) used channel × blue catfish in their study; and (2) the range of fish size used in these two studies was different, which could be a major cause for the opposite results. Juvenile catfish with relatively small body weight were used in this study, whereas two-year-old hybrid catfish with larger body weight were used in the other study (Dunham et al., 2014). For juvenile fish, the "vitality" would be expected to increase with body weight. For example, fish have to rely on anaerobic ATP production (glycolysis) for survival under sever hypoxic conditions, smaller fish with higher mass-specific metabolic rate can run out of glycogen or reach lethal levels of anaerobic end-products much faster than larger ones (Nilsson and Östlund‑Nilsson, 2008). However, once body weight beyond a

certain threshold, the increasing of body weight can reduce the relative ratio of surface areas of gill to body size, which may lead to the observed results by Dunham et al. (2014). Future studies are required to delineate which of these, or additional reasons accounted for these different results.

Six strains used in our study have not been widely used in aquaculture. However, these strains are heavily used in catfish breeding programs for future applications in aquaculture. This study, in spite of being an initial study, suggested that the genetic architecture of low-oxygen tolerance is quite complex. The strong strain effect on low-oxygen tolerance may suggest that strain effect is needed to be consider when conducting any QTL mapping studies, as QTL identified in one strain may not be operational in another. Effect of fish size on low-oxygen tolerance may be variable. Taken results from our study along with study by Dunham et al. (2014),

a quadratic relationship may exist between body weight and low-oxygen tolerance. Body weight positively affects the low-oxygen tolerance in a certain threshold, but beyond this threshold, body weight may negatively affect the low-oxygen tolerance. Future studies are warranted to determine the exact correlation of body size with low-oxygen tolerance.

References

Aboagye D.L., Allen P.J., 2014. Metabolic and locomotor responses of juvenile paddlefish *Polyodon spathula* to hypoxia and temperature. Comparative biochemistry and physiology. A: comparative physiology 169, 51-59.

Almeida-Val V.M., Val A.L., Duncan W.P., Souza F.C., Paula-Silva M.N., Land S., 2000. Scaling effects on hypoxia tolerance in the Amazon fish *Astronotus ocellatus* (Perciformes: Cichlidae): contribution of tissue enzyme levels. Comparative biochemistry and physiology. B: comparative biochemistry 125, 219-226.

Bland J.M., Altman D.G., 2004. The logrank test. Bmj 328, 1073.

Boyd C.E., Tucker C.S., 2012. Pond aquaculture water quality management. Springer Science & Business Media, New York, US.

Buentello J.A., Gatlin D.M., Neill W.H., 2000. Effects of water temperature and dissolved oxygen on daily feed consumption, feed utilization and growth of channel catfish (*Ictalurus punctatus*). Aquaculture 182, 339-352.

Burleson M., Wilhelm D., Smatresk N., 2001. The influence of fish size size on the avoidance of hypoxia and oxygen selection by largemouth bass. Journal of fish biology 59,1336-1349.

Burnham K.P., Anderson D.R., 2004. Multimodel inference understanding AIC and BIC in model selection. Sociological methods & research 33, .261-304

Cech Jr J.J., Massingill M.J., Vondracek B., Linden A.L., 1985. Respiratory metabolism of mosquitofish, *Gambusia affinis*: effects of temperature, dissolved oxygen, and sex difference. Environmental biology of fishes 13, 297-307.

Dunham R.A., Ramboux A.C., Perera D.A., 2014. Effect of strain on tolerance of low dissolved oxygen of channel × blue catfish hybrids. Aquaculture 420, 25-28.

Dunham R.A., Smitherman R.O., 1984. Ancestry and breeding of catfish in the United States, Cir. 273. Alabama Agricultural Experiment Station, Auburn, AL.

Fox J., Andronic L., Ash M., Boye T., Calza S., Chang A., Grosjean P., Heiberger R., Kerns G., Lancelot R., 2009. Rcmdr: R Commander. R package version 1.5-4.

Gjerde B., Odegard J., Thorland I., 2011. Estimates of genetic variation in the susceptibility of Atlantic salmon (*Salmo salar*) to the salmon louse *Lepeophtheirus salmonis*. Aquaculture 314, 66-72.

Grambsch P.M., Therneau T.M., 1994. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 81, 515-526.

Henryon M., Jokumsen A., Berg P., Lund I., Pedersen P.B., Olesen N.J., Slierendrecht W.J., 2002. Genetic variation for growth rate, feed conversion efficiency, and disease resistance exists within a farmed population of rainbow trout. Aquaculture 209, 59-76.

Kaplan E.L., Meier P., 1958. Nonparametric estimation from incomplete observations. Journal of the american statistical association 53, 457-481.

Lin D.Y., Wei L.J., 1989. The robust inference for the Cox proportional hazards model. Journal of the american statistical association 84, 1074-1078.

Miller Jr R.G., 2011. Survival analysis. John Wiley & Sons, New York, US

Nilsson G.E., Östlund-Nilsson S., 2008. Does size matter for hypoxia tolerance in fish? Biological review 83, 173-189.

Sloman K., Scott G., Wood S., Kajimura M., Johannsson O., Almeida-Val V., Val A., Wood C., 2005. The effect of size on the physiological and behavioural responses of Oscar to hypoxia. Comparative biochemistry and physiology. A: comparative physiology 141, 176-177.

Therneau T., 2013. A package for survival analysis in S. R package version 2.37-4. URL http://CRAN.R-project.org/package=survival. Box. 980032, 23298-20032.

Therneau T.M., Grambsch P.M., 2000. Modeling survival data: extending the Cox model. Springer Science & Business Media, New York, US.

Wang, X., Liu, S., Jiang, C., Geng, X., Zhou, T., Li, N., Bao, L., Li, Y., Yao, J., Yang, Y., Zhong, X., Jin, Y., Dunham, R., Liu, Z., 2017. Multiple across-strain and within-strain QTL suggest highly complex genetic architecture for hypoxia tolerance in channel catfish. Molecular genetics and genomics 292, 63-76.

Welker T.L., Mcnulty S.T., Klesius P.H., 2007. Effect of sublethal hypoxia on the immune response and susceptibility of channel catfish, *Ictalurus punctatus*, to enteric septicemia. Journal of the world aquaculture society 38, 12-23.

Chapter 3 Identification of across-strain and within-strain QTL for hypoxia tolerance in channel catfish

3.1 Abstract

The ability to survive hypoxic conditions is important for various organisms, especially for aquatic animals. Teleost fish, representing more than 50% of vertebrate species, are extremely efficient in utilizing low concentration of dissolved oxygen in water. However, huge variations exist among various taxa of fish in their ability to tolerate hypoxia. In aquaculture, hypoxia tolerance is among the most important traits because hypoxia can cause major mortalities. Genetic enhancement for hypoxia tolerance in catfish is of great interest, but little was done with analysis of the genetic architecture of hypoxia tolerance. The objective of this project was to conduct a GWA study to identify QTL for hypoxia tolerance using the catfish 250K SNP array with channel catfish families from six strains. Multiple significant and suggestive QTL were identified both across strains and within strains. One significant QTL and four suggestive QTL were identified across strains. Six significant QTL and many suggestive QTL were identified within strains. There were rare overlaps among the QTL identified within the six strains, suggesting a complex genetic architecture of hypoxia tolerance. Overall, within-strain QTL explained a larger proportion of phenotypic variation than across-strain QTL. Many of genes within these QTL identified across strains and within strains have known functions for regulation of oxygen metabolism and involvement in hypoxia responses. Pathway analysis indicated that most of these genes were

involved in MAPK or PI3K/AKT/mTOR signaling pathways that were known to be important for hypoxia-mediated angiogenesis, cell proliferation, apoptosis and survival.

3.2 Background

Oxygen is essential for life of all aerobic organisms. The ability to adapt hypoxic environment determine the life style and distribution of these organisms. Response to hypoxia is a highly complex biological process. Under hypoxic conditions, organisms usually develop alternative strategies of energy metabolism, enzymatic activities and signal transduction pathways to survive. Hypoxia inducible factors (HIFs), especially HIFα, are believed to be a master switch for hypoxia responses (Semenza, 2000). HIFα proteins, in spite of being continuously synthesized, are rapidly degraded under normoxic conditions. Under hypoxic conditions, HIFα rapidly accumulates, and dimerizes to its partner HIFβ, which then binds to the hypoxia-response element (HRE) of effector genes, leading to activation of a wide variety of genes involved in both the cellular and systematic responses to hypoxia (Kietzmann et al., 2016). Many of these genes, such as vascular endothelial growth factor, glucose transporter 1, and insulin-like growth factors, are involved in erythropoiesis and angiogenesis, apoptosis, vascularization, and metabolism (Semenza et al., 2000). In addition to regulation at post-translational level, HIF-1α gene was also found to be transcriptionally induced by hypoxia in catfish (Geng et al., 2014). HIF activation is also regulated by tumor suppressors such as p53 and von Hippel-Lindau, or oncogenes such as RAS, mTOR, PI3K, and AKT (Kietzmann et al., 2016). The PI3K/AKT/mTOR pathway or MAPK pathway has been shown to regulate the expression of HIF-1α (Xia et al., 2012).

At the genomic level, living in hypoxic conditions can place selection pressure for genotypes that are more tolerant to hypoxia (Bigham et al., 2010). Most of genome level research

of hypoxia tolerance has been conducted with plants. For instance, QTL for hypoxia tolerance have been mapped in some plant species, such as in rice (Aggeli et al., 2002) and in maize (Qiu et al., 2007). With animals, much less work has been conducted. However, QTL associated with tolerance to high altitude hypoxia have been identified in humans. High-altitude human populations such as Sherpas, Tibetans, Ethiopians, and Andeans were found to possess specific SNPs associated with hypoxia tolerance (Beall, 2000; Jha et al., 2016; Yi et al., 2010). With fish, especially aquaculture species, genetic analysis of hypoxia tolerance is important. Improving hypoxia tolerance is of great economic interest in aquaculture. Under aquaculture conditions, exposures to hypoxia can lead to large mortalities due to the high stocking density. In addition, exposures to hypoxia can also cause depression of the immune system, leading to increased susceptibility to diseases (Kvamme et al., 2013; Welker et al., 2007).

Traditional QTL mapping has been conducted with various fish species. For instance, some valuable QTL were identified for traits related to growth in rainbow trout (Wringe et al., 2010), body lipid percentage in Atlantic salmon (Derayat et al., 2007), sexual maturation in Arctic char (Moghadam et al., 2007), and disease resistance in Atlantic salmon (Houston et al., 2010; Moen et al., 2009). A number of studies have been conducted dealing with hypoxia tolerance, but most focused on gene regulation after exposure to hypoxia in various species such as zebrafish (Guan et al., 2011), Rainbow trout (Faust et al., 2004), Atlantic salmon (Anttila et al., 2013), and paddlefish (Aboagye and Allen, 2014).

GWAS is an alternative method for QTL study. It is to use a large number of markers in the whole genome level with many unrelated individuals (McCarthy et al., 2008). The aim of GWAS is to map susceptibility effects through the detection of associations between genotype frequency and phenotype trait. GWAS has shown a great power to the identification of SNP-based

variants with relatively large effects on phenotype. Based on linkage disequilibrium, GWAS has a high ability to detect very small effects of marker–trait associations (Stranger et al., 2011). However, the major disadvantage of GWAS is relative high possibility of false positives due to population stratification. With the rapid development of next-generation sequencing, SNP discovery, genotyping and other molecular procedures become more technique advanced and economy efficient. The opportunities to apply GWAS to aquaculture species have increased.

Channel catfish is the primary aquaculture species in the United States. Much progress has been made in recent years for integrated genetic and genomic analysis. These included the construction of its physical map (Xu et al., 2007), identification of over 8 million genome-wide SNPs (Sun et al., 2014), the development of the 250K SNP array (Liu et al., 2014), the development of high-resolution genetic linkage maps (Li et al., 2015), the production of the reference genome sequence and its annotation (Liu et al., 2016), and analysis of genomic regions associated with performance and production traits (Geng et al., 2016; Geng et al., 2015). For instance, QTL for columnaris disease resistance (Geng et al., 2015), and for head size (Geng et al., 2016) have been identified using the channel catfish $\times$ blue catfish interspecific system. However, no studies have been conducted with hypoxia tolerance in catfish. Here, a GWA study was conducted to identify QTL associated with hypoxia tolerance using the catfish 250K SNP array.

3.3 Materials and methods

3.3.1 Ethics statement

All procedures involving the handling and treatment of fish used during this study followed the protocols approved by the Institutional Animal Care and Use Committee (IACUC) at Auburn University. Blood samples were collected after euthanasia. All animal procedures were carried out

according to the Guide for the Care and Use of Laboratory Animals and the Animal Welfare Act in the United States. This article did not use any personal data.

3.3.2 Experimental fish, hypoxia challenge and sample collection

Channel catfish used in this experiment were produced at the E. W. Shell Fisheries Research Center, Alabama Agricultural Experiment Station, Auburn University, Alabama. A total of 521 fish were selected from six channel catfish strains, including Marion (54 fish), Marion Select (67 fish), Thompson (82 fish), 103KS (99 fish), Kansas (107 fish), and Kmix (112 fish). Detailed ancestry of the channel catfish strains used in this study is described in Dunham and Smitherman (Dunham and Smitherman, 1984). All experimental fish were PIT (passive integrated transponder) tagged before mixed together for hypoxia treatment. Information of strain, body weight and sex of each fish was recorded with the corresponding PIT tag number.

Fish were acclimated at ambient temperature of 20°C in the aerated flow-through water for one week. The DO concentration was monitored daily using the YSI dissolved oxygen meter. During the experiment, the DO concentration was reduced gradually from ~9.0 mg/L to 0.1 mg/L in an hour using sodium sulfite (Boyd, 1982). After that, the DO concentration was kept constant, and fish were monitored for signs of losing balance to hypoxic stress. The time and sequence of each fish that lost balance were recorded. Blood samples were carefully collected and immediately stored in a 15 mL tube containing 5 mL of Cell Lysis Solution (10 mL Lysis Solution, Proteinase K 20mg/mL) for DNA isolation. Fish were returned to the well-oxygenated water for recovery after sampling.

3.3.3 Preparation of genomic DNA and genotyping

From each strain, we regarded the first ~35% fish lost balance as the hypoxia sensitive fish, the last ~35% fish as the hypoxia tolerant fish. In total, 376 fish from six different strains were selected for this study (Table 4) based on the selective genotyping method, which is economical and efficient by using relatively small sample size to achieve high statistical power (Darvasi and Soller, 1992; Jin et al., 2004). DNA was isolated using standard phenol/chloroform method. Briefly, the blood cells were broken by cell lysis solution and proteins were removed by protease K and protein precipitation solution after incubated blood samples at 55°C overnight. DNA was precipitated by isopropanol and collected by brief centrifugation, washed with 70% ethanol, air-dried, and rehydrated in reduced EDTA TE buffer (10 mM Tris–HCl, 0.1 mM EDTA, pH 8.0). DNA was quantified using spectroscopy by Nanodrop (Thermo Scientific). The integrity of DNA samples was checked by 1% agarose gel electrophoresis stained with ethidium bromide.

Table 4. Information of catfish samples used for this GWA study.

Strain name	Total number of fish	Number of sensitive fish	Number of tolerant fish	Mean of body weight (g)
103KS	72	36	36	109.3
Kansas	82	41	41	78.7
Kmix	83	41	42	97.5
Thompson	60	30	30	134.4
Marion Select	48	24	24	100.6
Marion	31	16	15	61.3

Genomic DNA samples were arranged in a 96-well microtiter plate, and normalized to a final concentration of 50 ng/μL with a final volume of 10 μL. We have developed the catfish 250K SNP array with well-distributed markers using Affymetrix Axiom genotyping technology (Liu et al., 2014). Genotyping using the catfish 250K SNP array was performed at GeneSeek (Lincoln, NE, USA). Five samples were excluded due to low quality or low call rate (<95%). The SNP genotype calling was performed using the Affymetrix Genotyping Console. 176,798 SNPs were

kept for subsequent analysis after excluding SNPs failed one or more of the following requirements: a call rate > 95%, a minor allele frequency (MAF) > 1%.

3.3.4 Statistical analysis

To determine potential association between SNPs and phenotypic variants, statistical analysis was carried out using the SNP & Variation Suite (SVS, Version 8.0). To generate a set of independent SNPs, linkage disequilibrium (LD) pruning step was conducted with a window size of 50 SNPs (the number of SNPs at each LD testing), a step of 5 SNPs (shift the window 5 SNPs forward), and r^2 threshold of 0.2 using composite haplotype method (CHM) (Wang et al., 2009; Weir and Ott, 1997). After LD pruning step, 15,571 independent SNP markers and LD blocks were identified. Identical by state (IBS) between all pairs of samples were estimated using the independent SNPs. The population structure was assessed by principal component analysis (PCA) (Price et al., 2006). Wright's F_{st} was estimated between all pairs of six strains using all the qualified SNPs, which can range from zero (no genetic divergence between the strains) to one (complete isolation of the strains from each other and the overall population) (Nei, 1973).

To account for the sample structure in the association test, Efficient Mixed-Model Association eXpedited (EMMAX) model was conducted with the first two principal components (PC1, PC2) and body weight as covariates using all samples from six strains. The model is listed as follows:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{u} + \mathbf{e}$$

where $\mathbf{Y}$ is a vector of time before fish lost balance to hypoxic stress; $\mathbf{X}$ is a matrix of fixed effects including first two PCs and fish body weight; $\boldsymbol{\beta}$ is a coefficient vector; $\mathbf{Z}$ is a matrix of random additive genetic effects, $\boldsymbol{u}$ is a vector representing the coefficients of the random effect,

Var($\boldsymbol{u}$) = Gσ^2_g where G is a simple IBS allele-sharing matrix and σ^2_g is the additive genetic variance; $\boldsymbol{e}$ is the vector of random residuals.

The Manhattan plot of the P-values was produced according to the channel catfish genome sequence (version Coco1.0) (Liu et al., 2016). The threshold P-value for genome-wide significance was determined based on Bonferroni-correction (Duggal et al., 2008). The threshold P-value of the 5% Bonferroni genome-wide significance was 0.05/15571=3.21e^{-6} ($-\log_{10}$ (P-value) = 5.49). The threshold P-value for the significance of "suggestive association", which allowed one false positive effect in a genome-wide test, was 1/15571=6.42e^{-5} ($-\log_{10}$ (P-value) = 4.19).

To identify QTL within strains, the similar statistical analysis was conducted with samples from each of the six strains. For each strain, the Manhattan plot was produced and corresponding threshold P-value was calculated. In Kansas strain, the threshold $-\log_{10}$ (P-value) of the 5% Bonferroni genome-wide significance was 4.51, and the threshold $-\log10$ (P-value) for the significance of "suggestive association" was 3.21. Similarly, the $-\log10$ (P-value) thresholds for significant association and suggestive association are 4.77 and 3.47 in Kmix strain, 4.54 and 3.24 in Thompson strain, 4.54 and 3.24 in 103KS strain, 4.39 and 3.09 in Marion strain and finally, 4.68 and 3.38 in Marion Select strain.

3.3.5 Sequence analysis

Genes within ±0.5 Mb of the most significant SNPs associated with hypoxia tolerance both across strains and within strains were examined for candidate genes and identified using the catfish reference genome (Liu et al., 2016). Synteny analysis of these genes were conducted between channel catfish and zebrafish to provide evidence for orthology. Similarly, genes most close to the

suggestive SNPs identified in this GWAS were also identified as potential candidate genes and analyzed for their functions in hypoxia responses.

3.4 Results

3.4.1 Hypoxia challenge

The DO concentration in water was gradually reduced to 0.1 mg/L and kept for the remaining period of the hypoxia challenge. Under the hypoxic condition, the experimental fish reduced their movement and occasionally swim up to the surface of water to capture atmospheric air by gulping. The Kaplan-Meier analysis was performed to identify the association between probability of fish hypoxia tolerance (time before losing balance) and hypoxia treatment time (Goel et al., 2010). The first fish started to lose balance at 8 minutes after hypoxia treatment, and the last fish losing balance was observed at 107 minutes after hypoxia treatment. As expected, the probability of an experimental fish to maintain balance decreased as the hypoxia treatment continued (Figure 4). For instance, the probability of an experimental fish losing balance exceeds 50% after 50 minutes of hypoxia treatment.
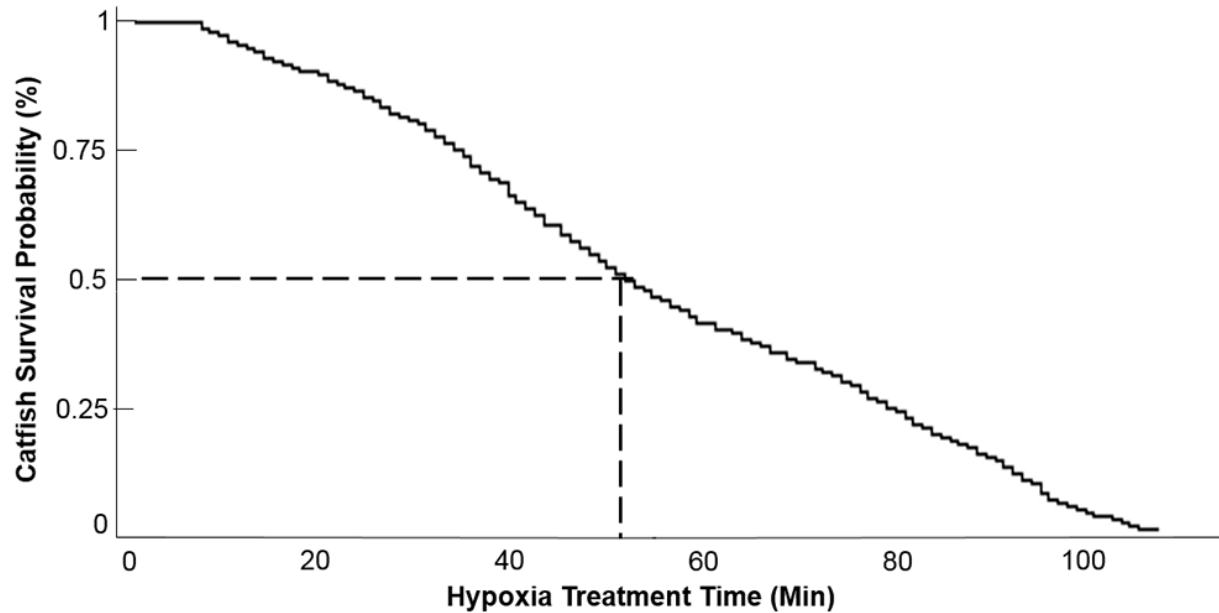
Figure 4. Overall tolerance of channel catfish under hypoxia treatment.

3.4.2 QTL associated with hypoxia tolerance across strains

The information of fish samples used for this GWA study was summarized in Table 4. After quality control filtering of unqualified samples and SNPs, principal component analysis (PCA) was conducted using eigenvalues as coordinates to visualize the population stratification with samples and SNPs passed quality control (Figure 5). To further determine the relationship of six strains, the values of Wright's F_{st} among six strains were calculated (Table 5). The values of F_{st} among strains varied from 0.11 to 0.25, indicating a relatively high level of genetic relatedness of the six strains. EMMAX model was conducted with the first two PCs to estimate the population structure in the association test.
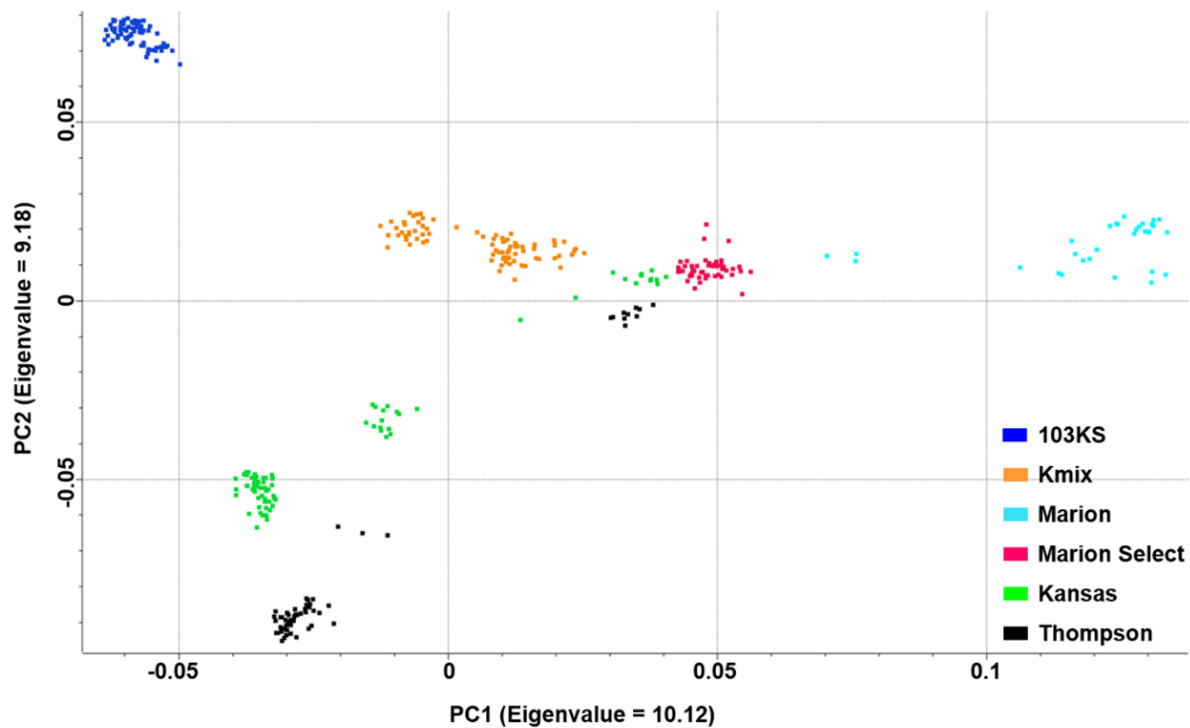
Figure 5. Population structure identified by PCA with the first two principal components. The coordinates are the first two principal components. Points in different colors represent different strains.

Table 5. Values of Wright's F_{st} reveal genetic relatedness among six strains.

Strain	Kmix	Marion	Marion Select	Kansas	Thompson	103KS
Kmix						
Marion	0.17					
Marion Select	0.15	0.2				
Kansas	0.11	0.2	0.16			
Thompson	0.16	0.22	0.2	0.14		
103KS	0.17	0.25	0.23	0.21	0.25	

First, QTL were analyzed using all samples across the six strains. A Manhattan plot of $-\log_{10}$ (P-value) for all SNPs associated with hypoxia tolerance was shown in Figure 6. Across the six strains, there were evidence for potential QTL in five genomic regions on four linkage groups, of which one genomic region on linkage group (LG) 6 was significantly associated with hypoxia

tolerance. Four additional genomic regions were found to be suggestively associated with hypoxia tolerance on LG5, LG10, and LG12. As shown in Figure 6, there appeared to be two suggestive genomic regions on LG5. On LG10 and LG12, there was only one genomic region each suggestively associated with hypoxia tolerance. The most significant SNP (AX-86110011) on LG6 is located at position 28,084,460 bp.
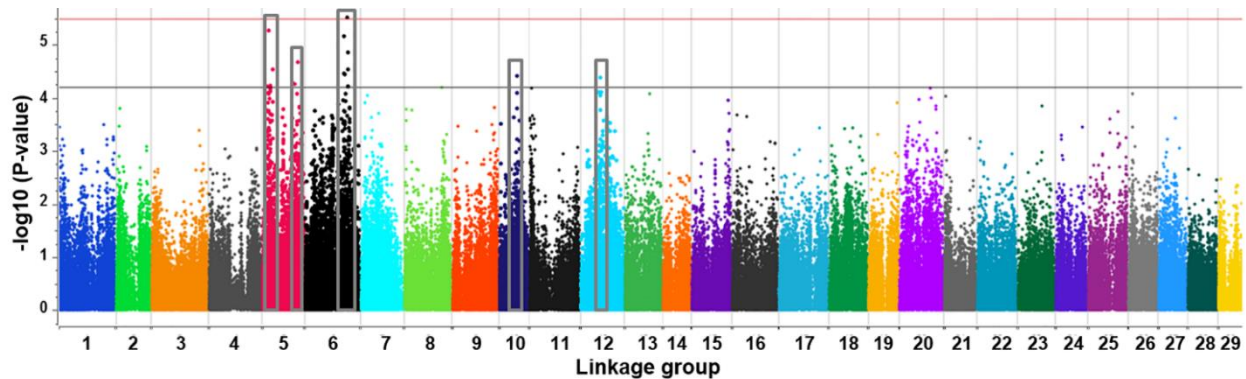
Figure 6. Manhattan plot reveals across-strain QTL associated with hypoxia tolerance. The red solid line indicates the threshold P-value for genome-wide significance. The black solid line indicates the threshold P-value for the significance of "suggestive association".

The phenotypic variation explained by significantly and suggestively associated SNPs were relatively small. The most significant SNP on LG6 explained approximately 6% of phenotypic variation, while other associated SNPs on LG6 explained 4-5% of phenotypic variation (Table 6). On LG5, two separate genomic regions were identified to be suggestively associated with hypoxia tolerance, including a set of SNPs with $-\log_{10}$(P-value) ranging from 4.19-5.27 (Table 6). One region spans from 3,073,377 bp to 6,575,020 bp, and the other region spans from 20,194,928 bp to 23,640,298 bp. They can explain 4-5% of phenotypic variation. One suggestively associated QTL was identified on LG10 and LG12, respectively (Table 6). On LG10, one SNP at position 11,248,369 bp was suggestively associated with hypoxia tolerance and explained ~4% of

phenotypic variation. On LG12, one SNP at position 12,924,457 bp was suggestively associated with hypoxia tolerance, explaining ~4% of phenotypic variation (Table 6).

Table 6. SNPs associated with hypoxia tolerance across strains. Grey and red shades indicate two separate associated genomic regions on LG5.

Linkage group	SNP ID	SNP position	-Log_{10} (P-value)	% Variance
	AX-85193644	25,756,477	4.45	4.51
	AX-85410024	25,789,234	5.15	5.30
	AX-85313934	26,633,015	4.43	4.49
LG6	AX-86110011	28,084,460	5.52	5.71
	AX-85214060	28,549,254	4.53	4.60
	AX-85254568	28,584,932	4.86	4.97
	AX-85252972	28,764,133	4.21	4.24
	AX-85353272	3,073,377	4.21	4.23
	AX-85216281	3,976,205	5.27	5.43
	AX-85324750	4,869,369	4.22	4.25
LG5	AX-85271686	5,189,256	4.19	4.22
	AX-85224364	6,575,020	4.53	4.60
	AX-85367587	20,194,928	4.26	4.30
	AX-85338386	21,541,353	4.25	4.28
	AX-85961782	23,640,298	4.66	4.75
LG10	AX-85217471	11,248,369	4.41	4.47
LG12	AX-85197805	12,924,457	4.38	4.43

3.4.3 QTL associated with hypoxia tolerance within strains

To investigate QTL within strains, EMMAX model was conducted with samples from each of the six strains. Many genomic regions were found to be associated with hypoxia tolerance within these strains. Significant SNPs were detected within three strains (Figure 7). With Kansas strain, significant SNPs were found on LG4, LG22, and LG25. In addition to these significant genomic regions, suggestive SNPs were found on many other linkage groups, including LG2, LG5, LG6, LG8, LG9, LG11, LG13, LG17, and LG20 (Figure 7A). Apparently, many of these could be pseudo-positive (see discussion), but some of these may potentially contain real QTL for hypoxia tolerance. With Kmix strain, two genomic regions were detected to contain significant SNPs on

LG7 and LG12. Additional suggestively associated SNPs were detected on LG2, LG3, LG5, LG10, LG17, LG18, and LG19 (Figure 7B). With Thompson strain, one SNP on LG7 was found to be significantly associated with hypoxia tolerance, and suggestively associated SNPs were detected on LG9, LG13, LG20, and LG21 (Figure 7C). With 103KS, Marion, and Marion Select strains, no significant SNPs were detected, but suggestive SNPs were found on LG10 for 103KS strain (Figure 8A), on LG4, LG13, and LG15 for Marion strain (Figure 8B), and on LG3, LG9, LG13, LG22, and on LG23 for Marion Select strain (Figure 8C).
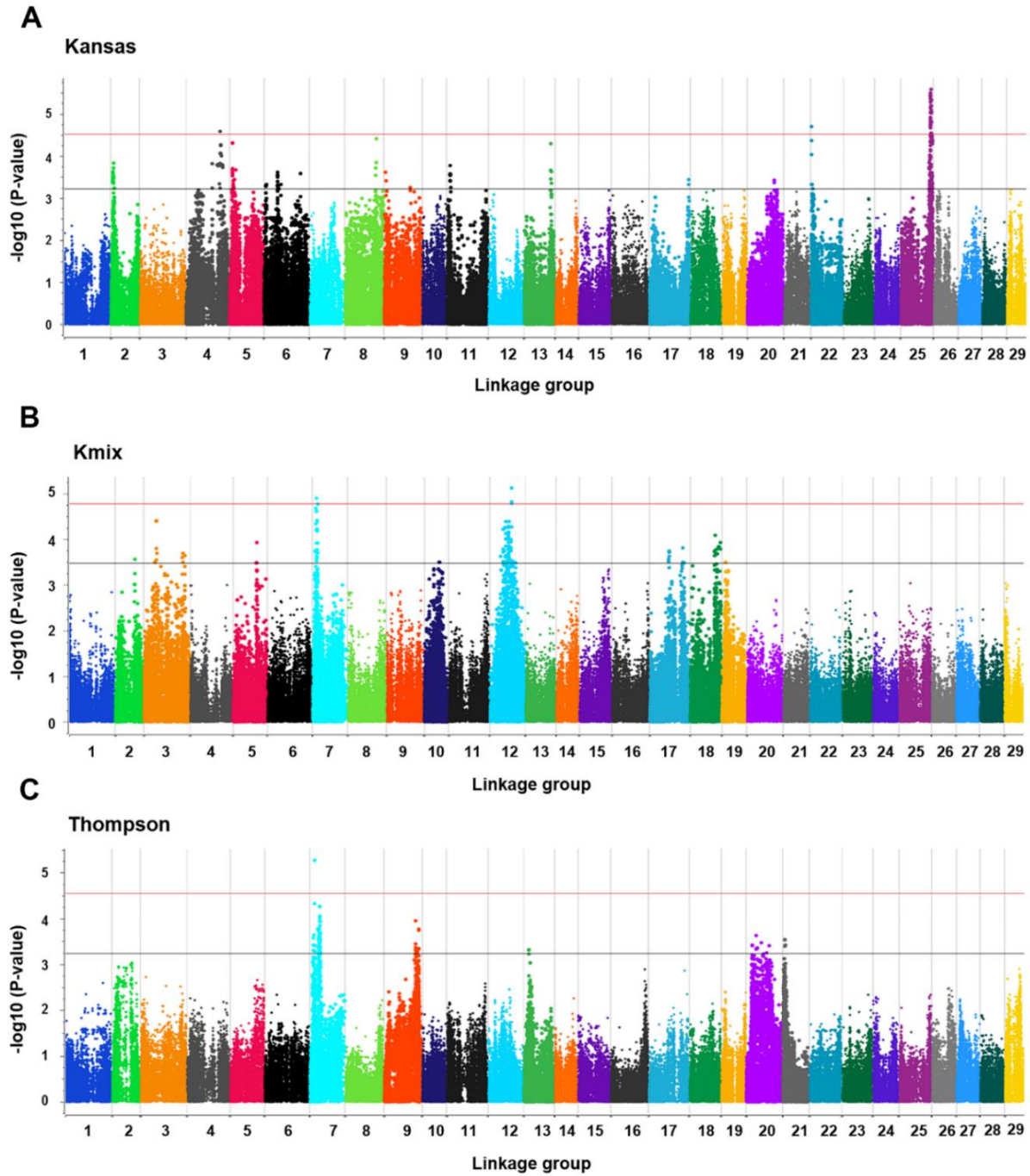
Figure 7. Manhattan plots reveal QTL for hypoxia tolerance within Kansas, Kmix and Thompson strains. A) QTL within Kansas strain. B) QTL within Kmix strain. C) QTL within

Thompson strain. The red solid line indicates the threshold P-value for genome-wide significance.

The black solid line indicates the threshold P-value for the significance of "suggestive association".
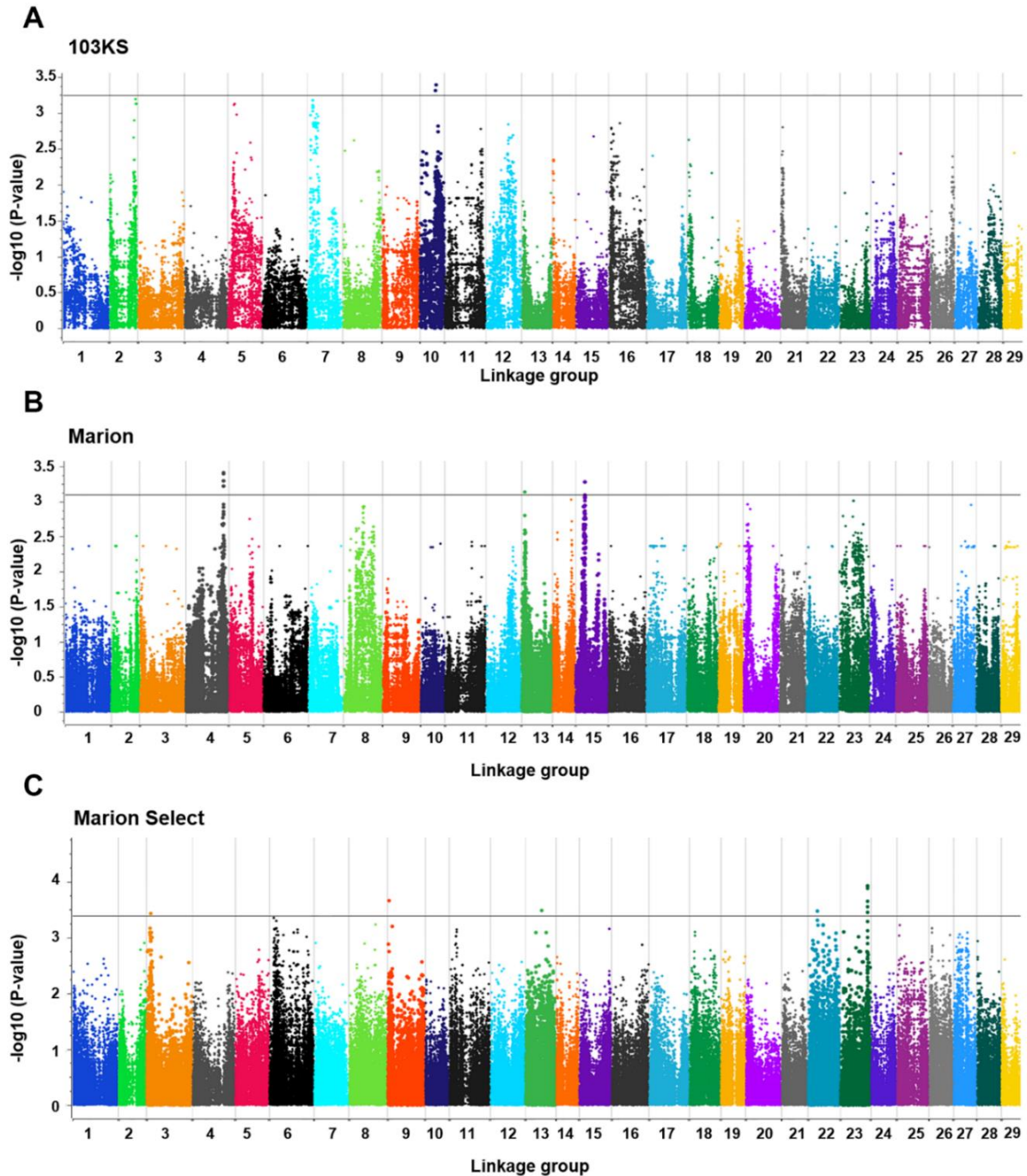

Figure 8. Manhattan plots reveal QTL for hypoxia tolerance within 103KS, Marion and Marion Select strains. A) QTL within 103KS strain. B) QTL within Marion strain. C) QTL within

Marion Select strain. The black solid line indicates the threshold P-value for the significance of "suggestive association".

The effects on phenotypic variation appeared to be relatively large for QTL significantly associated with hypoxia tolerance within strains (Table 7). With Kansas strain, the most significant SNPs on LG25, LG22 and LG4 explained approximately 25%, 21% and 20% of phenotypic variation, respectively. With Kmix strain, the most significant SNPs on LG12 and LG7 can explain approximately 23% and 22% of phenotypic variation, respectively. With Thompson strain, the most significant SNP on LG7 itself explained 32% of phenotypic variation.

Table 7. SNPs significantly associated with hypoxia tolerance within strains.

Linkage group	SNP ID	Position (bp)	-Log$_{10}$ (P-value)	% Variance
		Kansas		
	AX-85273690	23,848,240	4.83	22.02
	AX-85382689	24,066,260	4.73	21.53
	AX-85278184	24,083,110	4.72	21.51
	AX-85401604	24,116,919	4.73	21.53
	AX-85366674	24,133,568	4.77	21.75
	AX-85208027	24,147,963	5.41	24.60
	AX-85404384	24,149,673	5.02	22.84
	AX-85240400	24,161,130	5.14	23.39
	AX-86065291	24,169,339	5.48	24.91
	AX-85272411	24,171,584	4.90	22.32
	AX-85265728	24,189,546	5.26	23.95
	AX-85248269	24,210,520	5.36	24.38
LG25	AX-85382884	24,246,000	4.86	22.14
	AX-86065374	24,349,412	4.81	21.90
	AX-85368724	24,368,217	5.32	24.21
	AX-85317869	24,463,914	5.58	25.32
	AX-85333027	24,525,559	5.03	22.91
	AX-85305506	24,528,322	5.18	23.59
	AX-85234612	24,530,224	5.31	24.14
	AX-85208143	24,530,446	4.53	20.63
	AX-85242477	24,571,757	5.04	22.93
	AX-85297757	24,576,963	5.00	22.78
	AX-85953964	24,582,978	5.08	23.12
	AX-85280870	24,609,404	4.79	21.80
LG22	AX-85325454	383,931	4.69	21.37

LG4	AX-85294968	27,711,266	4.57	20.83
Kmix				
	AX-85394840	17,561,280	4.78	21.51
LG12	AX-85426612	17,792,919	5.12	23.04
	AX-85357080	17,800,330	4.81	21.68
LG7	AX-85370491	3,225,533	4.90	22.06
Thompson				
LG7	AX-85302554	3,746,317	5.26	32.04

3.4.4 Genes within the QTL regions significantly associated with hypoxia tolerance

As the reference genome sequence is available and the assembly was validated to be highly accurate, examination of the genes in genomic regions surrounding the significant SNPs may provide insights into potential candidate genes and potential mechanisms of hypoxia responses. We determined genes on LG6 within the ±0.5 Mb genomic region of the most significant SNP. A total of 25 genes exist in this region, and these genes and their locations relative to the most significant SNP are shown in Figure 9. Many of these genes are involved in hypoxia response pathways (see Discussion), such as *lrrc1* (leucine rich repeat containing 1), *tceb3* (transcription elongation factor B polypeptide 3), *mlip* (muscular LMNA-interacting protein), *fam83b* (family with sequence similarity 83 member b), *gclc* (glutamate-cysteine ligase catalytic subunit), *fgfr2* (fibroblast growth factor receptor 2), *plpp4* (phospholipid phosphatase 4), *fbxo9* (F-box protein 9), *bmp5* (bone morphogenetic protein 5), and *bag2* (BAG family molecular chaperone regulator 2).
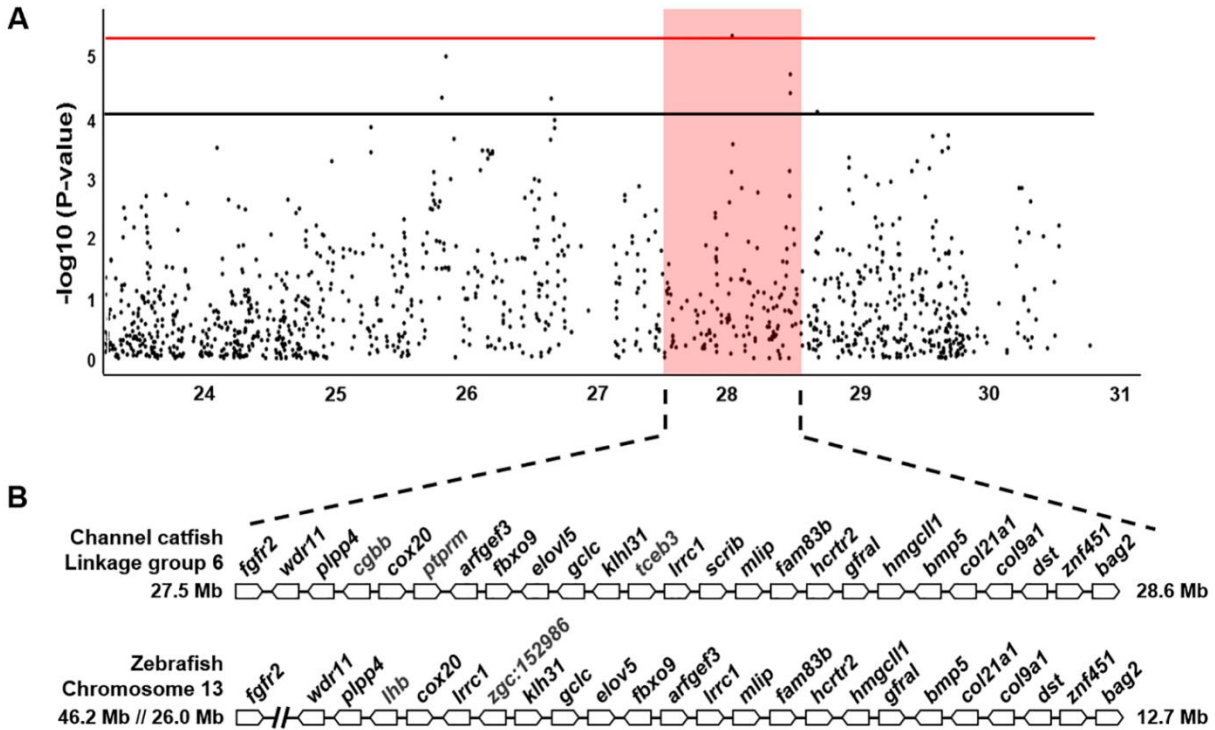
Figure 9. Genes surrounding the most significant SNP associated with hypoxia tolerance across strains on LG6. A) Regional -log10 (P-value) plot for the QTL. The red shade indicates ±0.5 Mb genomic region of the most significant SNP. The red solid line indicates the threshold P-value for genome-wide significance. The black solid line indicates the threshold P-value for the significance of "suggestive association". B) Genes within ±0.5 Mb genomic region of the most significant SNP. Synteny analysis was conducted between catfish and zebrafish. Genes with gray names are not conserved between channel catfish and zebrafish.

Although multiple linkage groups were detected to contain significant SNPs for hypoxia tolerance within strains, only one linkage group, LG25, within Kansas strain was found to contain multiple significant SNPs associated with hypoxia tolerance. Therefore, we determined the genes surrounding this specific genomic region (±0.5 Mb of the most significant SNP on LG25). A total of 34 genes exist in this genomic region (Figure 10). Similar to the situation for QTL across strains,

many genes included in this QTL region are involved in hypoxia response pathways, such as *nf1* (neurofibromin 1), *lgals9* (Lectin, galactoside binding soluble 9), *ucp2* (uncoupling protein 2), *slc1a3* (solute carrier family 1 member 3), *gdnf* (glial cell derived neurotrophic factor), and *dhrs13* (dehydrogenase/reductase 13). The closest genes surrounding the suggestive SNPs (Table 8) in both across-strain analyses and within-strain analyses.
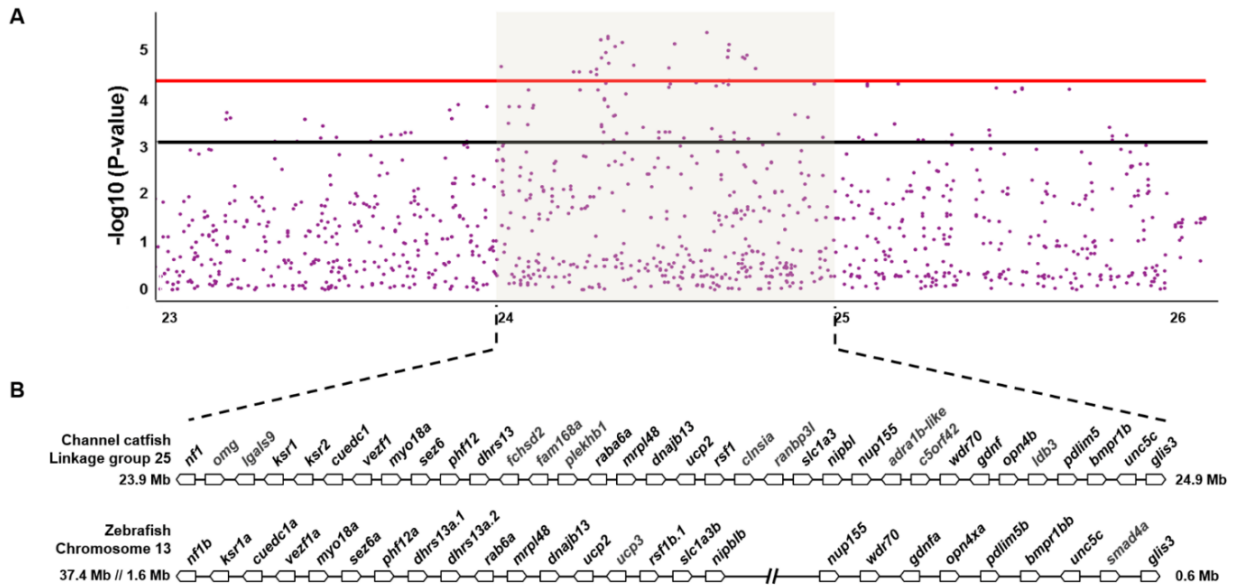
Figure 10. Genes surrounding the most significant SNP associated with hypoxia tolerance within Kansas strain on LG25. A) Regional -log10 (P-value) plot for the QTL. The grey shade indicates ±0.5 Mb genomic region of the most significant SNP. The red solid line indicates the threshold P-value for genome-wide significance. The black solid line indicates the threshold P-value for the significance of "suggestive association". B) Genes within ±0.5 Mb genomic region of the most significant SNP. Synteny analysis was conducted between catfish and zebrafish. Genes with gray names are not conserved between channel catfish and zebrafish.

Table 8. Closest genes to the suggestive SNPs associated with hypoxia tolerance for both across strains and within strains.

Type	Linkage group	SNP ID	SNP position	Gene name
across-strain	LG5	AX-85216281	3,976,205	BTB domain containing 17 (btbd17)
across-strain	LG5	AX-85961782	23,640,298	Mitogen-activated protein kinase kinase 6 (map2k6) Tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2 (tanc2)
across-strain	LG6	AX-85410024	25,789,234	Cytochrome c oxidase subunit 7A2 like (cox7a2l)
across-strain	LG10	AX-85217471	11,248,369	Intersectin 1 (itsn1)
across-strain	LG12	AX-85197805	12,924,457	Capicua (cic)
Kansas strain	LG2	AX-85331591	1,604,428	Glutamate decarboxylase 1 (gad1)
Kansas strain	LG5	AX-85304910	2,097,859	E3 ubiquitin-protein ligase TRIM39 (trim39)
Kansas strain	LG6	AX-85252247	1,352,645	Interferon regulatory factor 2 (irf2)
Kansas strain	LG6	AX-85196122	10,668,217	Gamma-synuclein (sncg)
Kansas strain	LG6	AX-85229118	29,371,054	CUB and sushi domain-containing protein 1 (csmd1)
Kansas strain	LG8	AX-85308773	25,315,833	Voltage-dependent R-type calcium channel subunit alpha-1E (cacna1e)
Kansas strain	LG9	AX-85362155	622,159	Synaptic vesicle glycoprotein 2B (sv2b)
Kansas strain	LG9	AX-85296181	21,395,031	Cytochrome P450 2K1 (cp2k1)
Kansas strain	LG11	AX-85222136	2,214,311	Ubiquitin-conjugating enzyme E2 D4 (ub2d4)
Kansas strain	LG13	AX-85285448	21,773,601	Delta-sarcoglycan (sgcd)
Kansas strain	LG17	AX-86076231	31,663,024	Glypican-6 (gpc6)
Kansas strain	LG20	AX-85352328	21,665,361	Mucolipin-2 (mcln2)
Kansas strain	LG22	AX-85325454	383,931	U2 snRNP-associated SURP motif-containing protein (u2surp)

Kmix strain	LG2	AX-85990897	15,998,754	CUGBP Elav-like family member 5 (celf5)
Kmix strain	LG3	AX-85237662	10,123,154	N-acetylglutamate synthase (nags)
Kmix strain	LG3	AX-85255451	32,163,410	Unconventional myosin-VI (myo6)
Kmix strain	LG5	AX-85287009	19,137,568	C-Jun-amino-terminal kinase-interacting protein 4 (jip4)
Kmix strain	LG7	AX-85370491	3,225,533	Neuronal growth regulator 1 (negr1)
Kmix strain	LG10	AX-85242426	11,808,873	Phospholipid-transporting ATPase IH (atp11a)
Kmix strain	LG12	AX-85426612	17,792,919	Retinoic acid receptor beta (rarb)
Kmix strain	LG17	AX-85218724	15,829,311	MYCBP Associated And Testis Expressed 1 (maats1)
Kmix strain	LG17	AX-85433101	27,141,224	Von Willebrand factor C domain-containing protein 2-like (vwc2l)
Kmix strain	LG18	AX-85211250	20,166,916	Cytoplasmic polyadenylation element-binding protein 4 (cpeb4)
Kmix strain	LG19	AX-85401386	2,580,625	Semaphorin-5A (sema5a)
Thompson strain	LG7	AX-85302554	3,746,317	Fatty-acid amide hydrolase 1 (faah)
Thompson strain	LG9	AX-85275004	25,282,329	Follistatin-related protein 5 (fstl5)
Thompson strain	LG13	AX-85390251	3,836,923	Serine/threonine-protein kinase (ulk1)
Thompson strain	LG20	AX-85231902	7,889,909	Leucine-rich repeat-containing protein 28 (lrc28)
Thompson strain	LG21	AX-86029063	1,435,994	Voltage-dependent calcium channel subunit alpha-2/delta-3 (cacna2d3)
103KS strain	LG10	AX-85387435	12,798,369	Transmembrane 9 superfamily member 2 (tm9s2)

Marion strain	LG4	AX-85237540	30,687,650	Ras-related protein Rap-2b (rap2b)
Marion strain	LG13	AX-85308641	1,990,490	LIM/homeobox protein Lhx5 (lhx5)
Marion strain	LG15	AX-85421803	7,248,445	Urokinase plasminogen activator surface receptor (upar)
Marion Select strain	LG3	AX-85337208	2,898,420	SRC kinase signaling inhibitor 1 (srcn1)
Marion Select strain	LG9	AX-85341574	579,165	Protein spire homolog 2 (spir2)
Marion Select strain	LG13	AX-85325943	12,700,036	Bcl2-associated agonist of cell death (bad)
Marion Select strain	LG23	AX-85238123	21,937,885	Transcription factor SOX-5 (sox5)

3.4.5 Correlation of the SNPs associated with hypoxia tolerance

Conditioned analyses were conducted to evaluate whether the SNPs associated with hypoxia tolerance were independent (Geng et al., 2015). The significant associated SNP on LG6 was included as a covariate in EMMAX model. After conditioning, the $-\log_{10}$ (P-value) of the remaining associated SNPs on LG6 drastically dropped below 2.0, indicating strong correlations among SNPs within the same linkage group. While the $-\log_{10}$ (P-value) of SNPs remain the same on different linkage groups. Similarly, when the conditioned analysis was conducted with associated SNPs on the same linkage group for QTL detected within strains, after conditioning, the $-\log_{10}$ (P-value) of the remaining associated SNPs on the same linkage group drastically dropped, while SNPs on different linkage group remain the same.

3.5 Discussion

In this study, we conducted a GWA study with six channel catfish strains using the catfish 250K SNP array to identify QTL associated with hypoxia tolerance. A large number of QTL were identified for hypoxia tolerance both across strains and within strains. Across the six strains, one

significant genomic region and four additional suggestively associated genomic regions were identified to contain QTL for hypoxia tolerance. Analyses of QTL associated with hypoxia tolerance within the six strains allowed detection of multiple significant QTL in three strains, and suggestive QTL in the other strains. In Kansas strain, three genomic regions were found to contain SNPs significantly associated with hypoxia tolerance, while additional nine genomic regions were found to contain SNPs suggestively associated with hypoxia tolerance. In Kmix strain, two significantly associated genomic regions and seven suggestively associated genomic regions were identified for hypoxia tolerance. In Thompson strain, one significant genomic region and four additional suggestively associated genomic regions were identified to contain QTL for hypoxia tolerance. With the other three strains, 103KS, Marion, and Marion Select, no significant SNPs were identified to be associated with hypoxia tolerance, but a large number of suggestive SNPs were identified.

The identification of a large number of QTL both across strains and within strains suggested a very complex genetic architecture of hypoxia tolerance in teleost fish. It is widely believed that hypoxia tolerance is a polygenic trait, involving both HIF-dependent and HIF-independent pathways (Li et al., 2013). With HIF-dependent pathway, HIF-1α is believed to be the master switch. Instead of degradation when under normoxic conditions, HIF-1α is no longer hydrolyzed, and it then coupled to HIF-1β, and the heterodimer bind to HRE to activate transcription of the effector genes under hypoxic conditions (Kietzmann et al., 2016). A large number of genes were reported to contain HRE, and therefore, any genomic sequence variations in the form of SNPs within HRE may potentially be involved in QTL controlling hypoxia tolerance. With HIF-independent pathways, multiple genes, such as PI3K, AKT, RAF, JNK, and COX2, were involved in response to hypoxia (Mizukami et al., 2007). Therefore, any SNPs within these

genes may also contribute to hypoxia tolerance. In addition, the preliminary results for analysis of hypoxia tolerance using various strains of catfish indicated strong strain difference in hypoxia tolerance, suggesting various QTL may be at work within various strains. However, this study is the very first QTL analysis for hypoxia tolerance in fish, there is no existing results to compare. The future analyses using even more strains, more families within each strain and large sample sizes are required to validate the genetic architecture for hypoxia tolerance.

QTL identified across strains barely overlapped with those identified within strains. This is expected because within-strain analyses may have more power to identify rare QTL within each strain, while across-strain analyses may be more efficient to detect common QTL for multiple strains (Ogut et al., 2015). In addition, the detected QTL across strains reflect the superiority of one allele over the other among strains, and they may explain variations of hypoxia tolerance among strains; while the allele differences within strains may reflect variations of hypoxia tolerance among families within each strain. As such, perhaps the QTL detected within strains are even more important for genetic enhancement programs. However, exercise caution must be noted because the sample sizes were quite small for both across-strain analyses and for within-strain analyses, especially for within-strain analyses. With small samples, it is prone to both type I and type II errors. One guiding principle can be used to provide some levels of confidence is the observation of many physically linked significant SNPs within the genomic region containing the most significant SNPs. For instance, for the QTL on LG25 of Kansas strain, twenty-four significant SNPs were detected within approximately 760 Kb genomic region containing the QTL associated with hypoxia tolerance (Figure 8A). The linked SNPs were all included in one haplotype block, suggesting that the association to hypoxia tolerance is related to the inclusion of the genomic segments due to co-segregation rather than genotyping errors or other related causes.

Analysis of genes included in the QTL regions associated with hypoxia tolerance suggested co-location in the genome and coordinated functions. Although it is not possible to delineate the candidate genes because of the relatively large numbers of genes existing in the identified QTL regions, analysis of the genes and their known functions are still important. First, such analysis can provide a pool of genes for the determination of potential candidate genes for hypoxia tolerance. To not miss any potentially important genes, we included all genes within the ±0.5 Mb genomic regions of the most significant SNPs (Figure 9 and Figure 10), and the closest genes surrounding the suggestive SNPs (Table 8) in both across-strain analyses and within-strain analyses. With channel catfish, 0.5 Mb physical distance correspond to, on average, almost 3 cM of genetic distance (Liu et al., 2016). Second, by examination of genes within these genomic regions and their known functions, insights can be gained as to if these genes have known functions related to hypoxia responses, and/or what hypoxia related pathways they are involved.

Genes surrounding the most significant SNP on LG6 for hypoxia tolerance across strains were identified using the catfish reference genome (Figure 9). Most of these candidate genes were involved in MAPK or PI3K/AKT/mTOR signaling pathways; and these pathways are known to be essential for hypoxia-mediated angiogenesis, cell proliferation, apoptosis and survival (Emerling et al., 2005). For instance, the most significant SNP (AX-86110011) was located in *lrrc1* gene on LG6. Gene *LRRC1*, encoding the cell polarity regulator, shares a high level of sequence and function similarity with *C/EBP* (Descombes et al., 1990). *C/EBP* plays a critical role in regulating cellular response to stress and is regulated by p38 MAPK (Huggins et al., 2015). Gene *tceb3*/*elongin-a* was located immediately upstream of *lrrc1* in the catfish genome. Knockout of *TCEB3* was reported to induce apoptosis and cellular senescence through activation of the p38 MAPK pathway and the hypoxia-response genes (Miyata et al., 2007). *GCLC* was also reported to

downregulate in hypoxic wild-type mice when compared with normoxic wild-type mice (Eba et al., 2013; Hoshikawa et al., 2001). *MLIP* is required for precocious cardiac adaptation to stress in mice by impacting cardiac activity of Akt/mTOR pathways, which may play a similar role for cardiac adaption to hypoxic stress in catfish (Cattin et al., 2015). Inhibition of *FAM83B*, a key regulator of RAF/MAPK signaling pathway, can decrease AKT phosphorylation by altering the subcellular location of multiple PI3K signaling components (Cipriano et al., 2014). *FGFR2* transduces *FGF* signals to PI3K-Akt signaling cascade, which is involved in cell survival and polarity control (Katoh, 2009). *PLPP4*, preferentially expressed in endothelial cells, may play a role in angiogenesis (Takeuchi et al., 2007). *COX20* acts as the last enzyme in the respiratory electron transport chain of mitochondria (Bourens et al., 2014). *FBXO9* can minimize energy-consuming and procure survival to avoid early cell death by directly regulating mTOR signaling pathway (Fernandez-Saiz et al., 2013). *BMP5* can induce apoptosis through activation of p38 MAPK pathway (Zuzarte-Luıs et al., 2004). *BAG2* is involved in the regulation of oxidative phosphorylation and energy metabolism; and is necessary for the assembly of mitochondrial respiratory supercomplexes (Ueda et al., 2004). Some genes neighboring the most significant SNP on LG25 for Kansas strain were also found to be contributed to hypoxia tolerance (Figure 10). For instance, *NF1*, as a tumor suppressor gene, may involve in hypoxia responses mediated by HIF-1α (Opocher et al., 2005). *LGALS9* and *VEZF1* were reported to involve in the regulation of angiogenesis (Miyashita et al., 2004; Thijssen and Griffioen, 2014). *SLC1A3*, *GDNF* and *DHRS13* were found to be upregulated in response to hypoxia (Hu et al., 2014; Wang et al., 2010; Yamagata et al., 2002). *UCP2* can inhibit ROS-mediated apoptosis under hypoxic conditions (Deng et al., 2012). The co-localization of these functional related genes for hypoxia tolerance also supported our hypothesis of "functional hubs" within the genomes (Geng et al., 2015).

In addition to the analysis of genes within the significant QTL regions, examination of the genes surrounding the suggestive SNPs may provide some insights as to if such genomic regions would include real QTL, upon increases of sample sizes. As listed in Table 8, some genes close to the suggestively associated SNPs on different linkage groups appeared to have known functions related to hypoxia tolerance. With across-strain QTL, for instance, *cox7a2l* gene is located 17 kb upstream of the suggestive SNP marker AX-85410024 on LG6. *COX7A2L* is involved in the regulation of energy metabolism and responsible for the assembly of mitochondrial respiratory supercomplexes (Ikeda et al., 2013). This gene was reported to be significantly differentially expressed under hypoxia as a cardiac gene in *Fundulus grandis* (Everett et al., 2012). On LG5, *btbd17* gene is located 1 Kb upstream, and *map2k6/mkk6* of is located 21 Kb downstream of the suggestive SNP marker AX-85216281. While the function of *BTBD17* remains unknown, *MAP2K6/MKK6* is an essential component of the MAPK pathway that mediates apoptotic cell death in thymocytes (Raingeaud et al., 1996). Cells deficient in *MKK6* failed to activate p38 MAPK and stabilize *HIF-1* during hypoxia, indicating that the hypoxic activation of p38 MAPK and *HIF-1* is in a *MKK6*-dependent manner (Emerling et al., 2005). On LG10, the suggestively associated SNP marker AX-85217471 is located within the *itsn1* gene. *ITSN1* codes a long (*ITSN1-L*) and a short (*ITSN1-S*) protein isoform, and hypoxia specifically upregulates the long isoform *ITSN1-L* (Weigand et al., 2012). *ITSN* also regulates PI3K-C2beta and AKT signaling pathway necessary for cell survival (Das et al., 2007). Additionally *ITSN1* was reported to negatively regulate the mitochondrial apoptotic pathway in endothelial cells (Predescu et al., 2007). Similarly, on LG12, the suggestively associated marker AX-85197805 is within the *cic* gene. *CIC* acts as an important regulator of the homeostatic control of mitochondria in tumor tissues, and high expression level of *CIC* in tumors allows adaptation for metabolic and respiration stress (Catalina-

Rodriguez et al., 2012). With within-strain QTL, for instance, *rap2b*, located in SNP AX-85237540 on LG4 for Marion strain, can be induced by various stresses and regulate pro-survival function in a p53-dependent manner (Zhang et al., 2013). Gene *bad*, located in SNP AX-85325943 on LG13 for Marion Select strain, was reported to induce apoptosis in cells by EGFR/MAPK and PI3K/Akt kinase pathways (She et al., 2005). Gene *aplp1*, located in SNP AX-85339135 on LG22 for Marion Select strain, is required for the proliferation in epithelial and fibroblastic cell types (Tang et al., 2007). Taken together, there are reasons to believe that the suggestive QTL deserve to be further explored. It is apparent that use of larger sample sizes may make these suggestive QTL statistically significant.

A number of gene pathways are known to be involved in hypoxia responses including HIF, NOTCH, MAPK and PI3K pathways. Here, the vast majority of the genes within QTL identified in this GWAS are involved in MAPK pathway or/and PI3K/AKT/mTOR pathway. In particular, p38 MAPK signaling, known as stress-activated protein kinase pathway, is essential for *HIF1* activation under hypoxic conditions (Emerling et al., 2005). Similar to MAPK pathway, constitutively activated PI3K/AKT involved in hypoxic activation of *HIF-1α*, while the activity of *mTOR1* is suppressed under hypoxic conditions in cancer cells (Cam et al., 2010). In our study, nine genes were involved in the MAPK pathway including *tceb3*, *klhl31*, *fam83b*, *hcrtr2*, *bmp5*, *bag2*, *map2k6*, *ucp2* and *bad*; seven genes were involved in PI3K/AKT/mTOR pathway including *elovl5*, *mlip*, *fam83b*, *fgfr2*, *fbxo9*, *itsn* and *bad*. All these genes are known to be involved in MAPK or PI3K/AKT/mTOR pathway through different mediators to regulate cell response to extracellular stress and maintain cell homeostasis for organisms' survival. Taken together, the presence of these genes within the detected QTL regions may suggest that the allelic variations within these genes could be potentially the causes of the observed phenotypic variations of hypoxia tolerance. Future

studies of the sequence variations in contrasted haplotypes of these QTL regions should elucidate the molecular basis of the detected QTL for hypoxia tolerance.

In summary, a large number of significant and suggestive QTL were identified for hypoxia tolerance in channel catfish. Within-strain QTL explained relatively large proportions (>20%) of phenotypic variations, while across-strain QTL had a relatively small effect on phenotypic variations (4-6%) for hypoxia tolerance. Many genes surrounding the identified QTL are known to be functionally related to cell adaption and response to hypoxic stress, and they are mostly involved in MAPK or/and PI3K/AKT/mTOR pathways. The fact that many QTL were detected both across strains and within strains, and that very few overlaps of the QTL among strains suggested a highly complex genetic architecture for hypoxia tolerance. If validated, these findings would suggest that marker-assisted selection for hypoxia tolerance would be essentially very difficult, while traditional and/or genome-based selection may have to be adopted for genetic enhancement programs in catfish.

References

Aboagye, D.L., Allen, P.J., 2014. Metabolic and locomotor responses of juvenile paddlefish *Polyodon spathula* to hypoxia and temperature. Comparative biochemistry and physiology. A: comparative physiology 169, 51-59.

Aggeli, I.K., Gaitanaki, C., Lazou, A., Beis, I., 2002. Hyperosmotic and thermal stresses activate p38-MAPK in the perfused amphibian heart. Journal of experimental biology 205, 443-454.

Anttila, K., Dhillon, R.S., Boulding, E.G., Farrell, A.P., Glebe, B.D., Elliott, J.A., Wolters, W.R., Schulte, P.M., 2013. Variation in temperature tolerance among families of Atlantic salmon (*Salmo salar*) is associated with hypoxia tolerance, ventricle size and myoglobin level. Journal of experimental biology 216, 1183-1190.

Beall, C.M., 2000. Tibetan and Andean patterns of adaptation to high-altitude hypoxia. Human biology, 72, 201-228.

Bigham, A., Bauchet, M., Pinto, D., Mao, X., Akey, J.M., Mei, R., Scherer, S.W., Julian, C.G., Wilson, M.J., Herráez, D.L., 2010. Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data. PLoS genetics 6, e1001116.

Bourens, M., Boulet, A., Leary, S.C., Barrientos, A., 2014. Human COX20 cooperates with SCO1 and SCO2 to mature COX2 and promote the assembly of cytochrome c oxidase. Human molecular genetics 23, 2901-2913.

Boyd, C.E., 1982. Water quality management for pond fish culture. Elsevier Scientific Publishing Co.

Cam, H., Easton, J.B., High, A., Houghton, P.J., 2010. mTORC1 signaling under hypoxic conditions is controlled by ATM-dependent phosphorylation of HIF-1α. Molecular cell 40, 509-520.

Catalina-Rodriguez, O., Kolukula, V.K., Tomita, Y., Preet, A., Palmieri, F., Wellstein, A., Byers, S., Giaccia, A.J., Glasgow, E., Albanese, C., 2012. The mitochondrial citrate transporter, CIC, is essential for mitochondrial homeostasis. Oncotarget 3, 1220-1235.

Cattin, M.E., Wang, J., Weldrick, J.J., Roeske, C.L., Mak, E., Thorn, S.L., DaSilva, J.N., Wang, Y., Lusis, A.J., Burgon, P.G., 2015. Deletion of MLIP (muscle-enriched A-type lamin-interacting protein) leads to cardiac hyperactivation of Akt/mammalian target of rapamycin (mTOR) and impaired cardiac adaptation. The journal of biological chemistry 290, 26699-26714.

Cipriano, R., Bryson, B.L., Miskimen, K.L., Bartel, C.A., Hernandez-Sanchez, W., Bruntz, R.C., Scott, S.A., Lindsley, C.W., Brown, H.A., Jackson, M.W., 2014. Hyperactivation of EGFR and downstream effector phospholipase D1 by oncogenic FAM83B. Oncogene 33, 3298-3306.

Darvasi, A., Soller, M., 1992. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. Theoretical and applied genetics 85, 353-359.

Das, M., Scappini, E., Martin, N.P., Wong, K.A., Dunn, S., Chen, Y.-J., Miller, S.L., Domin, J., O'Bryan, J.P., 2007. Regulation of neuron survival through an intersectin-phosphoinositide 3′-kinase C2β-AKT pathway. Molecular and cellular biology 27, 7906-7917.

Deng, S., Yang, Y., Han, Y., Li, X., Wang, X., Li, X., Zhang, Z., Wang, Y., 2012. UCP2 inhibits ROS-mediated apoptosis in A549 under hypoxic conditions. PloS one 7, e30714.

Derayat, A., Houston, R., Guy, D., Hamilton, A., Ralph, J., Spreckley, N., Taggart, J., McAndrew, B., Haley, C., 2007. Mapping QTL affecting body lipid percentage in Atlantic salmon (*Salmo salar*). Aquaculture 272, S250-S251.

Descombes, P., Chojkier, M., Lichtsteiner, S., Falvey, E., Schibler, U., 1990. LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein. Genes and development 4, 1541-1551.

Duggal, P., Gillanders, E.M., Holmes, T.N., Bailey-Wilson, J.E., 2008. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. BMC genomics 9, 516.

Dunham, R.A., Smitherman, R.O., 1984. Ancestry and breeding of catfish in the United States. Alabama agricultural experimental station circle 273, pp. 100.

Eba, S., Hoshikawa, Y., Moriguchi, T., Mitsuishi, Y., Satoh, H., Ishida, K., Watanabe, T., Shimizu, T., Shimokawa, H., Okada, Y., Yamamoto, M., Kondo, T., 2013. The nuclear factor erythroid 2-related factor 2 activator oltipraz attenuates chronic hypoxia-induced cardiopulmonary alterations in mice. American journal of respiratory cell and molecular biology 49, 324-333.

Emerling, B.M., Platanias, L.C., Black, E., Nebreda, A.R., Davis, R.J., Chandel, N.S., 2005. Mitochondrial reactive oxygen species activation of p38 mitogen-activated protein kinase is required for hypoxia signaling. Molecular and cellular biology 25, 4853-4862.

Everett, M.V., Antal, C.E., Crawford, D.L., 2012. The effect of short-term hypoxic exposure on metabolic gene expression. Journal of experimental zoology. Part A, Ecological genetics and physiology 317, 9-23.

Faust, H.A., Gamperl, A.K., Rodnick, K.J., 2004. All rainbow trout (*Oncorhynchus mykiss*) are not created equal: intra-specific variation in cardiac hypoxia tolerance. Journal of experimental biology 207, 1005-1015.

Fernandez-Saiz, V., Targosz, B.S., Lemeer, S., Eichner, R., Langer, C., Bullinger, L., Reiter, C., Slotta-Huspenina, J., Schroeder, S., Knorn, A.M., Kurutz, J., Peschel, C., Pagano, M., Kuster, B., Bassermann, F., 2013. SCFFbxo9 and CK2 direct the cellular response to growth factor withdrawal via Tel2/Tti1 degradation and promote survival in multiple myeloma. Nature cell biology 15, 72-81.

Geng, X., Feng, J., Liu, S., Wang, Y., Arias, C., Liu, Z., 2014. Transcriptional regulation of hypoxia inducible factors alpha (HIF-α) and their inhibiting factor (FIH-1) of channel catfish (*Ictalurus punctatus*) under hypoxia. Comparative Biochemistry and Physiology. B: Biochemistry and molecular biology 169, 38-50.

Geng, X., Liu, S., Yao, J., Bao, L., Zhang, J., Li, C., Wang, R., Sha, J., Zeng, P., Zhi, D., 2016. A genome wide association study identifies multiple regions associated with head size in catfish. G3: genes, genomes, genetics 10, 3389-3398.

Geng, X., Sha, J., Liu, S., Bao, L., Zhang, J., Wang, R., Yao, J., Li, C., Feng, J., Sun, F., 2015. A genome-wide association study in catfish reveals the presence of functional hubs of related genes within QTL for columnaris disease resistance. BMC genomics 16, 1.

Goel, M.K., Khanna, P., Kishore, J., 2010. Understanding survival analysis: Kaplan-Meier estimate. International journal of ayurveda research 1, 274-278.

Guan, B., Ma, H., Wang, Y., Hu, Y., Lin, Z., Zhu, Z., Hu, W., 2011. Vitreoscilla hemoglobin (VHb) overexpression increases hypoxia tolerance in zebrafish (*Danio rerio*). Marine Biotechnology 13, 336-344.

Hoshikawa, Y., Ono, S., Suzuki, S., Tanita, T., Chida, M., Song, C., Noda, M., Tabata, T., Voelkel, N.F., Fujimura, S., 2001. Generation of oxidative stress contributes to the development of pulmonary hypertension induced by hypoxia. Journal of applied physiology 90, 1299-1306.

Houston, R.D., Haley, C.S., Hamilton, A., Guy, D.R., Mota-Velasco, J.C., Gheyas, A.A., Tinch, A.E., Taggart, J., Bron, J., Starkey, W., 2010. The susceptibility of Atlantic salmon fry to freshwater infectious pancreatic necrosis is largely explained by a major QTL. Heredity 105, 318-327.

Hu, H.X., Takano, N., Xiang, L.S., Gilkes, D.M., Luo, W.B., Semenza, G.L., 2014. Hypoxia-inducible factors enhance glutamate signaling in cancer cells. Oncotarget 5, 8853-8868.

Huggins, C.J., Mayekar, M.K., Martin, N., Saylor, K.L., Gonit, M., Jailwala, P., Kasoji, M., Haines, D.C., Quiñones, O.A., Johnson, P.F., 2015. C/EBPγ is a critical regulator of cellular stress response networks through heterodimerization with ATF4. Molecular and cellular biology 36, 693-713.

Ikeda, K., Shiba, S., Horie-Inoue, K., Shimokata, K., Inoue, S., 2013. A stabilizing factor for mitochondrial respiratory supercomplex assembly regulates energy metabolism in muscle. Nature communications 4, 2147.

Jha, A.R., Zhou, D., Brown, C.D., Kreitman, M., Haddad, G.G., White, K.P., 2016. Shared genetic signals of hypoxia adaptation in *Drosophila* and in high-altitude human populations. Molecular biology and evolution 33, 501-517.

Jin, C., Lan, H., Attie, A.D., Churchill, G.A., Bulutuglo, D., Yandell, B.S., 2004. Selective phenotyping for increased efficiency in genetic mapping studies. Genetics 168, 2285-2293.

Katoh, M., 2009. FGFR2 abnormalities underlie a spectrum of bone, skin, and cancer pathologies. Journal of investigative dermatology 129, 1861-1867.

Kietzmann, T., Mennerich, D., Dimova, E.Y., 2016. Hypoxia-inducible factors (HIFs) and phosphorylation: impact on stability, localization, and transactivity. Frontiers in cell and developmental biology 4, 11.

Kvamme, B.O., Gadan, K., Finne-Fridell, F., Niklasson, L., Sundh, H., Sundell, K., Taranger, G.L., Evensen, O., 2013. Modulation of innate immune responses in Atlantic salmon by chronic hypoxia-induced stress. Fish and shellfish immunology 34, 55-65.

Li, Y., Liu, S., Qin, Z., Waldbieser, G., Wang, R., Sun, L., Bao, L., Danzmann, R.G., Dunham, R., Liu, Z., 2015. Construction of a high-density, high-resolution genetic map and its integration with BAC-based physical map in channel catfish. DNA Research 22, 39-52.

Li, Y., Padmanabha, D., Gentile, L.B., Dumur, C.I., Beckstead, R.B., Baker, K.D., 2013. HIF-and non-HIF-regulated hypoxic responses require the estrogen-related receptor in *Drosophila* melanogaster. PLoS genetics 9, e1003230.

Liu, S., Sun, L., Li, Y., Sun, F., Jiang, Y., Zhang, Y., Zhang, J., Feng, J., Kaltenboeck, L., Kucuktas, H., 2014. Development of the catfish 250K SNP array for genome-wide association studies. BMC research notes 7, 135.

Liu, Z.J., Liu, S.K., Yao, J., Bao, L.S., Zhang, J.R., Li, Y., Jiang, C., Sun, L.Y., Wang, R.J., Zhang, Y., Zhou, T., Zeng, Q.F., Fu, Q., Gao, S., Li, N., Koren, S., Jiang, Y.L., Zimin, A., Xu, P., Phillippy, A.M., Geng, X., Song, L., Sun, F.Y., Li, C., Wang, X.Z., Chen, A.L., Jin, Y.L., Yuan, Z.H., Yang, Y.J., Tan, S.X., Peatman, E., Lu, J.G., Qin, Z.K., Dunham, R., Li, Z.X., Sonstegard, T., Feng, J.B., Danzmann, R.G., Schroeder, S., Scheffler, B., Duke, M.V., Ballard, L., Kucuktas, H., Kaltenboeck, L., Liu, H.X., Armbruster, J., Xie, Y.J., Kirby, M.L., Tian, Y., Flanagan, M.E., Mu, W.J., Waldbieser, G.C., 2016. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. Nature communications 7, 11757.

McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P.A., Hirschhorn, J.N., 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nature reviews genetics 9, 356-369.

Miyashita, H., Kanemura, M., Yamazaki, T., Abe, M., Sato, Y., 2004. Vascular endothelial zinc finger 1 is involved in the regulation of angiogenesis: possible contribution of stathmin/OP18 as a downstream target gene. Arteriosclerosis, thrombosis, and vascular biology 24, 878-884.

Miyata, K., Yasukawa, T., Fukuda, M., Takeuchi, T., Yamazaki, K., Sakumi, K., Tamamori-Adachi, M., Ohnishi, Y., Ohtsuki, Y., Nakabeppu, Y., Kitajima, S., Onishi, S., Aso, T., 2007. Induction of apoptosis and cellular senescence in mice lacking transcription elongation factor, Elongin A. Cell death and differentiation 14, 716-726.

Mizukami, Y., Kohgo, Y., Chung, D.C., 2007. Hypoxia inducible factor-1–independent pathways in tumor angiogenesis. Clinical cancer research 13, 5670-5674.

Moen, T., Baranski, M., Sonesson, A.K., Kjøglum, S., 2009. Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*): population-level associations between markers and trait. BMC genomics 10, 368.

Moghadam, H.K., Poissant, J., Fotherby, H., Haidle, L., Ferguson, M.M., Danzmann, R.G., 2007. Quantitative trait loci for body weight, condition factor and age at sexual maturation in Arctic charr (*Salvelinus alpinus*): comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Atlantic salmon (*Salmo salar*). Molecular genetics and genomics 277, 647-661.

Nei, M., 1973. Analysis of gene diversity in subdivided populations. Proceedings of the national academy of sciences 70, 3321-3323.

Ogut, F., Bian, Y., Bradbury, P.J., Holland, J.B., 2015. Joint-multiple family linkage analysis predicts within-family variation better than single-family analysis of the maize nested association mapping population. Heredity 114, 552-563.

Opocher, G., Conton, P., Schiavi, F., Macino, B., Mantero, F., 2005. Pheochromocytoma in von Hippel-Lindau disease and neurofibromatosis type 1. Familial cancer 4, 13-16.

Predescu, S.A., Predescu, D.N., Knezevic, I., Klein, I.K., Malik, A.B., 2007. Intersectin-1s regulates the mitochondrial apoptotic pathway in endothelial cells. Journal of biological chemistry 282, 17166-17178.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., Reich, D., 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nature genetics 38, 904-909.

Qiu, F.Z., Zheng, Y.L., Zhang, Z.L., Xu, S.Z., 2007. Mapping of QTL associated with waterlogging tolerance during the seedling stage in maize. Annals of botany 99, 1067-1081.

Raingeaud, J., Whitmarsh, A.J., Barrett, T., Derijard, B., Davis, R.J., 1996. MKK3-and MKK6-regulated gene expression is mediated by the p38 mitogen-activated protein kinase signal transduction pathway. Molecular and cellular biology 16, 1247-1255.

Semenza, G.L., 2000. HIF-1: mediator of physiological and pathophysiological responses to hypoxia. Journal of applied physiology 88, 1474-1480.

Semenza, G.L., Agani, F., Feldser, D., Iyer, N., Kotch, L., Laughner, E., Yu, A., 2000. Hypoxia, HIF-1, and the pathophysiology of common human diseases. Advances in experimental medicine and biology 475, 123-130.

She, Q.-B., Solit, D.B., Ye, Q., O'Reilly, K.E., Lobo, J., Rosen, N., 2005. The BAD protein integrates survival signaling by EGFR/MAPK and PI3K/Akt kinase pathways in PTEN-deficient tumor cells. Cancer cell 8, 287-297.

Stranger, B.E., Stahl, E.A., Raj, T., 2011. Progress and promise of genome-wide association studies for human complex trait genetics. Genetics 187, 367-383.

Sun, L., Liu, S., Wang, R., Jiang, Y., Zhang, Y., Zhang, J., Bao, L., Kaltenboeck, L., Dunham, R., Waldbieser, G., 2014. Identification and analysis of genome-wide SNPs provide insight into signatures of selection and domestication in channel catfish (*Ictalurus punctatus*). PloS one 9, e109666.

Takeuchi, M., Harigai, M., Momohara, S., Ball, E., Abe, J., Furuichi, K., Kamatani, N., 2007. Cloning and characterization of DPPL1 and DPPL2, representatives of a novel type of mammalian phosphatidate phosphatase. Gene 399, 174-180.

Tang, X., Milyavsky, M., Goldfinger, N., Rotter, V., 2007. Amyloid-β precursor-like protein APLP1 is a novel p53 transcriptional target gene that augments neuroblastoma cell death upon genotoxic stress. Oncogene 26, 7302-7312.

Thijssen, V.L., Griffioen, A.W., 2014. Galectin-1 and-9 in angiogenesis: A sweet couple. Glycobiology 24, 915-920.

Ueda, K., Kosako, H., Fukui, Y., Hattori, S., 2004. Proteomic identification of Bcl2-associated athanogene 2 as a novel MAPK-activated protein kinase 2 substrate. Journal of biological chemistry 279, 41815-41821.

Wang, D., Sun, Y., Stang, P., Berlin, J.A., Wilcox, M.A., Li, Q., 2009. Comparison of methods for correcting population stratification in a genome-wide association study of rheumatoid arthritis: principal-component analysis versus multidimensional scaling. BMC proceedings 7, 109.

Wang, S.N., Zhou, Y.F., Seavey, C.N., Singh, A.K., Xu, X.L., Hunt, T., Hoyt, R.F., Horvath, K.A., 2010. Rapid and dynamic alterations of gene expression profiles of adult porcine bone marrow-derived stem cell in response to hypoxia. Stem cell research 4, 117-128.

Weigand, J.E., Boeckel, J.N., Gellert, P., Dimmeler, S., 2012. Hypoxia-induced alternative splicing in endothelial cells. PLoS one 7, e42697.

Weir, B.S., Ott, J., 1997. Genetic data analysis II. Trends in genetics 13, 379.

Welker, T.L., Mcnulty, S.T., Klesius, P.H., 2007. Effect of sublethal hypoxia on the immune response and susceptibility of channel catfish, *Ictalurus punctatus*, to enteric septicemia. Journal of the world aquaculture society 38, 12-23.

Wringe, B.F., Devlin, R.H., Ferguson, M.M., Moghadam, H.K., Sakhrani, D., Danzmann, R.G., 2010. Growth-related quantitative trait loci in domestic and wild rainbow trout (*Oncorhynchus mykiss*). BMC genetics 11, 63.

Xia, Y., Choi, H.-K., Lee, K., 2012. Recent advances in hypoxia-inducible factor (HIF)-1 inhibitors. European journal of medicinal chemistry 49, 24-40.

Xu, P., Wang, S., Liu, L., Thorsen, J., Kucuktas, H., Liu, Z., 2007. A BAC-based physical map of the channel catfish genome. Genomics 90, 380-388.

Yamagata, K., Tagami, M., Ikeda, K., Tsumagari, S., Yamori, Y., Nara, Y., 2002. Differential regulation of glial cell line-derived neurotrophic factor (GDNF) mRNA expression during hypoxia and reoxygenation in astrocytes isolated from stroke-prone spontaneously hypertensive rats. Glia 37, 1-7.

Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Cuo, Z.X.P., Pool, J.E., Xu, X., Jiang, H., Vinckenbosch, N., Korneliussen, T.S., 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. Science 329, 75-78.

Zhang, X., He, Y., Lee, K.-H., Dubois, W., Li, Z., Wu, X., Kovalchuk, A., Zhang, W., Huang, J., 2013. Rap2b, a novel p53 target, regulates p53-mediated pro-survival function. Cell cycle 12, 1279-1291.

Zuzarte-Luıs, V., Montero, J., Rodriguez-Leon, J., Merino, R., Rodrıguez-Rey, J., Hurle, J., 2004. A new role for BMP5 during limb development acting through the synergic activation of Smad and MAPK pathways. Developmental biology 272, 39-52.